Salinity adaptation in Tunisian and Portuguese *Medicago truncatula* populations

Matilde de Vasconcelos Manso Ataíde Cordeiro

Dissertation presented to obtain the Ph.D degree in Biology

Instituto de Tecnologia Química e Biológica António Xavier | Universidade Nova de Lisboa

Oeiras, June, 2016

Work performed at:

Laboratório de Biotecnologia de Células Vegetais
Instituto de Tecnologia Química e Biológica António Xavier
Universidade Nova de Lisboa
Av. da República, Quinta do Marques
2780-157 Oeiras, Portugal

Cook Lab – Plants, Microbes & Genes
University of California, Davis
Department of Plant Pathology
One Shields Ave., Robbins Hall
Davis, CA 95616, USA

Matilde Ataíde Cordeiro funded by:

Fundação para a Ciência e Tecnologia SFRH/BD/39905/2007

Ph.D supervisors

Professor Pedro Fevereiro

Head of laboratory, Laboratório de Biotecnologia de Células Vegetais, Instituto de Tecnologia Química e Biológica António Xavier, Universidade Nova de Lisboa.
Assistant Professor, Departamento de Biologia Vegetal, Faculdade de Ciências da Universidade de Lisboa.

Professor Douglas Cook

Professor, Department of Plant Pathology, University of California, Davis.
Director, Feed the Future Innovation Lab Climate Resilient Chickpea, University of California, Davis.

Table of contents

Acknowledgements

The work presented in this dissertation would have been impossible without the help, contribution, support and inspiration from several people that I would like to acknowledge:

Pedro Fevereiro, my supervisor, who gave me the opportunity to engage with science while I was still a teenager and to work in his lab throughout the years. Pedro motivated me to develop a scientific mind and gave me the chance to pursue this PhD, offering me support and independence.

Douglas Cook, my co-supervisor, who welcomed me in his lab and helped me grow as a scientist and as a person. Doug encouraged me to pursue and lead my own projects, while providing guidance and directing me towards the “big picture” broadening my scientific vision. More personally, Doug included me in his “extended family” and made it much easier to move alone to the other “corner” of the World.

Mafalda Lapa, who first introduced me to science and captured my interest with her unique energy and sharpness.

Varma Penmetsa, with whom troubleshooting always felt like a moment of leisure that constantly led me to new ideas and motivations.

Sergey Nuzhdin, who welcomed me into his lab and projects, offering me an “open road” of opportunities.

Eric von Wettberg, who initiated me to evolutionary ecology and guided me as I moved towards integrating molecular and ecological tools.

Maren Friesen, who welcomed and cultured me into ecological genomics with a distinctive enthusiasm.

Sharon Strauss, who steered me through ecology and ecological methods naturally and confidently.

My labmates and friends from the Plant Cell Biotechnology Lab:

Sofia Duque and Susana Araujo, who where the first to guide me through the process of lab work and gave me the chance to pursue my own project from beginning to end.

Ines Trindade, a friend in and out of the lab, without whom I don't think I would have been able to finish this thesis. There are no words to describe the support and understanding you gave me throughout this process.

Mara Alves, a friend who has accompanied me through science and laughter.

Zé Ricardo, Maria João, Susana Leitão, Nuno Almeida e Cátia Nunes, for walking into my life and making the lab such a nice place to be.

My labmates and friends from the Cook lab:

Noelia Carrasquilla, who received me in her life and in your house like family. Together with Estibaliz Larrainzar and Mariana Pereira, my friends and confidents in the lab that have always been there for whatever necessary.

Ken Moriuchi, who engaged me in data analysis, field work and evolutionary ecology rationales. But my fence is better!

Brendan Riely, for constant input and for making me company while the ghosts visited the lab.

Kelsey Miller, Lisa Vance, Hanna Yokohama-Hatch and Emily Bergmann, for being such good help, taking care of my plants and being so awesome!

My labmates and friends from the Nuzhdin Lab:

Wendy Vu, who accompanied me throughout the several field experiments and became such a special friend. You are the sweetest person and gave me more motivation to finish this thesis than you imagine.

Peter Chang and Vasantika Suryawanshi, my "plant group" friends who gave me insight and support.

But a lot of people that are not necessarily involved in science were key to give me moral and emotional support throughout the years. I'm not going to name all of you, but some I have to mention:

My friends and housemates that dealt with my mood swings as multitasking got harder and made me have a family to return to every day since I moved to Los Angeles. Brad Foley, as a friend and scientist, you've been so inspiring and such a good listener. Eliseo Melero, you are the best and you can still surprise me with your kindness and friendship. Tugba, Ulli, Thomas, Cynthia and Chad, you make my days so much easier and more fun.

My friends, older and newer, that will always have a special place and, mostly at a distance, helped me find the focus to finish this dissertation: Nuria Cotter, Paula Minas, Sofia Cardoso, Nuno Brás, Rita Gaspar, João Alho, Alexandra Almeida, Marta Ágoas, Marta Francisco, Barbara Gandolfi.

My family, who always supported me and gave me the tools and motivation needed to pursue and finalize this work. My parents, who gave me unconditional support and patiently listened to me through the good and the bad phases. My brother and sister, Fernando Cordeiro e Roberta Lucotti, and my nieces, Catarina and Francesca, who I wish I could have closer, but that make my path look clearer.

Ricardo Lopes, que me trazes calma durante a tempestade.

General abstract

Worldwide, salinity is a major constraint for plant growth and reproduction of both crop and wild species. Soil salinization is expected to aggravate, particularly in a scenario of climate change. There is a need to implement sustainable, long-term measurements to guarantee the global agricultural production levels necessary for food security. The study of wild populations that are viable and fertile in saline habitats has the potential to reveal adaptive mechanisms bringing new insights for crop improvement. Here, methods of molecular biology and ecology are used to investigate the genes and traits responsible for salinity adaptation in natural populations of the model legume *Medicago truncatula*.

Tunisian and Portuguese populations of *M. truncatula* with origin in saline and non-saline habitats were used to: i) study local adaptation; ii) identify the mechanisms of salinity adaptation evolved in these populations; iii) test whether the same or different mechanisms have evolved in populations from different countries; iv) identify candidate genes and traits underlying adaptive responses; and v) test evolutionary hypotheses regarding salinity adaptation.

Tunisian *M. truncatula* populations from environments contrasting in soil salinity were genotyped using existing Affymetrix microarrays. Patterns of polymorphism and correlation between adjacent molecular markers were calculated based on these genome-scale polymorphism scans. Single feature polymorphisms (SFP) were identified and linkage disequilibrium (LD) blocks containing 18 genomic regions that assort with habitat were defined, as they contain candidate genes and allelic variants for local adaptation. Salinity associated genetic differentiation makes Tunisian populations a good subject to study local adaptation and pursue the genetic and phenotypic basis of adaptive responses to salinity. Parallel evolutionary ecology studies revealed that both tolerating and avoiding excess salt are important strategies in the evolution of salinity adaptation in these populations.

Germination and initial root expansion are the first developmental stages exposed to salinity, and are among the most sensitive. Plate and aeroponical experiments were used to demonstrate differentiation of salinity adapted Tunisian populations for early seedling growth depending on salt (NaCl and KCl) and abscisic acid (ABA). Saline origin genotypes were less affected on germination and seedling traits, but revealed to be less sensitive to NaCl and more sensitive to ABA relative to non-saline origin genotypes, suggesting differential regulation. Also, parental exposure to salinity had negative effects on germination. Therefore, early developmental stages seem to be important for salinity adaptation.

Plants from *M. truncatula* populations were collected in saline and non-saline environments in Portugal. These plants were genotyped and used in a reciprocal transplant field experiment in Portugal together with an informative subset of Tunisian genotypes to test whether salinity responses involve the same or different mechanisms in distinct populations. Not all performance patterns could be explained by salinity *per se.* But populations from both countries showed signal for salinity adaptation at one of the locations and the same traits were under selection. The lack of common candidate genes between Portuguese and Tunisian populations suggests that they have evolved independently and, in spite of the signal for parallel adaptive evolution, selection acted on different genetic components.

Such evolutionary studies may lead to the identification of novel mechanisms of salinity adaptation, allowing its association with other environmental factors. Therefore, such studies constitute a more integrated approach to guide crop improvement.

Resumo geral (*in Portuguese*)

Mundialmente, a salinidade é um importante factor limitante para o crescimento e reprodução de espécies cultivadas e selvagens. É esperado o agravamento da salinização dos solos, particularmente num cenário de alterações climáticas. É necessária a implementação de medidas sustentáveis e a longo prazo para garantir os níveis de produção agrícola mundiais necessários à segurança alimentar. Estudar populações selvagens que são viáveis e férteis em habitats salinos tem o potencial de revelar mecanismos de adaptação, trazendo novas perspectivas para o melhoramento vegetal. Aqui são utilizados métodos de biologia molecular e de ecologia para revelar os genes e características fenotípicas responsáveis pela adaptação à salinidade da leguminosa modelo *Medicago truncatula.*

Populações Tunisinas e Portuguesas de *M. truncatula* com origem em habitats salinos e não salinos foram utilizadas para: i) estudar a adaptação local; ii) identificar os mecanismos de adaptação que evoluíram nestas populações; iii) testar se os mesmos ou distintos mecanismos evoluíram em populações de países diferentes; iv) identificar os genes e características fenotípicas responsáveis pelas respostas adaptativas; e v) testar hipóteses evolucionárias relacionadas com a adaptação à salinidade.

Populações Tunisinas de *M. truncatula* provenientes de solos com níveis contrastantes de salinidade foram genotipados com "microarrays" Affymetrix pré-existentes. Foram calculados os padrões de polimorfismo e as correlações entre marcadores moleculares adjacentes, com base na análise de polimorfismos à escala do genoma. Foram identificados polimorfismos de carácter único (SFP), e foram definidos blocos em desequilíbrio de ligação (LD) contendo 18 regiões do genoma concordantes com o habitat, os quais contêm genes e variantes alélicas candidatos para adaptação local. A diferenciação genética associada à salinidade faz das populações Tunisinas um objecto válido para o estudo de adaptação local e procura das bases genéticas e fenotípicas de resposta à salinidade. Estudos paralelos de ecologia evolutiva revelaram que tanto tolerar como evitar o excesso

de sal são estratégias importantes na evolução de adaptação à salinidade nestas populações.

As primeiras fases do desenvolvimento vegetal expostas à salinidade são a germinação e a expansão inicial da raiz estando estas fases entre as mais sensíveis. Foram realizadas experiências em placa e em aeroponia para demonstrar a diferenciação de populações Tunisinas adaptadas à salinidade. Foi avaliado o crescimento inicial de plântulas dependendo da presença de sal (NaCl e KCl) e de ácido abscísico (ABA). Genótipos de origem salina foram menos afectados na germinação e crescimento radicular, revelando-se menos sensíveis ao NaCl e mais sensíveis ao ABA relativamente aos genótipos de origem não salina, o que sugere uma regulação diferencial. Adicionalmente, a exposição parental à salinidade teve efeitos negativos na germinação. Assim, fases iniciais do desenvolvimento parecem ser importantes para a adaptação a salinidade.

Plantas de populações portuguesas de *M. truncatula* foram colectadas em ambientes salinos e não salinos. As plantas foram genotipadas e utilizadas numa experiência de transplante recíproco em Portugal, juntamente com um subconjunto testemunha de genótipos Tunisinos, de forma a testar se as respostas à salinidade envolvem os mesmos mecanismos em populações distintas. Nem todos os padrões de performance puderam ser explicados pela salinidade. Mas populações de ambos os países mostraram sinais de adaptação à salinidade numa das localidades e foram encontrados os mesmos caracteres sob seleção. A falta de genes candidatos comuns entre populações Tunisinas e Portuguesas sugere que estas evoluíram independentemente e que, apesar do sinal de evolução adaptativa paralela, a seleção agiu em diferentes componentes genéticas.

Estes estudos evolutivos podem levar à identificação de novos mecanismos de adaptação à salinidade, permitindo a sua associação com outros factores ambientais. Assim, estes estudos constituem uma abordagem mais integrada para guiar o melhoramento de culturas.

Chapter 1.

General Introduction

Chapter 1. Table of contents

Evolution and adaptation

Darwin's theory of evolution by natural selection (1859) anticipates that individuals expressing the most favorable traits in a particular environment will survive and reproduce more than other individuals from the same species in that environment. Consequently, the genes and alleles that underlie the expression of the traits favored in that environment will increase in frequency in the population over multiple generations (Barton et al. 2007).

Processes such as mutation, migration, genetic drift and natural selection can all directly affect allele frequencies in a population (Orr 2005; Barton et al. 2007). Mutations, here defined as heritable changes in the genetic sequence, can have several origins, including: i) single nucleotide substitutions that when occur in coding regions may represent synonymous or nonsynonymous mutations depending on whether they code for the same or different aminoacids, respectively); ii) deletions or insertions, which are generally called indels, that when occur in coding regions may cause frame-shifts; iii) inversion or translocation that consist of the reversion or movement of larger segments of the chromosomes, respectively; iv) or the movement of transposable elements that consist of small DNA fragments that can move within the genome. Depending on their effect on fitness, mutations can be classified as neutral, detrimental or beneficial. In genetic terms, migration is usually accompanied by gene flow, i.e., the movement of alleles from one population to another, which will change in frequency depending on reproductive success. Migration commonly increases genetic variation in a population, which may have opposite consequences: can be beneficial when it brings novel favorable traits or increases their frequency in the population; and it can be detrimental when it prevents or hinders adaptation or the evolution of reproductive barriers. Genetic drift encloses the change in allele frequencies in a population due to random chance, and its impact on allelic frequency within a population is expected to be greater the smaller the population.

Natural selection

Natural selection is the process by which populations adapt to the environment, i.e., the process by which genotypes that have favorable phenotypes have greater contribution to future generations within a population (Barton et al. 2007).

Natural selection is a complex process that is key for adaptive evolution, which at a minimum requires phenotypic variation, fitness differences associated with phenotype, and a genetic basis for the phenotypic variation. Because of the different components it encompasses, the consequences of natural selection can be defined depending on which perspective is it approached. In the perspective of how natural selection shifts the distribution of a trait within a population, selection can be defined as stabilizing (mean trait values are favored), directional (increase or decreased trait values are favored) or disruptive (different trait values are favored). On the other hand, approaching selection in terms of changes in genetic diversity, it can be defined as purifying (decreasing genetic variation) or balancing selection (maintaining genetic variation). While dependent on the overall fitness of genotypes within a population, natural selection can act on different stages of the life cycle. Viability selection acts on differential survival, while fecundity acts on differential reproduction (Barton et al. 2007).

For the past several decades, experimental and theoretical studies have been conducted to understand the genetic signatures of selection and demographic processes in populations (Fisher 1930; Orr 2005; Ehrenreich and Purugganan 2009; Vitti et al. 2013). In the current genomics Era, such knowledge could lead to the reconstruction of the evolutionary description of populations from sequence data (Vitti et al. 2013). Currently, the identification of such genetic signatures relies on the association of patterns in the DNA sequence that differ from the expected by neutrality (Ehrenreich and Purugganan 2009; Vitti et al. 2013). For example, Tajima's D and other variations may be used to test for single nucleotide polymorphisms (SNPs) frequency variation in a target sequence and potentially infer distinct signals of selection (Tajima 1989; Ehrenreich and

Purugganan 2009; Vitti et al. 2013). Other tests rely on the evolution of coding sequences, where the ratio between synonymous and non-synonymous substitutions within and among species is calculated and compared with the candidate allele variations to identify deviations from neutrality (Nei and Gojobori 1986; Ehrenreich and Purugganan 2009; Vitti et al. 2013). Currently, the genetic bases of adaptation are still poorly understood as most of the tests used fail to distinguish selection from demography (Luikart et al. 2004; Orr 2005; Ehrenreich and Purugganan 2009; Vitti et al. 2013). More experimental and theoretical research should be integrated to better understand the genomic basis of adaptive evolution.

Local adaptation

Local adaptation is a possible result from the evolution by natural selection within populations, and the expectation is that individuals should perform better in their home environment compared to a foreign environment (Kawecki and Ebert 2004). More specifically, local adaptation is a special case of a genotype by environment interaction – the different response of genotypes depending on the environment – that is detected when the fitness reaction norms of two populations cross when comparing between home and away environments (Kawecki and Ebert 2004; Hereford 2009; Atkins and Travis 2010; Savolainen et al. 2013). Local adaptation may accompany trade-offs, i.e., adaptation to the home environments may carry fitness costs in the away environment (Hereford 2009). But this pattern is not required for populations to exhibit local adaptation. For example, if the beneficial traits and/or alleles in the home environment are neutral in the away environment, then fitness in the new environment may be similar to that of local populations, and therefore the populations might not reveal crossing reaction norms for fitness, but still be locally adapted (Hereford 2009; Savolainen et al. 2013).

Although there are several reports of local adaptation in different species, not all populations evolve local adaptation (e.g., Galloway and Fenster 2000;

Kawecki and Ebert 2004; Hereford 2009; Atkins and Travis 2010; Friesen and von Wettberg 2010; Savolainen et al. 2013). For local adaptation to evolve selection must favor different traits in different environments and populations must contain significant standing genetic variation (Kawecki and Ebert 2004; Hereford 2009).

Demographic factors, such as genetic drift or gene flow, may outpace the strength of selection and hinder local adaptation (Kawecki and Ebert 2004; Hereford 2009; Savolainen et al. 2013). By limiting population's responses to selection, such processes often lead to misconceptions about the mechanisms driving evolution. For example, if there is high gene flow between populations, the homogenizing effect of allele frequencies may overcome selection and prevent populations from reaching their optimal fitness peak (Hereford 2009; Savolainen et al. 2013). Additionally, random events may drive alleles to fixation independently of selection, particularly when dealing with small population sizes (Hereford 2009; Savolainen et al. 2013).

Changing or unpredictable environments are commonly associated with spatial and temporal variation of selection, and may favor the evolution of phenotypic plasticity over local adaptation (Via and Lande 1985; Charmantier et al. 2008; Hereford 2009; Donohue et al. 2010; Savolainen et al. 2013). Phenotypic plasticity is detected when genotypes express different trait values in different environments, and it is adaptive if the expressed phenotypes increase fitness in their respective environments (Dudley and Schmitt 1996; Charmantier et al. 2008). Adaptive phenotypic plasticity is more likely to evolve when the scale of gene flow is larger than the scale of environmental heterogeneity (Via and Lande 1985).

Plant domestication and agriculture

Domestication consists of a complex evolutionary process that involves species interactions, i.e., it is a product of species co-evolution. For thousands of years, humans have been selecting, dispersing and growing animals and plants,

which in turn have evolved artificially selected traits that distinguish them from their wild relatives (Purugganan and Fuller 2009; Milla et al. 2015).

Consequences of plant domestication

Archeological findings help reveal the dynamics involved in plant adaptation to agricultural practices and the associated genetic signatures since the origin of crop species ~13,000 years, particularly in cereals (Purugganan and Fuller 2009). The ability to germinate and grow in disturbed agricultural soil, and the easiness and magnitude of harvest, incorporate important traits that were artificially selected since the beginning of agriculture, and are directly associated with increased seed sizes and reduced seed dispersal in domesticated species (Purugganan and Fuller 2009). Archeological data suggests that such traits were consecutively introduced in crop species over thousands of years, which is apparently contradictory to genomic data that suggests a single event because of the drastic reduction of standing genetic variation (Caicedo et al. 2007; Purugganan and Fuller 2009; Jiao et al. 2012; Milla et al. 2015; Marsden et al. 2016). Taken together, domestication of some species seems to envision sequential strong artificial selection events underlying their evolution.

Nevertheless, typically only a few wild genotypes are chosen in each domestication event resulting in the reduction of standing genetic variation and therefore in an extreme reduction in effective population size. Subsequently, artificial selection and inbreeding further reduce standing genetic variation in domesticated species (Marsden et al. 2016). Cultivated varieties and wild relatives have been sequenced and analyzed in several plant species such as soybean (Lam et al. 2010), rice (Xu et al. 2012), maize (Caicedo et al. 2007; Jiao et al. 2012), and cucumber (Qi et al. 2013) and genetic variation decreased substantially with domestication. While some of the fixed alleles underlie the expression of important traits for cultivation, several other neutral or deleterious mutations also come to fixation (Marsden et al. 2016).

Overall, low genetic diversity within domesticated germplasm is reflected on the extremely reduced adaptive potential of the current cultivated material. Consequently, it is difficult to colonize novel marginal environments in order to increase the global agricultural area with the existing germplasm. Moreover, it is currently hard to maintain yields in traditionally productive environments that are currently more prone to occurrences of biotic and abiotic stress, partially due to climate change and less predictable seasonal environmental shifts.

Agriculture and environmental stress

Food demand has been rising worldwide at such a rate that agricultural production is not able to meet this increase (FAO 2013). During the last decade, global food stocks have been decreasing substantially to account for the productive deficit and keep reaching minimal historical levels (IPCC 2007; FAO 2013). Such deficit in global production is due to several factors including the use of non-sustainable agricultural practices, the misusage of arable land, and the effects of climate change on the environment where crops were traditionally cultivated and high yielding (Challinor et al. 2014). Because of the yield reduction and the increase of demand, the prices are expected to increase up to 60% by 2050 (Nelson et al. 2014). Consequently, poorer and hunger prone regions are expected to be more vulnerable to food insecurity (Wheeler and von Braun 2013).

Predicting the impact of climate change on the environment and quantifying its effect on crop yields has been the object of several recent studies (IPCC 2007; Kates et al. 2012; FAO 2013; Nelson et al. 2014; Rippke et al. 2016). Global change affects several climatic variables aggravating the occurrence of extreme temperatures, reduced water availability and increased salinity, while typical seasonal events that farmers rely on for agricultural production are becoming unpredictable (FAO 2008, 2013). In natural populations, climate change is responsible for the shifting of species' range and biogeography (Etterson & Shaw 2001; Charmantier et al. 2008; Atkins and Travis 2010). In order to maintain global yields, agricultural practices will need to extend to environmental conditions that

are commonly considered marginal for crop production. But the low genetic variation found within cultivated germplasm may limit the extent to which crop varieties can maintain yield. Thus a major goal in agriculture will be to improve crop varieties to broader their agricultural range and, consequently, one corollary is conserved across studies: the urgent need to find novel adaptations, introduce them into crops and make available climate resilient germplasm (Etterson and Shaw 2001; Friesen and von Wettberg 2010; Nelson et al. 2014; Rippke et al. 2016).

Salinity stress

Salinity stress occurs when there are enough soluble salts in the soil that negatively affect plant growth and reproduction (Provin and Pitt 2001; Munns and Tester 2008). Soil salinization has been documented to negatively affect agriculture for more than 6000 years, dating back to ancient Mesopotamian agriculture (Jacobsen and Adams 1958). Archeological records describe salt accumulation in the soil, shifts from the production of wheat to more salt-tolerant barley, and decreases in yield to about one third in five centuries (Jacobsen and Adams 1958). Although several other factors lead to the breakup of Sumerian civilization, soil salinization is believed to have played a significant role (Jacobsen and Adams 1958).

Currently, soil salinity is estimated to affect over 45 Mha of irrigated arable land (FAO 2008, 2013). Salinized areas are expected to increase and salinity aggravate due to the persistence of unsustainable agricultural practices (particularly the overuse of fertilizers and irrigation that increase salt accumulation over time) coupled with climate change (temperature and precipitation and extremes promote salt build up in the soil), making salinity stress a major factor limiting crop yield worldwide (Flowers and Yeo 1995; Provin and Pitt 2001; IPCC 2007; FAO 2008, 2013; Nelson et al. 2014).

In spite of different salts potentially being responsible for soil salinization – such as several combination of sodium, calcium, potassium, magnesium, chlorides, nitrates, sulfates, bicarbonates and carbonates (Provin and Pitt 2001; FAO 2008) – for the sake of simplicity and focus, only NaCl driven salinity is discussed in the present chapter.

Salinity stress affects plant growth-yield

Salinity and drought stress effects on plant growth and yield are very similar. The major reason for this similarity is because osmotic stress is one of the components of salinity. Before having a direct toxic effect on the plant, excess ions in soil reduce water potential in the root zone and therefore salt stress is perceived and impacts plants functions in two phases, osmotic stress and ion toxicity (reviewed by Munns and Tester 2008). Accordingly, plants exhibit a broad range of responses to soil salinity. Drought and salinity are major stresses reducing crop yields worldwide.

The first indicator that salinity negatively influences a plant is osmotic stress. In the earliest stage of development, water flux towards the embryo is reduced resulting in the impediment or delay of germination (Wahid et al. 1999). In later stages of the plant's life cycle, water potential is decreased between the soil and the roots resulting in the reduction of water uptake potentially leading to water deficit (Boursiac et al. 2005). The effects of the osmotic phase are rapidly observed (within hours) inhibiting new growth and restraining the development of young leaves (Munns and Tester 2008).

On the other hand, ionic stress is perceived days, or even weeks, after excess salt is present in the soil. Ion toxicity results from excessive cellular Na^+ and can inhibit K^+ uptake interfering with K^+ dependent metabolic and physiological functions (Hauser and Horie 2010). During germination, ion toxicity can disrupt the metabolism of carbohydrates resulting in delayed seedling growth (Wahid et al. 1999). In later phases of plant development, ion accumulation may lead to a

gradual growth inhibition, premature senescence of older leaves, and maybe even death (Munns and Tester 2008; Hauser and Horie 2010).

Salinity sensing and signaling

Some plants have evolved mechanisms to deal with abiotic stress, including soil salinity, and maintain growth and reproduction under non-optimal conditions. Plants sessile nature disables them from moving out from a stressful into a non-stressful environment. Thus, some plants have evolved ways to adjust phenology, physiology and development to cope with environmental changes.

The earliest stages of the plant development exposed to the environment are germination and seedling growth, which may be among the most saline sensitive (Chang et al. 1961; Wilczek et al. 2009; Donohue et al. 2010). Soil salinity may impede or delay seed germination and seedling development (Wahid et al. 1999). If the seed germinates and the seedling develops in saline environments, early root development is crucial for survival because it determines water absorption capacity, which in turn may buffer against saline-induced osmotic stress (Galvan-Ampudia and Testerink 2011).

After plants are established, roots must sense both osmotic and ionic components of salinity and be able to respond to changes. Little is known about stress receptors, but the mostly studied candidates suggest that the plant can sense both stresses independently and, if so, that both stresses should play a role in salinity stress responses. For example, the *Arabidopsis thaliana* plasma membrane histidine kinase *AtHK1* gene, homolog to the yeast *Sln1* that senses cell water pressure, is believed to have a similar osmosensor role in plant roots initiating downstream responses to osmotic stress (Wohlbach et al. 2008). In the case of salt, the Na^+/H^+ plasma membrane antiporter SOS1 is a strong candidate to be a Na^+ sensor (Zhu 2002, 2003). The *A. thaliana AtSOS1* mutant was identified for being *salt overly sensitive* and when overexpressed increases cell salinity tolerance by reducing internal Na^+ (Shi et al. 2000; Qiu et al. 2002).

Stress perception triggers the biosynthesis of the hormone abscisic acid (ABA). ABA is known to integrate plant growth and development in response to both drought and high saline environments and may act as a long-distance signaling molecule (Zhu 2002; Davies et al. 2005). ABA synthesized in the roots is transported to the shoots and induces stomata closure to reduce water loss (Davies et al. 2005; Chaves et al. 2009). Additionally, ABA induces the expression of a wide set of genes, including particular families of transcription factors such as basic leucine zipper (bZip), NAC or MYB, which will assist in the coordination of stress specific responses (Conde et al. 2011; Janiak et al. 2015). Also, interactions with other hormones such as auxin, ethylene, jasmonic acid and cytokinin are essential to integrate stress responses with growth regulation (Davies et al. 2005; Janiak et al. 2015).

Extracellular Na^+ triggers quick and transient increases of Ca^{2+} in the cytosol, which are also believed to be involved in long-distance signaling for salinity stress (Tracy et al. 2008). Calcium concentration variations are stress and tissue dependent (Kiegle et al. 2000). Oscillations in Ca^{2+} signal the expression of proteins from the calmodulin (CaM), calcineurin B-like (CBL), class 2C phosphatase (PP2C), and from the calcium-dependent and CBL interacting protein kinase (CPK and CIPK, respectively) families, which then trigger downstream stress responses (Zhu 2002; Conde et al. 2011; Janiak et al. 2015).

Reactive oxygen species (ROS), such as hydrogen peroxide (H_2O_2), have also been suggested to signal ionic stress (Conde et al. 2011; Janiak et al. 2015). Mitogen-activated protein kinase (MAPK) cascades have been suggested to mediate signal transduction between salt dependent ROS increase and stress specific gene expression (Schmidt et al. 2013).

Overall, several signaling pathways seem to be involved in triggering plant responses to stress to enable survival under saline conditions.

Mechanisms of salinity stress resistance

Salinity resistance is the ability of plants to grow and complete their life cycle when grown in saline soils (Maas 1986). Plants have evolved various mechanisms to minimize the effects of salinity, including tolerance, avoidance and exclusion mechanisms (Munns and Tester 2008). These mechanisms are not mutually exclusive, may confer salinity resistance at different developmental stages, and their success is variable among species (Flowers et al. 1977; Bayuelo-Jiménez et al. 2002; Flowers and Colmer 2008).

Tolerance mechanisms can broadly be distinguished into osmotic stress tolerance and tissue tolerance to ionic stress. Osmotic stress can also result from high salt concentration in the cells. Under osmotic stress, plants respond by minimizing water loss (Boursiac et al. 2005; Chaves et al. 2009). Some plants are able to adjust photosynthetic and transpiration rates by reducing leaf open stomatal density (Yu et al. 2012), by inducing stomata closure and/or by regulating stomata aperture (Chaves et al. 2009). Cell osmotic adjustment allows for the recovery of cell turgor after water loss due to a sudden increase in soil salinity in a matter of hours, but cell elongation rates are reduced (Passioura and Munns 2000; Boursiac et al. 2005). Therefore, stomatal closure, consequential leaf temperature increase, and growth reduction are the main osmotic deficit responses (Munns and Tester 2008; Chaves et al. 2009).

Ion toxicity results mostly from excessive cellular Na^+, which can inhibit K^+ uptake and consequently restrict K^+ dependent metabolic and physiological functions (Hauser and Horie 2010). Being able to accumulate other nutrients such as K^+ when Na^+ is elevated, compartmentalize ions and synthesize organic compounds for osmotic adjustment, maintain transpiration and photosynthesis under high salt, or even limit the amount of salt in the xylem, comprise some of the mechanism that have evolved in plants that are able to grow and reproduce in highly saline environments (Munns and Tester 2008; Flowers and Colmer 2008; Hauser and Horie 2010).

Tissue ion tolerance usually involves Na^+ and Cl^- compartmentalization in the cell to avoid toxic concentrations in the cytoplasm (Munns and Tester 2008). To minimize the effects of water loss due to high salt concentrations, plants accumulate metabolites in the cytoplasm for osmoprotection and for osmotic adjustment. For example, this process is required to balance the osmotic potential of Na^+ and Cl^- being isolated into vacuoles (Hauser and Horie 2010). Moreover, ion balances, namely Na^+:K^+, may be more important to salinity tolerance than is the absolute Na^+ concentration (Maathuis and Amtmann 1999; Chartzoulakis et al. 2002; Meloni et al. 2008; Kronzucker and Britto 2011; Hauser and Horie 2010).

Avoidance mechanisms involve keeping the toxic ions away from parts of the plant where they can be harmful. Salt accumulation in the cells can have severe consequences, including cell dehydration, inhibition of enzymes involved in the carbohydrate metabolism or even affecting photosynthesis directly by accumulating in the chloroplasts (Munns and Tester 2008). Some plants have the ability to reduce ionic stress by limiting Na^+ and Cl^- accumulation in the cytosol, and controlling Na^+ transport through the expression regulation of particular ion channels and transporters (Munns and Tester 2008; Rajendran et al. 2009). Whether it involves ion compartmentalization in the vacuoles, ion exclusion, excretion, or even dilution, avoiding ion accumulation to toxic levels is a common avoidance mechanism (Munns and Tester 2008).

Besides, a common signaling cascade has been proposed to control germination and flowering time in *A. thaliana* under salt stress, two important phenological benchmarks for plant development (Kim and Park 2007). For example, another possible avoidance mechanism that may be particularly useful if salinity stress is seasonal and predictable is reproducing before the stress becomes too intense. While rarely studied, stress resistance in plants may incorporate adjusts in phenology, morphology, and development.

Root architecture under salinity stress

To fully benefit from the soil environment, plants should be able to modulate root growth, and mechanisms have evolved to allow for the control of root architecture. Early primary and lateral root growth plays an important role in the ability of the plant to detect and potentially respond to the environment (Galvan-Ampudia and Testerink 2011; Janiak et al. 2015). Moreover, root traits such as length, diameter, biomass and specific surface have been associated with differential shoot growth and reproductive success under stress (Comas et al. 2013). Besides facilitating water uptake *per se*, redirection of root growth can allow plants to explore microenvironments that may be more favorable for plant growth (Nibau et al. 2008). Interactions between the plant's intrinsic developmental program and the environment will dictate the final root architecture (Galvan-Ampudia and Testerink 2011).

Not surprisingly, salinity stress has a large effect on root architecture and most mediators in the salinity-signaling cascade are able to interfere in root design. For example, most root morphological changes due to the presence of salt can also be stimulated by ABA (DeSmet et al. 2003; Galvan-Ampudia and Testerink, 2011). Moreover, the putative osmosensor *AtHK1* also has a role in regulating root elongation under water stress: mutants have shorter roots under stress, and overexpression allows the maintenance of primary root elongation (Wohlbach et al. 2008). Nevertheless, there might be advantages of continuously producing root tissue given that young roots absorb most of the water for the plant, a limited resource under high soil salinity (Peterson et al. 1993).

Towards salinity resistance

Soil salinity is a major selective factor that affects plant growth and reproduction. A common expectation of plants that inhabit saline environments is that they have evolved salinity adaptation, i.e., that they will survive and reproduce to a greater extent under higher salt conditions than populations that evolved in

non-saline environments (Lowry et al. 2008, 2009; Munns and Tester 2008; Friesen et al. 2014; Busoms et al. 2015). Recent studies include the genetic quantification of adaptation (e.g., Galloway and Fenster 2000; Hall et al. 2010; Turner et al. 2010; Fournier-Level et al. 2011), but the understanding of the specific traits and underlying genetics of local adaptation are limiting and more experimental data is necessary to enable distinguishing adaptive from demographic evolution (Hereford 2009; Ehrenreich and Purugganan 2009; Vitti et al. 2013).

Associating genes and traits

In the simplest model, local adaptation involves major genes with strong contributions to fitness and the candidate genes inform us about ecologically relevant mechanisms (MacNair 1983; Courbot et al. 2007; Baxter et al. 2010; Friesen et al. 2014). Such situations are amenable to direct genetic analysis using the logic of selective sweeps, genome-wide association studies (GWAS) or quantitative trait loci (QTL) analyses. Generally, association mapping relies on the identification of genetic markers that are correlated with the traits of interest and the expectation that the candidate genes are in linkage disequilibrium (LD – non-random association of alleles at different *loci*) with such markers, which depends on the population's patterns of selection, but also on mating system, mutation, recombination, and migration (Barton et al. 2007, Slatkin 2008).

More commonly evolution involves multiple genes of small effect, each influencing distinct traits that make incremental contributions to fitness (Pritchard and Di Rienszo 2010; Corre and Kremer 2012; Gould et al. 2014). Moreover, depending on the frequency and distribution of causal alleles within a population, different individuals may contain different combinations of alleles and thus subtly different adaptive capacities. This combination of genetic heterogeneity and incremental effects is intractable to standard genetic tests. However, the recent advent of genome-scale analysis to identify genomic variants and computational methods to quantify and relate variation within and among genomes, provides

increasing power to nominate candidate genes and genomic intervals associated with adaptive phenotypes (Galloway and Fenster 2000; Hall et al. 2010; Turner et al. 2010; Fournier-Level et al. 2011; Ingvarssom and Street 2011). For example, whole-genome scans enabled the identification of candidate genes for local adaptation to serpentine soils in *Arabidopsis lyrata* populations (Turner et al. 2010).

Salinity adaptation

As the number of candidate genes and traits potentially associated with adaptive evolution to salinity and other environmental constraints are being discovered, the knowledge gathered can be used to linking ecology and evolution to crop improvement (Friesen and von Wettberg 2010; Milla et al. 2015).

Plants that can constitutively grow and reproduce under high soil salinity are generally designated halophytes and the converse, i.e. plants that mostly inhabit low salinity environments and that are generally affected by high salinity levels, are commonly named glycophytes (Flowers et al. 1977). Glycophytes have evolved to grow and reproduce successfully in low salinity environments and to successfully grow under higher salinity they need to evolve mechanisms to resist the additional stress (Flowers et al. 1977; Flowers and Colmer 2008). Despite, there is a considerable variation to salinity resistance among and within populations of glycophytes (Munns and Tester 2008). Interestingly, halophytes independently evolved multiple times sharing plant families with glycophytes and, consequently, there are frequently closely related species to explore evolved mechanisms of salinity resistance (Flowers et al. 1977; Flowers and Colmer 2008).

Several candidate genes and alleles that contribute to enhanced salinity resistance have been identified in model species such as *A. thaliana,* rice and wheat and considered for crop improvement (Rus et al. 2006; Munns and Tester 2008; James et al. 2012; Schmidt et al 2013; Roy et al. 2014). When relevant genetic variation is found in close relatives, including from different species if they can hybridize and produce viable offspring, controlled crosses and backcrosses

can be performed such that the gene and/ or trait of interest are introduced in the new crop variety (Friesen and von Wettberg 2010; Warschefsky et al. 2014; Roy et al. 2014). For example, an allele coding for the *A. thaliana* HKT1 (a high affinity potassium transporter that regulates Na^+ accumulation in the shoot from the root) has been identified and characterized in wild populations for conferring increased salinity tolerance (Rus et al. 2006). Subsequently, alleles with the same characteristics have been identified in an ancestral species of wheat and, through the production of near-isogenic lines, have been successfully introduced into the cultivated background enabling the high quality production with enhanced salinity tolerance (James et al. 2012). On the other hand, when relevant natural variation is not available one possible approach is the production of transgenic crop varieties, which carry caveats such as: a single gene approach may not solve the problem; the effect of the transgene in the crop may have different trait effects relative to the ones expected from the original organism; it takes a long time to produce, test and approve such varieties (Roy et al. 2014).

Studying natural populations and their adaptations, particularly in the genomics era, is unraveling mechanisms of stress resistance together with their most likely genetic basis (Galloway and Fenster 2000; Hall et al. 2010; Turner et al. 2010; Fournier-Level et al. 2011; Ingvarssom and Street 2011). For a broader impact in crop improvement, it is essential to gather a better understanding of what are the genes and traits underlying adaptive evolution of the domesticated wild relatives (Friesen and von Wettberg 2010; Warschefsky et al. 2014; Milla et al. 2015). Additionally, because plant domestication is usually accompanied by reproductive isolation and, consequently, speciation, studies on hybridization are facilitating the introduction of genes and traits identified in the wild relatives in cultivated germplasm (Friesen and von Wettberg 2010; Warschefsky et al. 2014). There is a need to introduce sustainable agricultural practices globally, while trying to make available new crop varieties that are able to maintain high yields in the scenario of climate change, as well as enable the use of marginal soils and environments (Friesen and von Wettberg 2010). Novel traits that potentiate the

widening of crop species breadth and maximize fitness in different environments need to be incorporated into cultivated species to allow for a continued sustainable production for global food safety under new environmental conditions.

Legumes

Fabaceae, most commonly called the legume family, is among the largest and more diverse flowering plant families, containing close to 20,000 species distributed over at least 700 genera, that occupy wide ecogeographical niches varying from desert to tropical or even aquatic environments and ranging from low to high altitudes (Pohill et al. 1981; Graham and Vance 2003; Doyle and Luckow 2003; Gepts et al. 2005; Lavin et al. 2005). Legume diversity provides a rich resource to perform integrative ecological, molecular and agricultural studies.

Fossil records date legume divergence in the Late Paleocene with the first conclusive legume identified ~56M years ago (Herendeen et al. 1992). Based on the same criteria, it is believed that within clade divergence occurred mostly within ~1M year after the initial separation (Doyle and Luckow 2003; Lavin et al. 2005). Molecular data suggest that a whole-genome duplication preceded legume divergence ~58M years ago, and that it played a major role in the evolution of legume-rhizobia symbiotic nitrogen fixation (Schlueter et al. 2004; Lavin et al. 2005; Young et al. 2011).

Global economical and nutritional importance

Legumes have been important in agriculture for thousands of year: from the domestication of lentil (*Lens esculenta*) ~9,500 to 8,000 years ago in Iran (Cohen 1977), to the consumption of beans (*Phaseolus vulgaris*) in America ~3,000 years ago (Kaplan and Lynch 1999), and to the use for forage and soil improvement since the Roman Era (Fred et al. 1932). Currently, Fabaceae are second after the Graminiae in their importance to humans for food and feed and are grown in 13% to 15% of the World's arable land (Graham and Vance 2003; Gepts et al. 2005).

Grain legumes are the major source of vegetable protein for human nutrition, combining about one third of human total nitrogen intake (Vance et al. 2000; Graham and Vance 2003). Bean, lentil, broad bean (*Vicia faba*), chickpea (*Cicer arietinum*), cowpea (*Cajanus cajan*) and pea (*Pisum sativum*) are amongst the most used for food; soybean (*Glycine max*) and peanut (*Arachis hypogeae*) are used in ~35% of the World's vegetable oil production; and forage legumes such as medics (*Medicago spp.*) and clovers (*Trifolium spp.*) are of great importance for animal feed (Graham and Vance 2003).

Given legume's global agricultural, nutritional and economical importance, together with the broad ecological niche range and associated wide natural diversity, then the importance of integrating ecological, genomic and agricultural studies becomes particularly rich within this plant family.

Symbiosis and sustainable agriculture

Most legumes (> 88% of the species) have the unique capacity for symbiotic nitrogen fixation; ability that likely underlies their evolutionary, ecological and economical success (de Faria et al. 1989; Graham and Vance 2003).

Legume and soil bacteria that fix atmospheric nitrogen (rhizobia) are able to engage in a symbiotic interaction where the bacteria provides nitrogen in the form of ammonia to the plant, and the plant feeds back energy in the form of carbon to the bacteria. This interaction starts with "communication" between plant and bacteria: the plant releases flavonoids that when are recognized by the bacteria, trigger the production of a lipooligosaccharide called "Nod factor" (Long 1996). When and if the plant recognizes the Nod factor, several responses are triggered starting with root hairs curling around the bacteria and culminating in the formation of a new organ containing nitrogen fixing bacteroids in the core – the nodule (Long 1996; Penmetsa et al. 2003; Larrainzar et al. 2015). Therefore, nodulation is a very tightly regulated process that involves signaling and transcriptional changes coordinated between host and symbiont or could otherwise become too costly for

the plant (Penmetsa and Cook 2000; Penmetsa et al. 2003; Oldroyd and Downie 2006; Larrainzar et al. 2015).

Soil salinity tends to decrease nodulation in legume plants and may influence salinity tolerance (Bianco and Defez, 2009; Flowers et al. 2010). Interestingly, the signaling pathways that modulate this interaction are at least partially overlapping with the ones involved in adjusting root architecture (Gonzales-Rizzo 2006). Symbiotic nitrogen fixation allows legume species to grow and reproduce in low nitrogen environments and therefore potentially allows legumes to colonize a wider range of environments relative to non-nodulating species. Besides, in agricultural terms, symbiosis allows for a low input and less pollutant way of facilitating the nitrogen cycle and increasing soil nutrition.

The model legume Medicago truncatula

Medicago truncatula var. *truncatula* (Fabaceae) is a highly selfing (selfing rate in wild populations > 95%) annual nodulating legume that occurs in a wide range of environments throughout the Mediterranean region (Bonnin et al. 2001; Lazrek et al. 2009; Branca et al. 2011). Phylogenetically, the genus Medicago belongs to the Papilionoideae subfamily, such as most of the economically important legumes (Doyle and Luckow 2003). Within the *Papilionoideae*, *Medicago* falls in the same clade (Galegoid) with cool season legumes from the genera *Cicer*, *Pisum*, *Vicia* and *Lens*, and is believed to have split from the tropical legumes clade (Phaseolid), that contains the *Cajanus*, *Glycine*, *Vigna* and *Phaseolus* genera, about 54M year ago (Doyle and Luckow 2003; Zhu et al. 2005). This close phylogenetic distance with economically important crops is a major reason why *M. truncatula* became a model to study legume specific processes.

Some of the intrinsic characteristics that make *M. truncatula* a good legume study system include having a short life cycle (~3 months from seed to seed), and a small (~ 375 Mb), diploid (2n = 16), and currently sequenced, annotated and available genome (Cook 1999; Branca et al. 2011; Young et al. 2011). Most of the molecular, developmental and physiological studies have been conducted using

the reference *M. truncatula* genotype A17, the first of this species to be sequenced, and therefore used as the reference genome (Young et al. 2011). Because of the high synteny (similar relative position of *loci* along each chromosome) with other legumes, the knowledge from studying legume specific processes in *M. truncatula* can be more easily applied to other important legume crops (Zhu et al. 2005; Young et al. 2011).

Being a model species, a rich set of genetic resources was built over the years and is currently available for the scientific legume community. For example, at least two mutant collections were developed: one TILLING (**T**argeting **I**nduced **L**ocal **L**esions **IN G**enomes) population with point nucleotide mutations across the A17 genome (Penmetsa and Cook 2000); a fast neutron bombardment collection that contains single nucleotide deletions (also called De-TILLING) across the A17 genome (Rogers et al. 2009); and a *Tnt1* retrotransposon insertional mutant family based on the closely related R108 (Tadege et al. 2008). Such mutant collections have been used to either discover the genes underlying a certain phenotype, or to test the function of particular genes, i.e., these mutant populations can be used in forward or reverse genetics (starting in phenotype or in genotype, respectively). While the TILLING population is more likely to offer weaker alleles, the Tnt1-insertion mutant population usually delivers knockout mutants, and the De-TILLING collection tends to provide knockouts, but may also produce weaker alleles. Additionally, efficient and reproducible transformation protocols have been developed for *M. truncatula*, namely using the highly embriogenic line M9-10a obtained from the same cultivar as A17 (Araújo et al. 2004; Duque et al. 2007). Moreover, *M. truncatula* has been used in several genomic and ecological studies, some of which are reported in the present thesis (Friesen et al. 2010; Branca et al. 2011; Young et al. 2011; Castro et al. 2013; Cordeiro et al. 2014; Friesen et al. 2014).

For this dissertation, natural populations of *M. truncatula* collected in saline and non-saline soils from northern Tunisia and southern Portugal were used to study local adaptation to salinity. Initially, Tunisian populations were genotyped and

searched for genome-wide signatures of salinity-dependent differentiation, which revealed candidate regions and genes for salinity adaptation (Friesen et al. 2010). After this, whole genome sequencing of the Tunisian populations gave further support to the previously identified regions (Friesen et al. 2014). Additionally, reciprocal soil field and greenhouse experiments were conducted in Tunisia, together with a greenhouse experiment in UC Davis (California, USA) to phenotype the Tunisian genotypes (Friesen et al. 2014). Furthermore, plate and aeroponical experiments were performed to understand how germination and early root growth differ in these populations (Cordeiro et al. 2014). Parental environmental effects due to salinity were tested in germinations and adult plants of the Tunisian *M. truncatula* populations (Cordeiro et al. 2014; Moriuchi et al. 2016). Finally, to understand whether isolated populations that evolved under saline conditions have evolved the same or different mechanism to deal with salinity, saline and non-saline populations were collected in southern Portugal (Cordeiro et al. submitted). The whole genome of these genotypes was sequenced and compared with Tunisian genotypes. To test for local adaptation to salinity, a field reciprocal experiment was then conducted in the Portuguese original sites using also the best and worst performers under salt and no salt of the Tunisian saline and non-saline origin populations, respectively. Summarily, Tunisian and Portuguese natural populations of the model legume *M. truncatula* are used to study local adaptation, mechanisms of salinity resistance, and the genes and traits underlying such adaptive responses. Particularly, chapter two focuses on the population genomics of Tunisian *M. truncatula* (Friesen et al. 2010); chapter three on the population differentiation on germination and early root growth traits in saline and non-saline adapted Tunisian *M. truncatula* (Cordeiro et al. 2014); and chapter four on the independent evolution of salinity tolerance in Tunisian and Portuguese populations of *M. truncatula* (Cordeiro et al. submitted).

References

Araújo SS, Duque A, dos Santos D, Fevereiro PS (2004) An efficient transformation method to regenerate a high number of transgenic plants using a new embryogenic line of Medicago truncatula cv. Jemalong. *Plant cell, tissue and organ culture*, **78**, 123-131.

Atkins KE, Travis JMJ (2010) Local adaptation and the evolution of species' ranges under climate change. *Journal of Theoretical Biology*, **266**, 449-457.

Barton NH, Briggs DE, Eisen JA, Goldstein DB, Patel NH (2007) Evolution, Cold Spring Harbor Laboratory Press, pp 833.

Baxter I, Brazelton JN, Yu D, Huang YS, Lahner B, Yakubova E, Li Y, Bergelson J, Borevitz JO, Nordborg M, et al (2010) A coastal cline in sodium accumulation in Arabidopsis thaliana is driven by natural variation of the sodium transporter AtHKT1;1. *PLoS genetics,* **6**: e1001193.

Bayuelo-Jiménez JS, Craig R, Lynch JP (2002) Salinity tolerance of Phaseolus species during germination and early seedling growth. *Crop Science*, **42**, 1584–1594.

Bianco C, Defez R (2009) *Medicago truncatula* improves salt tolerance when nodulated by an indole-3-acetic acid-overproducing *Sinorhizobium meliloti* strain. *Journal of Experimental Botany*, **60**, 3097–3107.

Bonnin I, Ronfort J, Wozniak F, Olivieri I (2001) Spatial effects and rare outcrossing events in Medicago truncatula (Fabaceae). *Molecular ecology,* **10**, 1371-1383.

Boursiac Y, Chen S, Lu DT, Sorieul M, van den Dries N, Maurel C (2005) Early effects of salinity on water transport in Arabidopsis roots. Molecular and cellular features of aquaporin expression. *Plant physiology,* **139**, 790-805.

Branca A, Paape TD, Zhou P, Briskine R, Farmer AD, Mudge J, Bharti AK, Woodward JE, May GD, Gentzbittel L, et al (2011) Whole-genome nucleotide diversity, recombination, and linkage disequilibrium in the model legume *Medicago truncatula*. *Proceedings of the National Academy of Sciences of the*

U S A, **108**, E864-E870.

Busoms S, Teres J, Huang XY, Bomblies K, Danku J, Douglas A, Weigel D, Poschenrieder C, Salt DE (2015) Salinity Is an Agent of Divergent Selection Driving Local Adaptation of Arabidopsis to Coastal Habitats. *Plant physiology,* **168**, 915-929.

Caicedo AL, Williamson SH, Hernandez RD, Boyko A, Fledel-Alon A, York TL, Polato NR, Olsen KM, Nielsen R, McCouch SR, et al (2007) Genome-wide patterns of nucleotide polymorphism in domesticated rice. *PLoS genetics,* **3**, 1745-1756.

Challinor AJ, Watson J, Lobell DB, Howden SM, Smith DR, Chhetri N (2014) A meta-analysis of crop yield under climate change and adaptation. *Nature Climate Change*, **4**, 287-291.

Chang, C. W. 1961. Effect of irrigation water and exchangeable sodium on soil properties and growth of alfalfa. *Soil Science*, **91**, 29–37.

Charmantier A, McCleery RH, Cole LR *et al.* (2008) Adaptive phenotypic plasticity in response to climate change in a wild bird population. *Science*, **320**, 800-803.

Chartzoulakis K, Patakas A, Kofidis G, Bosabalidis A, Nastou A (2002) Water stress affects leaf anatomy, gas exchange, water relations and growth of two avocado cultivars. *Scientia Horticulturae*, **95**, 39-50.

Chaves MM, Flexas J, Pinheiro C (2009) Photosynthesis under drought and salt stress: regulation mechanisms from whole plant to cell. *Annals of Botany*, **103**, 551-560.

Cohen MN (1977) The Food Crisis in Prehistory: Overpopulation and the Origins of Agriculture. Yale University Press, New Haven, CT.

Comas LH, Becker SR, Cruz VM, Byrne PF, Dierig DA (2013) Root traits contributing to plant productivity under drought. *Frontiers in Plant Sciences*, **4**, 442.

Conde A, Chaves MM, Gerós H (2011) Membrane transport, sensing and signaling in plant adaptation to environmental stress. *Plant Cell Physiology*, **52**, 1583-1602.

Cook DR (1999) *Medicago truncatula*—a model in the making! *Current Opinion Plant Biology*, **2**, 301–304.

Cordeiro MA, Moriuchi KS, Fotinos TD, Miller KE, Nuzhdin SV, et al (2014) Population differentiation for germination and early seedling root growth traits under saline conditions in the annual legume *Medicago truncatula* (Fabaceae). *American Journal of Botany*, **101**, 488-498.

Corre V, Kremer A (2012) The genetic differentiation at quantitative trait loci under local adaptation. *Molecular Ecology*, **21**, 1548-1566.

Courbot M, Willems G, Motte P, Arvidsson S, Roosens N, et al (2007) A major quantitative trait locus for cadmium tolerance in *Arabidopsis halleri* colocalizes with HMA4, a gene encoding a heavy metal ATPase. *Plant Physiology*, **144**, 1052-1065.

Darwin, C (1859) On the origin of species by means of natural selection, or the preservation of favored races in the struggle for life. Nature (1st ed.) John Murray, London. **5**: 502.

Davies, W. J., G. Kudoyarova, and W. Hartung. 2005. Long-distance ABA signaling and its relation to other signaling pathways in the detection of soil drying and the mediation of the plant's response to drought. *Journal of Plant Growth Regulation*, **24**, 285–295.

de Faria SM, Lewis GP, Sprent JI, Sutherland JM (1989) Occurrence of nodulation in the Leguminosae. *New Phytologist*, **111**, 607-619.

De Smet I, Signora L, Beeckman T, Inzé D, Foyer CH, Zhang H (2003) An abscisic acid-sensitive checkpoint in lateral root development of Arabidopsis. *The Plant journal : for cell and molecular biology,* **33**, 543-555.

Donohue K, de Casas RR, Burghardt L, Kovach K, Willis CG (2010) Germination, postgermination adaptation, and species ecological ranges. *Annual Review of*

Ecology, Evolution and Systematics, **41**, 293-319.

Doyle JJ, Luckow MA (2003) The rest of the iceberg. Legume diversity and evolution in a phylogenetic context. *Plant Physiology*, **131**, 900-910.

Dudley SA, Schmitt J (1996) Testing the adaptive plasticity hypothesis: density-dependent selection on manipulated stem length in Impatiens capensis. *American Naturalist*, **147**, 445-465.

Duque AS, de Sousa Araújo S, Cordeiro MA, Santos DM, Fevereiro MP (2007) Use of fused gfp and gus reporters for the recovery of transformed *Medicago truncatula* somatic embryos without selective pressure. *Plant Cell, Tissue and Organ Culture*, **90**, 325-330.

Ehrenreich IM, Purugganan MD (2006) The molecular genetic basis of plant adaptation. *American Journal of Botany*, **93**, 953-962.

Etterson JR, Shaw RG (2001) Constraint to adaptive evolution in response to global warming. *Science*, **294**, 151-154.

FAO 2008. FAO Land and Plant Nutrition Management Service. Available at http://www.fao.org/ag/agl/agll/spush.

FAO FAOSTAT (2013); Available at http://faostat.fao.org

Fisher R (1930) The genetical theory of natural selection. Oxford University Press, Oxford, UK.

Flowers TJ, Colmer TD (2008) Salinity tolerance in halophytes. *New Phytologist*, **179**, 945–963.

Flowers TJ, Gaur PM, Gowda CLL, Krishnamurthy L, Samineni S, Siddique KH, Turner NC, Vadez V, Varshney RK, Colmer TD (2010) Salt sensitivity in chickpea. *Plant, cell & environment,* **33**, 490-509.

Flowers TJ, Troke PF, Yeo AR (1977) The mechanism of salt tolerance in halophytes. *Annual Review of Plant Physiology*, **28**, 89–121.

Flowers TJ, Yeo AR (1995) Breeding for salinity resistance in crop plants: where next? *Australian Journal of Plant Physiology*, **22**, 875–884.

Fournier-Level A, Korte A, Cooper MD, Nordborg M, Schmitt J, Wilczek AM (2011) A map of local adaptation in Arabidopsis thaliana. *Science*, **334**, 86-89.

Fred EB, Baldwin IL, McCoy E (1932) Root nodule bacteria and leguminous plants. University of Wisconsin Press, Madison.

Friesen ML, Cordeiro MA, Penmetsa RV, Badri M, Huguet T, Aouani ME, Cook DR, Nuzhdin SV (2010) Population genomic analysis of Tunisian Medicago truncatula reveals candidates for local adaptation. *The Plant journal : for cell and molecular biology,* **63**, 623-635.

Friesen ML, von Wettberg EJ (2010) Adapting genomics to study the evolution and ecology of agricultural systems. *Current opinion in plant biology,* **13**, 119-125.

Friesen ML, von Wettberg EJ, Badri M, Moriuchi KS, Barhoumi F, et al (2014) The ecological genomic basis of salinity adaptation in Tunisian *Medicago truncatula*. *BMC Genomics*, **15**, 1160.

Galloway LF, Fenster CB (2000) Population differentiation in an annual legume: local adaptation. *Evolution*, **54**, 1173-1181.

Galvan-Ampudia CS, Testerink C (2011) Salt stress signals shape the plant root. *Current Opinion in Plant Biology*,**14**, 296–302.

Gepts P, Beavis WD, Brummer EC, Shoemaker RC, Stalker HT, Weeden NF, Young ND (2005) Legumes as a model plant family. Genomics for food and feed report of the cross-legume advances through genomics conference. *Plant physiology,* **137**, 1228-1235.

Gonzalez-Rizzo S, Crespi M, Frugier F (2006) The *Medicago truncatula* CRE1 cytokinin receptor regulates lateral root development and early symbiotic interaction with *Sinorhizobium meliloti. Plant Cell*, **18**, 2680-2693.

Gould B, McCouch S, Geber M (2014) Variation in soil aluminium tolerance genes is associated with local adaptation to soils at the Park Grass Experiment. *Molecular Ecology*, **23**, 6058-6072.

Graham PH, Vance CP (2003) Legumes: importance and constraints to greater use. *Plant Physiology*, **131**, 872-877.

Hall BG, Ehrlich GD, Hu FZ (2010) Pan-genome analysis provides much higher strain typing resolution than multi-locus sequence typing. *Microbiology*, **156**, 1060-1068.

Hauser F, Horie T (2010) A conserved primary salt tolerance mechanism mediated by HKT transporters: a mechanism for sodium exclusion and maintenance of high K^+/Na^+ ratio in leaves during salinity stress. *Plant, Cell & Environment,* **33**, 552–565.

Hereford J (2009) A quantitative survey of local adaptation and fitness trade-offs. *American Naturalist*, **173**, 579-588.

Herendeen PS, Crepet WL, Dilcher DL (1992) The fossil history of the Leguminosae: Phylogenetic and biogeographic implications. Pages 303-316 in Advances in legume systematics, part 4, the fossil record (P. S. Herendeen and D. L. Dilcher, eds.). Royal Botanic Gardens, Kew, UK.

Ingvarsson PK, Street NR (2011) Association genetics of complex traits in plants. *New Phytologist*, **189**, 909-922.

IPCC (2007) Climate change 2007: the physical science basis. *Agenda*, **6**, 333.

Jacobsen T, Adams RM (1958) Salt and Silt in Ancient Mesopotamian Agriculture: Progressive changes in soil salinity and sedimentation contributed to the breakup of past civilizations. *Science*, **128**, 1251-1258.

James RA, Blake C, Zwart AB, Hare RA, Rathjen AJ, Munns R (2012) Impact of ancestral wheat sodium exclusion genes Nax1 and Nax2 on grain yield of durum wheat on saline soils. *Functional Plant Biology*, **39**, 609-618.

Janiak A, Kwaśniewski M, Szarejko I (2015) Gene expression regulation in roots under drought. *Journal of Experimental Botany*, erv512.

Jiao Y, Zhao H, Ren L, Song W, Zeng B, et al (2012) Genome-wide genetic changes during modern breeding of maize. *Nature Genetics*, **44**, 812-815.

Kaplan L, Lynch TF (1999) Phaseolus (Fabaceae) in archeology: AMS radio-carbon dates and their significance in pre-Colombian agriculture. *Economic Botany*, **53**, 261-272

Kates RW, Travs WR, Wilbanks TJ (2012) Transformational adaptation when incremental adaptations to climate change are insufficient. *Proceedings of the National Academy of Sciences of the U S A*, **109**, 7156-7161.

Kawecki TJ, Ebert D (2004) Conceptual issues in local adaptation. *Ecology Letters*, **7**, 1225-1241.

Kiegle E, Moore CA, Haseloff J, Tester MA, Knight MR (2000) Cell-type-specific calcium responses to drought, salt and cold in the Arabidopsis root. *Plant Journal*, **23**, 267-278.

Kim SG, Park CM (2007) Membrane-mediated salt stress signaling in flowering time control. *Plant Signaling and Behavior*, **2**, 517-518.

Kronzucker HJ, Britto DT (2011) Sodium transport in plants: a critical review. *New Phytologist*, **189**, 54-81.

Lam HM, Xu X, Liu X, Chen W, Yang G, Wong FL, Li MW, He W, Qin N, Wang B, et al (2010) Resequencing of 31 wild and cultivated soybean genomes identifies patterns of genetic diversity and selection. *Nature genetics,* **42**, 1053-1059.

Larrainzar E, Riely BK, Kim SC, Carrasquilla-Garcia N, Yu HJ, Hwang HJ, Oh M, Kim GB, Surendrarao AK, Chasman D, et al (2015) Deep Sequencing of the Medicago truncatula Root Transcriptome Reveals a Massive and Early Interaction between Nodulation Factor and Ethylene Signals. *Plant physiology,* **169**, 233-265.

Lavin M, Herendeen PS, Wojciechowski MF (2005) Evolutionary rates analysis of Leguminosae implicates a rapid diversification of lineages during the tertiary. *Systems Biology*, **54**, 575-594.

Lazrek F, Roussel V, Ronfort J, Cardinet G, Chardon F, Aouani ME, Huguet T (2009) The use of neutral and non-neutral SSRs to analyse the genetic structure of a Tunisian collection of Medicago truncatula lines and to reveal associations with eco-environmental variables. *Genetica,* **135**, 391-402.

Long SR (1996) Rhizobium symbiosis: nod factors in perspective. *Plant Cell,* **8**, 1885.

Lowry DB, Hall MC, Salt DE, Willis JH (2009) Genetic and physiological basis of adaptive salt tolerance divergence between coastal and inland Mimulus guttatus. *New Phytologist*, **183**, 776-788.

Lowry DB, Rockwood RC, Willis JH (2008) Ecological reproductive isolation of coast and inland races of *Mimulus guttatus*. *Evolution*, **62**, 2196-2214.

Luikart G, England PR, Tallmon D, Jordan S, Taberlet P (2003) The power and promise of population genomics: from genotyping to genome typing. *Nature reviews genetics*, **4**, 981-994.

Maas EV (1986) Salt tolerance of plants. *Applied Agricultural Research*, **1**, 12–26.

Maathuis FJM, Amtmann A (2004) K Nutrition and Na toxicity: The basic of cellular K/Na ratios. *Annals of Botany*, **84**, 123-133.

Macnair MR (1983) The genetic control of copper tolerance in the yellow monkey flower, *Mimulus guttatus*. *Heredity*, **50**, 283-293.

Marsden CD, Ortega-Del Vecchyo D, O'Brien DP, Taylor JF, Ramirez O, et al (2016) Bottlenecks and selective sweeps during domestication have increased deleterious genetic variation in dogs. *Proceedings of the National Academy of Sciences of the U S A*,**113**, 152-157.

Meloni DA, Gulotta MR, Martinez CA (2008) Salinity tolerance in Schinopsis quebracho colorado: Seed germination, growth, ion relations and metabolic responses. *Journal of Arid Environments*, **72**, 1785-1792.

Milla R, Osborne CP, Turcotte MM, Violle C (2015) Plant domestication through an ecological lens. *Trends in Ecology & Evolution*, **30**, 463-469.

Moriuchi KS, Friesen ML, Cordeiro MA, Badri M, Vu WT, Main BJ, Aouani ME, Nuzhdin SV, Strauss SY, von Wettberg EJ (2016) Salinity Adaptation and the Contribution of Parental Environmental Effects in Medicago truncatula. *PloS one,* **11**: e0150350.

Munns, R., and M. Tester. 2008. Mechanisms of salinity tolerance. *Annual Review of Plant Biology*, **59**, 651–681.

Nei M, Gojobori T (1986) Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. *Molecular Biology and Evolution*, **3**, 418-426.

Nelson GC, Valin H, Sands RD, Havlík P, Ahammad H, et al (2014) Climate change effects on agriculture: economic responses to biophysical shocks. *Proceedings of the National Academy of Sciences of the U S A,* **111**, 3274-3279.

Nibau C, Gibbs DJ, Coates JC (2008) Branching out in new directions: the control of root architecture by lateral root formation. *New Phytologist*, **179**, 595-614.

Oldroyd GE, Downie JA (2006) Nuclear calcium changes at the core of symbiosis signalling. *Current Opinion in Plant Biology*, **9**, 351-357.

Orr HA (2005) The genetic theory of adaptation: a brief history. *Nature Reviews Genetics*, **6**, 119-127.

Passioura JB, Munns R (2000) Rapid environmental changes that affect leaf water status induce transient surges or pauses in leaf expansion rate. *Functional Plant Biology*, **27**, 941-948.

Penmetsa RV, Frugoli JA, Smith LS, Long SR, Cook DR (2003) Dual genetic pathways controlling nodule number in *Medicago truncatula*. *Plant Physiology*, **131**, 998-1008.

Penmetsa, R. A., and D. R. Cook. 2000. Production and characterization of diverse developmental mutants of *Medicago truncatula*. *Plant Physiology*, **123**, 1387–1398.

Peterson CA, Murrmann M, Steudle E (1993) Location of the major barriers to water and ion movement in young roots of *Zea mays* L. *Planta*, **190**, 127-136.

Pohill RM, Raven PH, Stirton CH (1981) Evolution and systematics of the Leguminosae. *Advances in legume systematics*, **1**, 1-26.

Pritchard JK, Di Rienzo A (2010) Adaptation--not by sweeps alone. *Nature Reviews Genetics*,**11**, 665-667.

Provin T, Pitt JL (2001) Managing soil salinity. Texas AgriLife Extension Service publication E-60. Texas A&M Univ. Publication. College Station, TX.

Purugganan MD, Fuller DQ (2009) The nature of selection during plant domestication. *Nature*, **457**, 843-848.

Qi J, Liu X, Shen D, Miao H, Xie B, et al (2013) A genomic variation map provides insights into the genetic basis of cucumber domestication and diversity. *Nature Genetics*, **45**, 1510-1515.

Qiu Q-S, Guo Y, Dietrich MA, Schumaker KS, Zhu J-K (2002) Regulation of SOS1, a plasma membrane Na+/H+ exchanger in *Arabidopsis thaliana*, by SOS2 and SOS3. *Proceedings of the National Academy of Sciences of the U S A*, **99**, 8436-8441.

Rajendran K, Tester M, Roy SJ (2009) Quantifying the three main components of salinity tolerance in cereals*. Plant Cell & Environment*, **32**, 237-249.

Rippke U, Ramirez-Villegas J, Jarvis A, Vermeulen SJ, Parker L, Mer F, Diekkrüger B, Challinor AJ, Howden M (2016) Timescales of transformational climate change adaptation in sub-Saharan African agriculture. *Nature Climate Change,* NCLIMATE2947.

Rogers C, Wen J, Chen R, Oldroyd G (2009) Deletion-based reverse genetics in *Medicago truncatula*. *Plant Physiology*, **151**, 1077-1086.

Rus A, Baxter I, Muthukumar B, Gustin J, Lahner B, Yakubova E, Salt DE (2006) Natural variants of AtHKT1 enhance Na+ accumulation in two wild populations of Arabidopsis. *PLoS genetics,* **2**: e210.

Savolainen O, Lascoux M, Merilä J (2013) Ecological genomics of local adaptation. *Nature Reviews Genetics*, **14**, 807-820.

Schlueter JA, Dixon P, Granger C, Grant D, Clark L, Doyle JJ, Shoemaker RC (2004) Mining EST databases to resolve evolutionary events in major crop species. *Genome,* **47**, 868-876.

Schmidt R, Mieulet D, Hubberten HM, Obata T, Hoefgen R, Fernie AR, Fisahn J, San Segundo B, Guiderdoni E, Schippers JH, et al (2013) Salt-responsive

ERF1 regulates reactive oxygen species-dependent signaling during the initial response to salt stress in rice. *The Plant cell,* **25**, 2115-2131.

Shi H, Ishitani M, Kim C, Zhu J-K (2000) The *Arabidopsis thaliana* salt tolerance gene SOS1 encodes a putative Na+/H+ antiporter. *Proceedings of the national academy of sciences of the U S A*, **97**, 6896-6901.

Slatkin, M. (2008) Linkage disequilibrium – understanding the evolutionary past and mapping the medical future. *Nature Reviews Genetics*, **9**, 477–485.

Tadege M, Wen J, He J, Tu H, Kwak Y, Eschstruth A, Cayrel A, Endre G, Zhao PX, Chabaud M, et al (2008) Large-scale insertional mutagenesis using the Tnt1 retrotransposon in the model legume Medicago truncatula. *The Plant journal : for cell and molecular biology,* **54**, 335-347.

Tajima F (1989) Statistical methods to test for nucleotide mutation hypothesis by DNA polymorphism. Genetics*, **123**,585-595.*

Tracy FE, Gilliham M, Dodd AN, Webb AA, Tester M (2008) NaCl-induced changes in cytosolic free Ca2+ in *Arabidopsis thaliana* are heterogeneous and modified by external ionic composition. *Plant Cell and Environment*, **31**, 1063-1073.

Turner TL, Bourne EC, Von Wettberg EJ, Hu TT, Nuzhdin SV (2010) Population resequencing reveals local adaptation of Arabidopsis lyrata to serpentine soils. *Nature Genetics*, **42**, 260-263.

Vance CP, Graham PH, Allan DL (2000) Biological nitrogen fixation. Phosphorus: a critical future need. In FO Pedrosa, M Hungria, MG Yates, WE Newton, eds, Nitrogen Fixation: From Molecules to Crop Productivity. Kluwer Academic Publishers, Dordrecht, The Netherlands, pp 506–514.

Via S, Lande R (1985) Genotype-environment interaction and the evolution of phenotypic plasticity. *Evolution*, **39**, 505-522.

Vitti JJ, Grossman SR, Sabeti PC (2013) Detecting natural selection in genomic data. *Annual Reviews Genetics*, **47**, 97-120.

Wahid A, Rasul E, Rao A (1999) Germination of seeds and propagules under salt stress. *In*: M. Pessarakli [ed.], Handbook of Plant and Crop Stress, Second Edition, Revised and Expanded. Pp. 153-168. Marcel Dekker, Inc. New York, New York. USA

Warschefsky E, Penmetsa RV, Cook DR, von Wettberg EJ (2014) Back to the wilds: tapping evolutionary adaptations for resilient crops through systematic hybridization with crop wild relatives. *American Journal of Botany*, **101**, 1791-1800.

Wheeler T, von Braun J (2013) Climate change impacts on global food security. *Science*, **341**,508-513.

Wilczek AM, Roe JL, Knapp MC, Cooper MD, Lopez-Gallego C, Martin L, Muir C, et al (2009) Effects of genetic perturbation on seasonal life history plasticity. *Science*, **323**, 930–934.

Wohlbach DJ, Quirino BF, Sussman MR (2008) Analysis of the Arabidopsis histidine kinase ATHK1 reveals a connection between vegetative osmotic stress sensing and seed maturation. *Plant Cell*, **20**, 1101-1117.

Xu X, Liu X, Ge S, Jensen JD, Hu F, et al (2012) Resequencing 50 accessions of cultivated and wild rice yields markers for identifying agronomically important genes. *Nature Biotechnology*, **30,** 105-111.

Young ND, Debellé F, Oldroyd GE, Geurts R, Cannon SB, Udvardi MK, Benedito VA, et al (2011) The Medicago genome provides insight into the evolution of rhizobial symbioses. *Nature*, **480**, 520–524.

Yu S, Wang W, Wang B (2012) Recent progress of salinity tolerance research in plants. *Russian Journal of Genetics*, **48**: 497-505.

Zhu H, Choi HK, Cook DR, Shoemaker RC (2005) Bridging model and crop legumes through comparative genomics. *Plant Physiology*, **13**, 1189-1196.

Zhu J-K (2002) Salt and drought stress signal transduction in plants. *Annual review of plant biology,* **53**, 247.

Zhu J-K (2003) Regulation of ion homeostasis under salt stress. Current Opinion in Plant Biology, **6**, 441-445.

Chapter 2.

Population genomic analysis of Tunisian *Medicago truncatula*

***This work is published as**:*

Friesen ML, **Cordeiro MA**, Penmetsa RV, Badri M, Huguet T, Aouani ME, Cook DR, Nuzhdin SV (2010) Population genomic analysis of Tunisian Medicago truncatula reveals candidates for local adaptation. *The Plant journal,* **63**, 623-635.

Matilde Ataide Cordeiro (MAC) joined this project after the microarray data was generated and therefore was involved in all the subsequent molecular and biological aspects of this chapter, while the first author (Maren Friesen) developed the mathematical and computational parts. Speciffically, MAC validated the array, defined the LD blocks and candidate genes assorting by habitat, and wrote the manuscript.

Chapter 2. Table of contents

Abstract

Genome-wide association studies rely upon segregating natural genetic variation, particularly the patterns of polymorphism and correlation between adjacent markers. A genome-scale polymorphism scan was performed using existing Affymetrix microarrays to facilitate association studies in the model legume *Medicago truncatula*. A method that uses a simple information-criteria algorithm to call polymorphism from microarray data without reliance on a reference genotype was developed and validated. Twelve inbred *M. truncatula* lines sampled from four wild Tunisian populations were genotyped and polymorphisms were found at approximately 7% of features, comprising 31 419 probes. Only approximately 3% of these markers assort by population, and of these only 10% differentiate between populations from saline and non-saline sites. Fifty-two differentiated probes with unique genome locations correspond to 18 distinct genome regions. Sanger resequencing was used to characterize a subset of maker loci and develop a single nucleotide polymorphism (SNP)-typing assay that confirmed marker assortment by habitat in an independent sample of 33 individuals from the four populations.

Genome-wide linkage disequilibrium (LD) extends on average for approximately 10 kb, falling to background levels by approximately 500 kb. A similar range of LD decay was observed in the 18 genome regions that assort by habitat; these LD blocks delimit candidate genes for local adaptation, many of which encode proteins with predicted functions in abiotic stress tolerance and are targets for functional genomic studies. Tunisian *M. truncatula* populations contain substantial amounts of genetic variation that is structured in relatively small LD blocks, suggesting a history of migration and recombination. These populations provide a strong resource for genome-wide association studies.

Introduction

The genus Medicago contains 83 species, including alfalfa (*Medicago sativa*), that are typically either tetraploid perennial or diploid annual species

(Lesins and Lesins 1979; Small and Jomphe 1989). *Medicago truncatula* is an exemplar of the annual diploid 'Medics' that occur spontaneously throughout the Mediterranean basin across a wide range of habitats. Because *M. truncatula* has been domesticated in Western Australia for use as forage in dry land agriculture, it is able to serve as a reference species for crop legumes, as well as a model species to understand the molecular genetic basis of legume processes. Of particular interest are the mutualistic interactions with nitrogen-fixing rhizobia and symbiotic mycorrhizal fungi (Heath and Tiffin 2007, 2009; Rangin et al. 2008; Chen et al. 2009; Gomez et al. 2009), properties shared by the majority of legume species. Medicago populations occur naturally across a broad range of stressful habitats, including serpentine soils in California, soils contaminated with heavy metals, drought-impacted regions of Mediterranean countries, and naturally occurring saline soils in North Africa and Western Europe. Salinity is a major factor affecting agricultural production worldwide, with one-fifth to one-third of irrigated agricultural land at risk (Tester and Davenport 2003). Thus, identifying genes involved in salinity adaptation in the model legume *M. truncatula* holds promise for enabling the improvement of economically important legume crops through translational genomics (Young and Udvardi 2009).

Phenotypic differences between populations can be due to drift or selection. In wild populations of *Arabidopsis thaliana*, flowering time and major genes known to underlie it vary with latitude in a manner consistent with evolutionary-ecological theory (Aranzana et al. 2005; Zhao et al. 2007; Wilczek et al. 2009). Spatial differences in selection can occur in response to a range of abiotic and biotic factors to produce locally adapted genotypes. For example, wild populations of the native North American grass *Andropogon gerardii* and their arbuscular mycorrhizal fungi are both locally adapted to the physico-chemical properties of their home soils as well as to one another (Johnson et al. 2010). Whole-genome scans can identify candidate genes for local adaptation, as in the case of *Arabidopsis lyrata* populations growing on serpentine and nearby non-serpentine soils (Turner et al. 2010).

The advent of reference genome sequence enables genome-wide scans of polymorphism in a wide range of organisms. Knowledge of genome-wide polymorphism has great utility for association mapping of ecological or agronomic traits of interest. The goal of association mapping is to identify genetic markers and candidate genes that are correlated with particular phenotypes; this can only succeed if enough markers are sampled to capture each linkage disequilibrium (LD) block within the population. Linkage disequilibrium in a given population depends on mating system, mutation, recombination, and migration rates, as well as patterns of selection (Slatkin 2008). Since individuals mate within their local population, isolated populations can diverge genetically from one another via genetic drift; within-population LD can be quite low but differences in allele frequencies between populations will lead to high global LD in the species via the Wahlund effect (Slatkin 2008). Population structure can also induce a large number of false positives in genome-wide association studies, since alleles that differ due to drift can become spuriously correlated with phenotypes that differ among populations (Aranzana et al. 2005; Rosenberg and Nordborg 2006).

Currently, there is little genome-wide information available about population differentiation and LD in *M. truncatula*. Since *M. truncatula* is highly selfing, with estimates from 95 to 99% (Chaulet and Prosperi 1994; Bonnin et al. 2001; Siol et al. 2008), LD is expected to be high. However, existing studies show that natural populations vary greatly in their genetic diversity, fine-scale spatial structure, and the number of loci that are in LD. Within French populations, the frequency of loci that exhibit significant LD ranges from 6% of 22 random amplified of polymorphic DNA (RAPD) loci (Aude) to 78% of 13 simple sequence repeat (SSR) loci (Salses 1999); within subpopulations of Aude, only 0–2% of loci show significant LD (Bonnin et al. 1996; Siol et al. 2007). Analysis of 10 Tunisian populations found on average only 6.3% of 18 loci to be in significant LD, with 30.5% in one population but 0% in others (Lazrek et al. 2009), while another study focused on four populations found on average 20.5% of 20 loci to have significant LD (Badri et al. 2007). A range-wide survey of 13 SSRs in 384 *M. truncatula* lines found 37.2% to

have significant LD, which could be due to population structure (Ronfort et al. 2006). Finally, sequencing of three regions spanning a symbiosis gene in 28 lines distributed around the Mediterranean showed that LD did not decay over 50 kb (De Mita et al. 2006). However, this study inferred that positive selection may have acted at this locus; if recent, this would cause extended LD due to a selective sweep.

To enable genome-wide association studies in the model legume *M. truncatula,* a genome level picture of polymorphism and LD is required. Here, existing Affymetrix microarrays are used to perform a genome-wide study of natural variation in *M. truncatula*. Data were obtained from 12 inbred genotypes sampled from four Tunisian populations, with two of the populations occurring on high-salinity soils. A new algorithm was developed to call single feature polymorphisms (SFPs) based on information criteria and traditional Sanger sequencing was used to validate the experiment and algorithm, in order to describe the patterns of: i) polymorphism, ii) population differentiation, and iii) LD. Finally, genetic regions that are consistently differentiated between saline-source and non-saline-source genotypes were identified. Genotyping an additional 33 individuals for six test loci confirms that these loci are significantly differentiated across habitats. These genome regions contain candidate genes for local adaptation to high-salinity habitats, including several genes with putative roles in abiotic stress responses.

Material and methods

Plant genotypes

Genotypes were collected in Tunisia in July 1999 by MEA with the assistance of Chedly Abdelly. At each of 10 sites, 30–100 pods of *M. truncatula* were collected at random in an area of radius 500 m. Pods were selected to maximize the variation in pod morphology at a site, thus minimizing the chance that pods from the same parent are sampled. Twelve lines per site were multiplied in

the greenhouse; although *M. truncatula* is highly selfing in nature, each line was selfed twice to lower remaining heterozygosity. The collection is housed at the CBBC (Centre of Biotechnology of Borj Cedria, Tunisia) and germplasm is available upon request (contact Dr Mounawer Badri). From among the 10 sites, four populations were selected, each in the north of Tunisia: TN1 (Enfidha), TN8 (Soliman), TN10 (Rhayet), and TN9 (Bulla Regia). Enfidha and Soliman sites have highly saline soil (8.65 and 4.40 g l−1) while Rhayet and Bulla Regia sites have low levels of salt (0.80 and 0.95 g l−1 (Lazrek et al. 2009) (Figure 1).

Microarray experiment

Genomic DNA was extracted from young leaves grown in growth rooms (ENSA Toulouse, France) using the DNeasy Plant Mini Kit (Qiagen, http://www.qiagen.com/). Genomic DNA was subsequently amplified using the Repli-g Midi (Qiagen). Amplified DNA samples from individual genotypes were extracted with phenol–chloroform, and the resulting purified DNA was fragmented by partial digestion with DNase, as follows: 10.5 µg DNA was dissolved in 30 µl double-distilled (dd) H2O, plus 4 µl One-Phor-All buffer (Amersham Biosciences 27-0910-02, http://www.gelifesciences.com), 0.2975 µl DNase (Promega M6101-RQ1, http://www.promega.com/), and 0.14 µl acetylated BSA (Invitrogen 15561-020, http://www.invitrogen.com/). The DNase digestion was allowed to proceed for 16 min at 37°C, followed by heat inactivation at 99°C for 15 min, and cooling to 12°C for 15 min. All reactions were carried out in a MJ Research thermocycler (Waltham, MD, USA). Three microliters of each DNA sample was visualized by ethidium bromide staining, following separation by gel electrophoresis on a 4% agarose SFR 0.5 TRIS-borate-EDTA (TBE; TRIS = 2-amino-2-(hydroxymethyl)-1,3-propanediol) gel (50 V for 120 min) with 10 bp and 100 bp DNA ladders (Promega). Samples that yielded bright smears from 20 to 100bp were selected and labeled with biotin by adding 2µl Biotin-N6-ddATP (Enzo 42809) and 3 µl RTdT (diluted to 15U µl−1; Promega M1875) and running the following program in a MJ Research thermocycler: 90 min at 37°C, 15 min at 99°C, 5 min at 12°C.

Labeled samples were frozen and delivered on dry ice to the Microarray Core facility at Children's Hospital, Los Angeles, USA, where they were hybridized to Medicago Genechips (Affymetrix) using the Affymetrix Hyb, Wash, and Stain Kit with the following hybridization cocktail: 125 µl 2 × hybridization mix, 4.17 µl control oligo B2, 12.5 µl 20 × hybridization controls, 25 µl DMSO, labeled target DNA (9.585 µg), ddH2O to 250 µl, and wash protocol FS450_0001. A single array was hybridized with each individual genotype's DNA to maximize the number of individuals assayed for a given cost. While this does not enable the estimation of technical error in genotype calls, the resulting data are sufficient to identify many SFPs that are replicated at the population level, namely those that occur in two or more individuals.

Validation of the SFPs

Genomic DNA was extracted from 45 individuals from four Tunisian populations using the DNeasy Plant Mini kit (Qiagen). Six loci around Affymetrix probes and seven COS markers were amplified in the genotypes used for the microarray experiment. Primer3Plus software was used to design primers to amplify 700–1000 bp around the Affy probes *M. truncatula*.21891.1.S1 (F-tatcagaggaagctgcaaaagc; R-tcagcctcttcatcaatgtcc), *Mtr*.48956.1.S1 (F-ttgacagctacaacaaggaagc; R-gtaacctttctcccaaagttgc), *Mtr.*42442.1.S1 (F-ctcttccggacaagtgttcacc; R-cacaagccacaaacacataagagc), *Mtr.*20573.1.S1 (F-tctctactagttccctctctattagttcc; R-cagtaaaaatcgcgctacgg), *Mtr*.20569.1.S1 (F-tctgccatagccatgtttcg; R-aaccggtcatcttacacaacg), and *Mtr*.8358.1.S1 (F-taaacccatcagtcccatcacc; R-tgtagatttgttgttggcaagg). The COS markers 1433P, AAT, AGT, CALTL, CNGC4, SHMT, and SUSY were selected (Choi et al. 2006). The PCR reactions were performed in a Tetrad 2 Thermal Cycler PTC-0240G using Takara's (http://www.takara-bio.com/) Ex Taq® DNA Polymerase (3 min at 95°C; 40 cycles of 30 sec 95°C, 30 sec 55°C (except for *Mtr.*20573.1.S1 and *Mtr*. 20569.1.S1 where 57°C and 60°C were used, respectively), 90 sec 68°C, and 3 min 68°C. Amplicons (3 µl) were visualized by gel electrophoresis on a 1.2%

agarose 0.5 TBE gel running at 100 V for 30–40 min with All Purpose Hi-Lo® DNA Marker (Bionexus, http://www.bionexus.net/). The PCR products were cleaned using 0.5 U shrimp alkaline phosphatase (SAP) (USB, http://www.usbweb.com/), SAP buffer (USB), and 0.2 U Exo I (USB) (30 min, 37°C), and were sequenced using an ABI 3730 XL capillary sequencer (Applied Biosystems, http://www.appliedbiosystems.com/). Sequences were analyzed using CodonCode Aligner 2.0.6 and mapped into Mt3.0 using BLAST. Genes around the sequenced Affymetrix probes were assessed using the IMGAG Genome Annotation Version 3.0.

The identified SNPs were surveyed in the 45 individuals from the four TN populations using the ABI Prism® SNaPshot® Multiplex kit: 1 µl SNaPshot mix, 2 µl of clean PCR product and 2 pmol of extension primers for *Mtr*.21891.1.S1 (TTTGAAGGAATCTGCACC), *Mtr*.48956.1.S1 (GTTGGACG- -TGGTGGCGAGCTTA), *Mtr*.42442.1.S1 (TCTACTTGCTTGTTGTTC), *Mtr*.20573.1.S1 (AAAATGCAACTGGAAATAAGAC), and *Mtr*.8358.1.S1 (TCTTTCACTATTACTTCAACTA). The extension reaction consisted of 40 cycles of 10 sec 95°C, 5 sec 50°C, and 30 sec 60°C. The individuals were genotyped using an ABI 3730 XL capillary sequencer (Applied Biosystems) and the polymorphisms were analyzed in the GeneMapper® software v.3.7 (Applied Biosystems). Fisher's exact test Bonferroni corrected on the 33 independent inbred lines was performed to assess the statistical significance of the distribution of the polymorphisms with respect to habitat type.

Microarray data analysis

We first examined the .jpeg for each array for defects and found none. All statistics were performed in R 2.6.2 (R Team 2009); code is available upon request. The Affymetrix® GeneChip Medicago Genome Array contains 673 880 probe pairs in total; 560 206 of these (50 902 probe sets) are specific to *M. truncatula* with an additional 1896 probe sets specific to *M. sativa*, 8305 probe sets specific to the bacterial symbiont *Sinorhizobium meliloti*, and 14 control probe sets.

A comparison of raw perfect-match (pm) probe intensities and mismatch (mm) probe intensities revealed that 0.8608 of targets hybridized more strongly to the pm probe; when only *M. truncatula* specific probe pairs were considered the pm intensity was greater than the mm intensity for 0.9148 of targets, indicating a substantial amount of signal in our data. Examining the log pm intensity distribution did not reveal large differences between arrays. We used two standard background corrections implemented in Bioconductor (Gentleman et al. 2004), rma and mas. Rma ('robust multi-array averaging') models each pm intensity as having signal and error components, while mas (Affymetrix's 'Micro Array Suite') performs a spatial correction for each array by considering the lowest-intensity probes in each grid. We performed two standard normalizations: slide mean normalization, to scale each slide to have the same mean intensity, and quantile normalization, which scales each slide to have the same intensity distribution. The six potential combinations (raw, rma, mas correction by slide mean, quantile normalization) were compared. Mas correction changed the second peak corresponding to low hybridization intensities into a shoulder that obscured differences between strong and weak hybridization signals, so it was not used in further analyses. Raw and rma histograms had similar shapes to one another, with clear peaks for normal hybridization intensities and weak intensities that presumably correspond to sequence divergence.

For each data processing, more markers are called for quantile normalization than for slide-mean normalization with many of these markers present in only two of the 12 genotypes. Since low-frequency markers will deflate estimates of LD, we focus on the slide-mean normalized raw data for the analyses in this paper.

Algorithm for determination of SFPs

We develop a new algorithm to determine whether the 12 individuals are polymorphic at a site as reflected by Affymetrix probe hybridization intensity. We presume that each individual is homozygous, since *M. truncatula* is highly selfing in

nature and wild plant genotypes were further selfed for at least two generations in the greenhouse. Our method uses simple information criteria to compare two types of models: Model1 where a probe does not cover a polymorphism and Model2 where a probe does detect polymorphism and is therefore a marker. Information criteria are used in model selection to balance the explanatory power of each model against the number of model parameters, thus identifying which model is closer to the truth. Since this is a non-parametric procedure without established significance cutoffs, we use simulation to determine the significance threshold.

For each probe, we first order the log-transformed hybridization intensities I from lowest to highest (I1, I2, …, I12). We then consider all possible two-way splits that divide the data into contiguous sets of values, where each subset contains at least two observations. For 12 observations there are 10 possible splits [(I1, I2),(I3, …, I12)]; [(I1, I2, I3),(I4,…, I12)]; …; [(I1, …, I10), (I11, I12)]. We next calculate the likelihood of the data under Model1, where the data are drawn from a single Normal (μ, θ) model, where μ = mean (I) and θ = standard deviation (I). The likelihood is then the product of probabilities: $L(I \mid Model_1) = \prod_{k=1}^{12} \Pr(I_k \mid Normal(\mu,\theta))$.

For each of the ten two-group models, Model2_i with i = 1, …, 10, we assume that the data are drawn from two groups. For example, under Model2_1, group g1 is Ig1 = (I1, I2) and group g2 is Ig2 = (I3, …, I12). Each group gi is assumed to have a Normal (μgi,θgi) distribution, where μgi = mean (Igi) and θgi = standard deviation (Igi). The likelihood of the data under this model is:

$$L(I \mid Model_{2\_1}) = \prod_{k=1}^{2} \Pr(I_k \mid Normal(\mu_{g1},\theta_{g1})) \prod_{k=3}^{12} \Pr(I_k \mid Normal(\mu_{g2},\theta_{g2})).$$

Next, we use Akaike's AIC with the small sample bias correction term (Burnham and Anderson 2002): *AICc* = *-2 Ln*(*Likelihood*(*Data*|*Model*)) + *2 k* + *2 k* (*k+1*)/(*n-k-1*), where k is the number of parameters in Model and n is the number of observations. We assign the model with the lowest AICc value to have score of 0

and subtract this AICc value from the AICc value for the other models to obtain deltaAICc values. Models with higher AICc values and thus larger deltaAICc values should be rejected; we use simulation to determine the significance threshold of deltaAICc.

To formulate a null distribution for the deltaAICc test statistic, we simulate extensively under the no-polymorphism model, i.e. a single normal distribution having a mean and standard deviation drawn from the empirical distribution of means and standard deviations. This ensures that the range of variances in the simulation reflects the variances present in the data set. We simulated 5 million data sets and calculated the deltaAICc values for Model1, Model2_1, …, Model2_10. In order to set the significance threshold, we employ false discovery rate (FDR) criteria. The FDR is the frequency of false rejections of the null hypothesis within all null hypothesis rejections. We calculate it by dividing the expected frequency of false positives, as seen in the simulated null distribution, by the number of probes in the real data that exceed each threshold and then selecting the threshold giving the desired FDR. For example, setting the deltaAICc threshold to 28.0812 gives on average 1568 (out of 560 206) simulated null model probes that have higher deltaAICc values while 31 419 of the 560 206 empirical probes have deltaAICc≥ 28.0812. Since 1568/31 419 = 0.05, this deltaAICc threshold corresponds to 5% FDR. Affymetrix probes whose hybridization data support a two-group model with deltaAICc ≥ 28.0812 are referred to as ‘markers’ or equivalently ‘SFPs’.

Analysis of population genetics and structure

Population genetic parameters were calculated using DNAsp v. 5.0 (Librado and Rozas 2009) running under ‘wine’ on a Mac OSX PC. Only polymorphic sites supported by both strands of Sanger sequencing were included. Results are presented in Table 5. To explore population structure, the Bayesian clustering program STRUCTURE v. 2.3 (Pritchard et al. 2000) was compiled and run on a 64-bit node. We ran each number of clusters (K) three times. As

STRUCTURE has a maximum marker number of 10 000 and further requires that markers be unlinked, we sampled one marker per 50 kb (genome-wide LD is approximately 0.2 at this distance, see Figure 3) for a total of 3429 markers. For K = 1, 2, 3, 4, 5, and 6 the ln probabilities of the data were as follows: run1, −48445.5, −44279.7, −41359.8, −39861.5, −37868.6, −37333.1; run2, −48428, −44341.3, −41357.8, −39438.7, −37759.5, −37070.1; run3, −48435.2, −44353.4, −41305.3, −39497.5, −37787.1, −37744.5. As these numbers are congruent between runs, we conclude that the model converged. The 'optimal' number of clusters is five, since the change in probability between four and five is much greater than between five and six (Pritchard et al. 2010). However, since 12 is a small number of individuals relative to the number of groups tested we do not have a high degree of confidence in these clusters. Individual assignment was broadly consistent between runs; representative assignments plotted with 'distruct' (Rosenberg 2004) are shown in Figure 2b.

Results

Microarray genotyping

Inbred lines of *M. truncatula* sampled from four wild populations in Tunisia were used, as shown in Figure 1. Genomic DNA from 12 individual genotypes (three per population) were hybridized to existing Affymetrix microarrays to perform a genome-wide study of natural variation in this organism. DNA from a single inbred genotype was hybridized to each array so that haplotypes could be inferred. Oligonucleotide microarrays are being used as a cost-effective technique for simultaneously interrogating hundreds of thousands to millions of genome positions simultaneously. Sequence differences lead to altered hybridization intensity of particular probes, termed SFPs. The first of these studies was performed in yeast and compared recombinants to their two parental genotypes at the observed 3714 SFPs (Winzeler 1998). Similarly, in A. thaliana two reference lines were hybridized to replicate microarrays to identify 3806 SFPs (Borevitz et al.

2003). A larger sample of 23 accessions uncovered 77 420 SFPs relative to the reference genotype Columbia (Borevitz et al. 2007). However, in natural populations such as those that we are investigating there is no 'reference' hybridization with which to compare hybridization intensity. Studies thus far typically compare between subpopulations or other pre-defined groups to discover segregating SFPs (Turner et al. 2005, 2008a,b). We extend this approach by developing a new algorithm that uses information criteria to identify probes in our sample that give polymorphic hybridization signal intensities and thus are likely to contain segregating sequence polymorphisms.

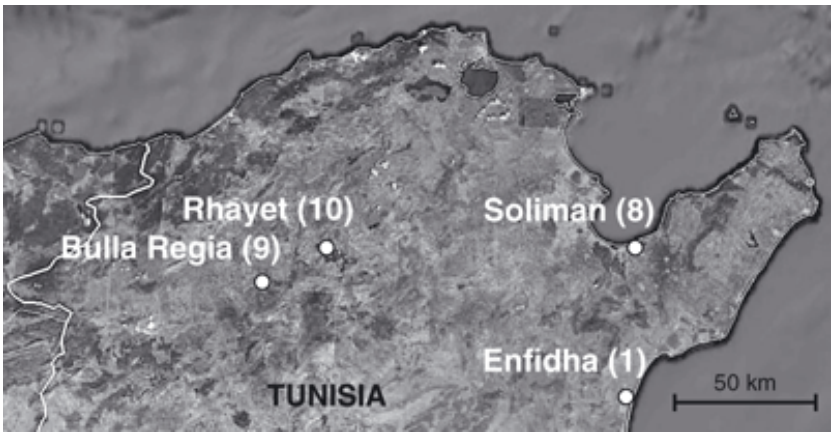


Figure 1. Map showing sampling locations for *Medicago truncatula* germplasm. Sites 1 (Enfidha) and 8 (Soliman) are highly saline, while sites 9 (Bulla Regia) and 10 (Rhayet) have very low levels of soil salinity.

Statistical identification of marker probes

This experiment uses unknown genotypes from natural populations, and therefore the loci where individuals differ from one another are not know *a priori*. Indeed, since SSR analysis suggests that gene flow is common between populations (Lazrek et al. 2009), identifying loci by their population-level divergence is expected to drastically misestimate the pattern of polymorphism. The first step is to determine whether the hybridization intensities observed for each probe reflect one allele or two, then LD and population-level patterns can be determined after assigning individuals' genotypes at loci where two alleles are believed to be present. While sophisticated clustering and partitioning algorithms

exist, these tend to be computationally expensive, particularly when hundreds of thousands of probes need to be considered. Thus, a simple algorithm based on information criteria was developed (Akaike's an information criterion, AICc; Burnham and Anderson 2002) to decide whether a two-group model, where the values of each locus are drawn from two distinct distributions, fits the data substantially better than a model where each intensity at a locus is drawn from the same underlying distribution. All computations were performed using R (R Team 2009); code is available upon request.

Polymorphism rate, LD, population structure

At 5% FDR the slide-mean normalized raw data contains 31,419 probes called as markers by our algorithm. Clustering the 12 haplotypes with the unweighted pair group method with arithmetic mean (UPGMA) shows evidence of population structure, since individuals from the same population tend to be more similar to one another, as seen in Figure 2(a).

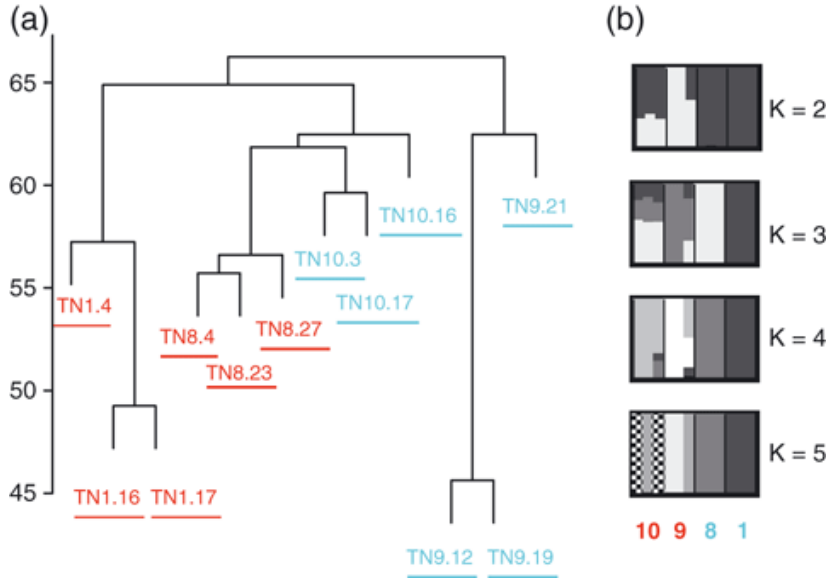


Figure 2. Population structure in 12 Medicago truncatula genotypes. (a) Haplotype clustering based on the inferred markers, using the unweighted pair group method with arithmetic mean (UPGMA). Genotypes from saline sites are red and those from non-saline sites are blue. (b) STRUCTURE analysis of a subset of the single feature polymorphisms (SFPs) showing ancestral population assignments assuming historical populations K = 2, 3, 4, and 5.

Running ‘Structure’ (Pritchard et al. 2000) on a subset of the data suggests that five groups best fit the data, but with only 12 individuals we regard this clustering method as highly preliminary since groups contain only two or three individuals (Figure 2b). Nonetheless, ‘Structure’ provides evidence against the one- and two-population models, which suggests that there is complex population structure in these data. However, despite these indications of population structure it is important to note that the vast majority of SFPs do not assort along population subdivisions: only 938 (3% of probes) are structured by the four populations, and just 90 probes assort with saline habitat. For comparison, 30 probes discriminate populations 1 and 9 from 8 and 10, and 21 probes discriminate 1 and 10 from 8 and 9. Analysis of polymorphisms in an independent set of conserved orthologous sequences (COSs) (Choi et al. 2006) confirms that the majority of the genome is not structured by population.

All probes on the Affymetrix array were mapped to the Mtr3.0 assembly of the genome using ‘bowtie’ (Langmead et al. 2009). There were 301 055 probes that have a perfect unique hit in Mtr3.0; of these 20 208 are called as SFPs at 5% FDR, giving a polymorphism rate of 6.7%. Note that these polymorphisms may include nucleotide changes, indels, and copy-number variants. To investigate patterns of LD across the genome, the pairwise correlation coefficient (r2) between SFPs across all 12 individuals was computed. As seen in Figure 3, the maximum value is approximately 0.8 for markers within 1 kb of one another; this decays to approximately 0.4 on average by 10 kb, to approximately 0.2 by 100 kb, and to background levels by approximately 500 kb. Useful LD, i.e. correlations that would enable markers to tag causal single nucleotide polymorphisms (SNPs), thus extends on average approximately 10 kb in these populations.

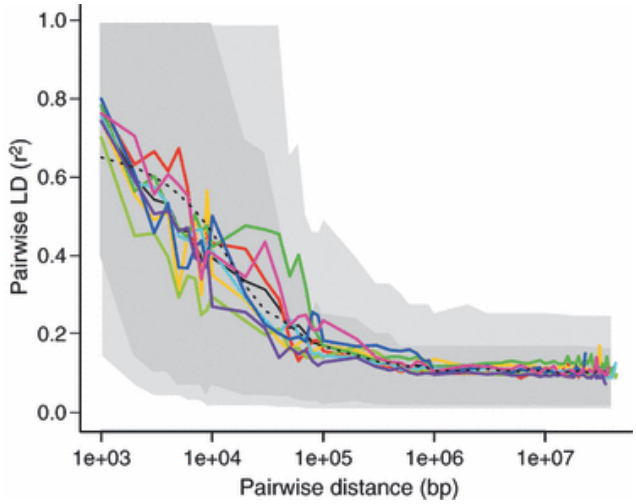


Figure 3. Pairwise correlation between markers averaged over markers within a given distance.

Each chromosome is depicted by a different line. Linkage is half-decayed (from 0.8 to 0.4) at approximately 10 kb. Dark shading shows the 20 to 80% quantiles across all markers, while light shading shows the 10 to 90% quantiles. The dotted line shows a second-degree lowess fit across all data with span parameter set to 0.01; linkage disequilibrium (LD) is fully decayed between markers approximately 500 kb apart.

Comparison between new algorithm and the t-test/q-value approach

Typically, studies of population differentiation with microarray genotyping without a reference employ t-tests between habitat types to determine which markers are differentiated. This test is appropriate when populations are pooled, since the t-test then tests the difference in allele frequencies (Turner et al. 2008a,b). However, when applied to hybridization data from individuals the t-test confounds marker detection and habitat assortment because non-differentiated markers are not identified (e.g. Turner et al., 2005); to the best of our knowledge, the statistical properties of this scenario have not been investigated. Our approach represents an alternative that yields information about overall diversity and LD across the genome in addition to habitat differentiation.

Since our AICc method is new, it was compared with the t-test approach by computing t-tests between the log-transformed hybridization intensities of individual genotypes from two saline habitats and the two non-saline habitats. After correcting for multiple testing by converting the resulting P-values into q-values, the AIC approach calls fewer markers that assort by habitat. The intersection of the AICc and t-test approaches is relatively high, with the intersection containing 83 and 71% of each at a FDR of 0.05.

Analysis of regions containing markers that assort by habitat

Markers that assort by habitat across multiple populations are potential evidence of selection operating on genes that confer adaptation to the corresponding habitat. Although this current dataset is limited by marker density and analysis of a small set of individual genotypes, it is possible to estimate the extent of local LD and thus circumscribe a set of candidate genes for adaptation to saline habitats. A total of 52 SFPs assort with habitat and are mapped in Mt3.0; these cluster in 18 genome regions.

Table 1. Overview of the 18 genomic regions that are differentiated between saline and non-saline habitats.

Reg	Chr	Coordinates	Size (Kbp)	Probes	Markers	Assort with habitat	Amp	Candidate genes
1.1	1	chr1:17066007..17130116	64.1	187	29	*Mtr*.21891.S1_at:635:557; *Mtr*.48008.S1_at:674:567	Chr1	13
2.1	2	chr2:18976699..18994201	2.2	26	2	*Mtr*.23081.S1_at:1046:237; *Mtr*.23081.S1_at:421:337	Chr2.1 Chr2.2	5
2.2	2	chr2:21927619..21970724	43.1	68	6	*Mtr*.20573.S1_at:839:803; *Mtr*.20569.S1_at:603:565	Chr3	6
3.1	3	chr3:21844852..21847013	2.2	19	3	*Mtr*.8358.S1_at:427:1115		1
3.2	3	chr3:22351489..22401448	50.0	67	2	*Mtr*.10504.S1_at:798:1029		12
4.1	4	chr4:5884997..5913860	28.9	42	10	*Mtr*.33441.S1_at:147:233		8
4.2	4	chr4:28624487..28676670	52.2	114	1	*Mtr*.37707.S1_at:1145:225		10
5.1.1	5	chr5:8617978..8683602	65.6	104	13	*Mtr*.48956.S1_at:865:623; *Mtr*.48956.S1_at:763:453; *Mtr*.48956.S1_at:1070:921; *Mtr*.48956.S1_at:464:815; *Mtr*.48956.S1_at:586:901; *Mtr*.48956.S1_at:846:457; *Mtr*.17919.S1_at:414:683	Chr5.1.1 Chr5.1.2	15

Reg	Chr	Coordinates	Size (Kbp)	Probes	Markers	Assort with habitat	Amp	Candidate genes
5.1.2	5	chr5:8779956..8845569	65.6	48	1	*Mtr*.42442.S1_at:633:1151		11
5.2	5	chr5:9272181..9296836	24.7	22	4	*Mtr*.43580.S1_at:1008:607; *Mtr*.43580.S1_at:1012:429; *Mtr*.43580.S1_at:442:351; *Mtr*.43580.S1_at:394:669		7
5.3	5	chr5:11059938..11066007	6.1	21	2	*Mtr*.11801.S1_at:95:849		2
5.4	5	chr5:30719205..30723486	4.3	11	6	*Mtr*.18994.S1_at:596:269; *Mtr*.18994.S1_at:283:687		1
6.1	6	chr6:10237963..10289680	51.7	85	23	*Mtr*.2455.S1_at:223:475; *Mtr*.2455.S1_at:369:1055; *Mtr*.2455.S1_at:14:455; *Mtr*.2455.S1_at:188:447; *Mtr*.2455.S1_at:970:139; *Mtr*.2455.S1_at:230:469; *Mtr*.2455.S1_at:71:559; *Mtr*.11624.S1_at:590:975; *Mtr*.11624.S1_at:763:435; *Mtr*.6977.S1_at:901:537; *Mtr*.6977.S1_at:97:989; *Mtr*.6977.S1_at:690:1091; *Mtr*.6977.S1_at:235:655; *Mtr*.6977.S1_at:84:1057; *Mtr*.6977.S1_at:967:359; *Mtr*.6977.S1_at:310:235; *Mtr*.6977.S1_at:727:319		10
6.2	6	chr6:10662451..10663318	0.86	10	10	*Mtr*.48980.S1_at:668:433		2
6.3	6	chr6:11332555..11332783	0.23	6	5	*Mtr*.22401.S1_at:135:569		1
6.4	6	chr6:14923810..14927210	3.4	29	11	*Mtr*.30229.S1_at:386:321		1
7.1	7	chr7:22444707..22448619	3.9	20	1	*Mtr*.14073.S1_at:689:685		2
7.2	7	chr7:26351771..26373082	21.3	84	1	*Mtr*.12942.S1_at:303:1091		9
8.1	8	chr8:868185..892217	24.0	31	9	*Mtr*.32725.S1_at:117:409; *Mtr*.32725.S1_at:678:723; *Mtr*.32725.S1_at:902:1077; *Mtr*.35740.S1_at:805:725		9

A region is defined by the presence of one or more markers that assort by habitat, flanked by two consecutive marker probes that have low LD with the assorting focal probes (r2 < 0.5). When there are habitat-assorting markers within 100 kb, these were considered to be a single region even if there are a few low-LD markers present in the region (Table 1). These regions range in size from 229 bp (region 6.3) to 65.62 kb (region 5.1.1) with an average size of 27 kb and containing

from 1 to 15 genes (average of 6.6 genes per region). LD decay around these regions is shown in Figure 4 and the complete list of 125 candidate genes is given in Table S1.

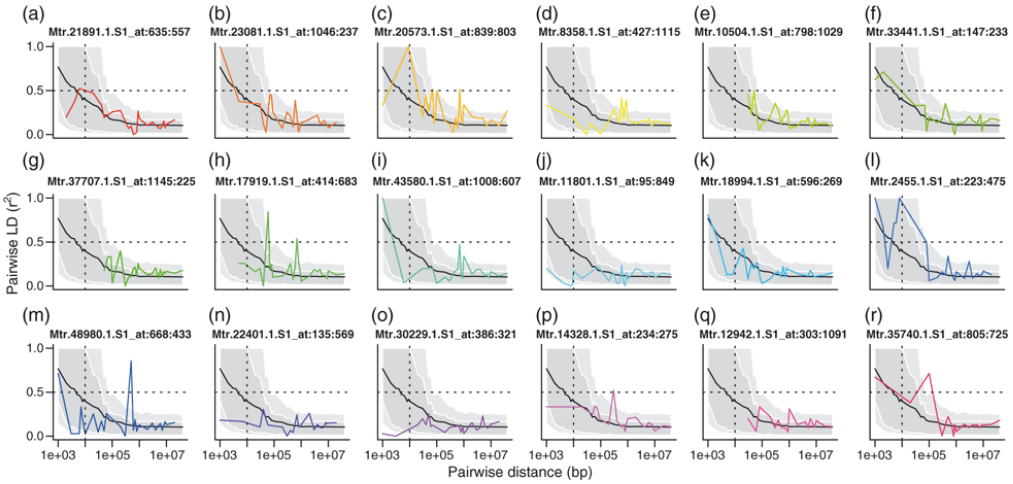


Figure 4. Linkage disequilibrium (LD) decay around focal single feature polymorphisms (SFPs) that assort by habitat. The first habitat-assorting SFP at a locus is given in the title. The black line gives genome-wide average LD, while the gray regions are as in Figure 3 (dark gray shows the 20 to 80% quantiles across all markers, while light gray shows the 10 to 90% quantiles).

Although probe density and polymorphism rates constrain detail in the analysis, it is evident that rates of LD decay around candidate regions are reflective of our genome-wide estimates (Figure 4). Thus, correlations between the focal and linked marker probes typically decay to r2 < 0.3–0.5 within 10 kb. In four regions (Figure 4, panels h, l, m and r), significant LD is apparent between probes that are separated by distances considerably greater than 10 kb (up to 100s of kb) with intervening regions of low LD. Such instances might arise from either biological (i.e. mutation or recombination) or technical (array design or sample size) circumstances, as discussed in the example below. In support of the former possibility, two of these regions (Figure 4, panels l and m) occur on chromosome 6, a genome segment that is notoriously rich in fast-evolving NBS-LRR disease resistance genes (e.g. Zhu et al. 2002). NBS-LRR genes evolve by processes that can involve high rates of recombination by unequal crossing over and gene

conversion, as well as diversifying selection, all of which are factors that could underlie the observed patterns of LD. Interestingly, two additional regions of chromosome 6 also include probes that assort by habitat (shown in Figure 4, panels n and o), suggesting that this linkage group, which is apparently highly dynamic and not conserved in other analyzed legume genera (e.g. Choi et al. 2006), may have been under selection in saline habitats.

It is important to note that the saline habitats sampled here differ in many aspects from the non-saline habitats in terms of soil characteristics and vegetation composition, any of which could potentially mediate habitat-specific selection. The NBS-LRR genes mentioned above could themselves be targets of selection, but are more likely to be linked to such targets. While it is premature to view this differentiation as final evidence of selection, differentiated SFPs represent candidate genomic regions for selection based on habitat type.

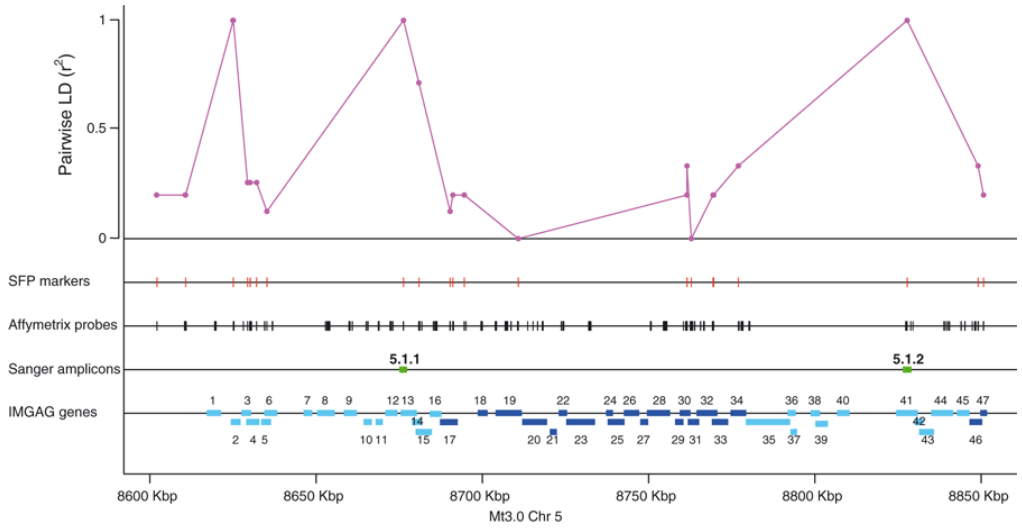


Figure 5. Regions 5.1.1 and 5.1.2 that are differentiated between saline and non-saline habitats in the four Tunisian populations.Top: Pairwise correlation coefficient between each marker with habitat. Bottom rows show the positions of: (i) SFPs, (ii) all uniquely mapped Affymetrix probes, (iii) the Sanger sequences regions used for validation, and (iv) International Medicago Genome Annotation Group (IMGAG) gene predictions. The IMGAG genes in cyan are considered to fall within the candidate regions, while those in dark blue are considered outside the regions. Gene numbers in this figure correspond to genes listed in Table 2.

As an example, a single 250 kb genome interval from chromosome 5 that contains eight focal marker probes that assort by habitat separated by non-assorting SFPs is presented (Figure 5). Whether this region is actually under selection by habitat-mediated factors awaits future confirmation, but it is as a strong candidate is presented as an example of processes that could be acting. The left region (5.1.1) contains seven markers that assort with habitat, but these are interrupted by five marker probes in low LD with the assorting SFPs. The right-most region (5.1.2) contains a single assorting SFP, separated from region 5.1.1 by 11 SPFs in low LD with the focal SFPs.

The eight habitat-assorting SFPs detailed in Figure 5 could be the product of distinct selection events separated by historical recombination, or a single region with a common selection history and relatively unstable intervening genome features. In either case, the annotation of predicted genes obtained from the International Medicago Genome Annotation Group (IMGAG) provides a starting point for estimating gene function and narrowing the list of candidate genes in these regions for subsequent functional analysis. Table 2 lists the candidate genes from the region detailed in Figure 5. Several of the deduced proteins have potential roles in physiological and/or regulatory adjustments to abiotic stress, as well as in biotic stress responses. Ultimately it will be important to narrow the candidate genome intervals by more precise genome characterization on larger numbers of individuals, e.g. using second-generation sequencing methodologies, and to test candidate gene function by means of reverse genetic and/or biochemical characterization.

Table 2. Annotation of candidate genes within the chromosome 5 region 5.1, shown in Figure 5. Genes shaded in grey are either flanked by or border on probes that assort with saline habitats, with corresponding linkage disequilibrium (LD) values of flanking probes indicated. Gene numbers (Gene No.) correspond to numbering in Figure 5.

Gene nr	Gene ID	LD	Location	IMGAG annotation
1	5g022170	0.20-1.00	8617978	Armadillo
2	5g022180	0.20-1.00	8625089	hypothetical protein
3	5g022190	1.00-0.26	8628391	hypothetical protein
4	5g022200	0.26-0.26	8629743	Peptidase C48

Gene nr	Gene ID	LD	Location	IMGAG annotation
5	5g022210	0.26-1.0	8634233	hypothetical protein
6	5g022220	1.00-1.00	8635311	FAR1
7	5g022230	1.00-1.00	8647104	HAT dimerisation
8	5g022240	1.00-1.00	8651184	Oxidoreductase FAD/NAD(P)-binding
9	5g022250	1.00-1.00	8659225	Thaumatin
10	5g022260	1.00-1.00	8665015	hypothetical protein
11	5g022270	1.00-1.00	8668657	Short-chain dehydrogenase/reductase SDR
12	5g022280	1.00-1.00	8671722	Thaumatin
13	5g022290	1.00-0.71	8676311	hypothetical protein
14	5g022300	1.00-0.71	8679578	hypothetical protein
15	5g022310	1.00-0.71	8680577	Nodulin-like
16	5g022320	0.71-0.13	8684964	Rhodanese-like
17	5g022330	0.71-0.13	8687859	Mlo-related protein
18	5g022340	0.20-0.00	8699479	Cyclin-like F-box; F-box interaction domain
19	5g022350	0.20-0.00	8704849	von Willebrand factor
20	5g022360	0.00-0.20	8712687	Senescence-associated
21	5g022370	0.00-0.20	8720896	hypothetical protein
22	5g022380	0.00-0.20	8723815	hypothetical protein
23	5g022390	0.00-0.20	8725953	Heavy metal sensor kinase
24	5g022400	0.00-0.20	8738079	NB-ARC
25	5g022410	0.00-0.20	8738399	Leucine-rich repeat
26	5g022420	0.00-0.20	8743465	Haem peroxidase
27	5g022430	0.00-0.20	8748205	hypothetical protein
28	5g022440	0.00-0.20	8750392	Protein of unknown function DUF630
29	5g022450	0.00-0.20	8758684	hypothetical protein
30	5g022460	0.00-0.20	8760270	Peptidase C1A
31	5g022470	0.33-0.00	8762509	hypothetical protein
32	5g022480	0.00-0.20	8765278	3-dehydroquinate synthase
33	5g022490	0.20-0.33	8769738	D-galactoside/L-rhamnose binding
34	5g022500	0.20-0.33	8775550	Ras GTPase
35	5g022510	0.33-1.00	8779956	Vacuolar protein sorting-associated protein 35
36	5g022520	0.33-1.00	8792564	Heavy meta transport/detoxification protein
37	5g022530	0.33-1.00	8793254	hypothetical protein
38	5g022540	0.33-1.00	8799416	Lipolytic enzyme
39	5g022550	0.33-1.00	8800774	Lipolytic enzyme
40	5g022560	0.33-1.00	8807604	Lipolytic enzyme
41	5g022570	0.33-1.00	8825347	Uncharacterized Cys-rich domain
42	5g022580	1.00-0.33	8830696	hypothetical protein
43	5g022590	1.00-0.33	8831951	Splicing factor motif WD40-like

Gene nr	Gene ID	LD	Location	IMGAG annotation
44	5g022600	1.00-0.33	8835936	Diaminopimelate decarboxylase
45	5g022610	1.00-0.33	8843627	SKP1 component
46	5g022620	0.33-0.20	8847219	Cupin
47	5g022630	0.33-0.20	8850664	hypothetical protein

Validation of the array results by resequencing

To validate the algorithm and microarray data, and to potentially extend the correlation between habitat of origin and molecular polymorphisms, chosen loci were analyzed across a larger number of individual genotypes. Initially, Sanger resequencing was used to characterize molecular variation underlying different hybridization intensities, focusing on six genome regions (including regions 5.1.1 and 5.1.2, detailed in Figure 5) that assort between saline and non-saline populations. A total of 14 differentiated features were analyzed and polymorphisms that correlate with probe hybridization intensities were identified. The nature and location of the polymorphisms corresponding to the 12 resequenced assorting probes are described in Table 3. An additional 39 non-polymorphic probes were also confirmed in the sequencing data; all validated probes are described in Table S2. One SFP did not possess polymorphism in the Sanger data, for a false positive rate of 7%. Resequencing revealed additional polymorphism at 33% of probes; this high false negative rate is expected since probes are not sensitive to polymorphism near their edges. Of the polymorphic SFPs, one locus was shown to be tri-allelic and 7/12 genotypes at this locus were miscalled by the SFP approach. Excluding this special case, only two of the remaining 92 sequenced alleles were miscalled (2%) with 13 alleles ambiguous due to lack of PCR amplification.

Table 3. Sequence polymorphism associated with single feature polymorphisms (SFPs) for the six loci validated by Sanger sequencing

Chr	Amp	Probeset	Type	Position	S-like	NS-like	Type	Gene ID
1	Chr_1	*Mtr*.21891.S1	SNP	chr1:17088612	G	A	exon/silent	Medtr1g086410
2	Chr_2.1	*Mtr*.20573.S1	indel	chr2:21934468..4962	+	-	intergenic	Medtr2g095720 - Medtr2g095730
2	Chr_2.2	*Mtr*.20569.S1	indel	chr2:21942438..3217	+	-	intergenic	Medtr2g095740 - Medtr2g095750
3	Chr_3	*Mtr*.8358.S1	SNP	chr3:21846708	G	A	intron	Medtr3g092650
5	Chr_5.1.1	*Mtr*.48956.S1	SNP	chr5:8676302	C	G	UTR	Medtr5g022290
5	Chr_5.1.2	*Mtr*.42442.S1	SNP	chr5:8827741	T	C	exon/silent	Medtr5g022570

To extend the correlation between genetic polymorphism and habitat of origin, a larger set of 33 individuals that were derived from the same four Tunisian populations represented by the original 12 genotypes was analyzed. For each of the six genome regions, one resequenced polymorphism was selected for analysis; towards this end, SNP polymorphisms were converted to a simple allele-specific oligonucleotide assay, while insertion–deletion polymorphisms were monitored by direct resequencing. As shown in Table 4, all of the six analyzed loci revealed a significant assortment of genetic polymorphism by habitat. These results extend the initial observations, which suggest that genes contained within these regions could function in adaptation to saline habitats.

Table 4. Expanded genotyping of polymorphisms that assort with habitat. Sites were identified by Sanger sequencing of loci containing the probe sets that assort with habitat in 45 TN genotypes from saline (TN1 and TN8) and non-saline (TN9 and TN10) populations, from which 12 (in bold) were used in the microarray experiment.

Amp	p-value	**TN 1.4**	**TN 1.16**	**TN 1.17**	TN 1.3	TN 1.5	TN 1.8	TN 1.9	TN 1.11	TN 1.12	TN 1.14	TN 1.19	TN 1.20	TN 1.21	**TN 8.4**	**TN 8.23**	**TN 8.27**	TN 8.3	TN 8.15	TN 8.20	TN 8.21	TN 8.22	TN 8.24	TN 8.29	**TN 9.12**	**TN 9.19**	**TN 9.21**	TN 9.1	TN 9.3	TN 9.4	TN 9.5	TN 9.17	TN 9.18	TN 9.20	TN 9.22	**TN 10.3**	**TN 10.16**	**TN 10.17**	TN 10.1	TN 10.2	TN 10.4	TN 10.8	TN 10.9	TN 10.15	TN 10.19	TN 10.23
1	1.7E-2	S	S	S	S	NS	NS	S	S	S	NS	NS	NS	S	S	S	S	NS	S	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS
2.1	3.7E-5	S	S	S	NS	NS	NS	S	S	S	S	S	S	S	S	S	S	NS	S	NS	S	NS	S	NS	NS	NS	S	NS	NS	S	NS	S	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS
2.2	1.1E-7	S	S	S	NS	S	S	S	S	S	S	S	S	S	S	S	S	NS	S	NS	S	S	NS	NS	NS	NS	NS	NS	S	S	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS
3	5.4E-7	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	-	S	-	S	S	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	S	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS
5.1.1	2.2E-3	S	S	S	S	NS	NS	S	S	-	-	-	S	S	S	S	S	-	-	S	-	S	-	S	NS	NS	NS	S	-	S	NS	S	S	NS	NS	S	NS	NS	NS	NS	NS	S	-	NS	NS	NS	NS
5.1.2	5.2E-7	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	NS	NS	NS	S	NS	S	NS	S	S	NS	S	S	NS	S	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS

Standard population genetic statistics were computed on the Sanger sequences obtained for six differentiated loci, as well as for a set of seven ‘control’ loci corresponding to COS markers. As shown in Table 5, some candidate loci have higher levels of polymorphism than control loci (Pi for candidate loci 0.00082–0.028, control loci, 0.0014–0.0094; Theta per nucleotide candidate loci 0.00075–

0.038, control loci 0.001698–0.00857). F_{ST} tends to be higher for candidate loci than control loci (F_{ST} for candidate loci 0.32–0.71, control loci −0.03 to 0.54). However, Bonferroni-corrected t-tests do not support differences between candidate and control loci for any parameter, other than a marginally significant difference in F_{ST} (Pi, P = 0.188; ThetaNuc, P = 0.233; Tajima's D, P = 0.515; F_{ST}, P = 0.0106; Nm, P = 0.310).

Conclusions

Tunisian populations of *Medicago truncatula* harbor substantial amounts of polymorphism with relatively low levels of LD. LD is half of its maximal value at approximately 10 kb and at background levels by approximately 500 kb on average. For comparison, in global samples of the model plant *A. thaliana*, LD extends on average by approximately 10 kb (Kim et al. 2007). In *Hordeum vulgare* (barley), cultivated Hordeum germplasm has LD extending across a 212 kb region, while in wild *Hordeum spontaneum* LD does not extend past genic regions, i.e. 28 kb (Caldwell et al. 2006). Similarly, cultivated species of rice, including *Oryza indica* and tropical and temperate *Oryza japonica*, have average LDs of 75, 150, and 500 kb respectively, while their wild relative *Oryza rufipogon* has LD < 40 kb (Mather et al. 2007). More limited data are available for legume species, but in the case of soybean (*Glycine max*) compared to its wild progenitor *Glycine soja*, LD extends up to 500 kb in cultivated accessions while it decays within 100 kb in non-domesticated genotypes, though different genomic regions show slightly different patterns (Hyten et al. 2007). These *M. truncatula* samples span four subpopulations, which could explain why LD takes approximately 500 kb to completely decay.

Table 5. Population genetic parameters for sequences covering differentiated single feature polymorphisms (SFPs) and control genomic regions (conserved orthologous sequence (COS) markers). N_i, number of individuals with Sanger sequence data; N_pops, number of populations, Sites, total length of sequenced locus; NetSites, sites with no missing data; S, number of segregating sites; Pi, average pairwise nucleotide polymorphism; ThetaNuc, per nucleotide estimate of Watterson's theta; Tajima D, measure of allele frequency skew from neutral; FST, measure of population differentiation; Nm, estimate of gene flow between populations (m) scaled by effective population size (N), n.d., not determined.

Locus	N_i	N_pops	NetSites	S	Pi	ThetaNuc	TajimaD	Fst	Nm
Control loci									
COS1_1433	12	4	480	10	0.003756	0.006898	-1.874	-0.02061	-12.38
COS10_SHMT	13	4	564	15	0.00941	0.00857	0.4096	0.14493	1.48
COS11_SUSY	12	4	487	4	0.001368	0.002719	-1.7469	*n.d.*	*n.d.*
COS2_AAT	12	4	807	7	0.002008	0.002872	-1.1763	-0.03448	-7.5
COS3_AGT	12	4	724	5	0.002218	0.002286	-0.1105	0.07954	2.89
COS6_CALTL	12	4	390	2	0.00202	0.001698	0.5542	0.18182	1.13
COS7_CNGC4	11	4	303	2	0.00228	0.002253	0.0361	0.53846	0.21
Differentiated loci									
Mtr.20569.S1	21	3	741	2	0.000822	0.00075	0.2222	0.71247	0.1
Mtr.20573.S1	19	3	477	31	0.023269	0.018594	0.9991	0.3169	0.54
Mtr.21891.S1	24	4	625	4	0.001472	0.001713	-0.3852	0.48108	0.27
Mtr.42442.S1	24	4	582	6	0.001556	0.00276	-1.3194	0.43333	0.33
Mtr.48956.S1	23	4	500	8	0.004695	0.004335	0.2698	0.3704	0.42
Mtr.8358.S1	24	4	383	50	0.028096	0.037756	-0.9958	0.50345	0.25

Extending the microarray genotyping polymorphism rates to the whole genome, predicts on average 2.6 polymorphic sites kb^{-1}. High levels of linkage disequilibrium in these populations extend 10–100 kb on average, so around 26–260 segregating sites are expected per LD block. With a target of 10 markers per LD block and an estimated genome size of 500 Mb, a dense marker set in these *M. truncatula* populations would require half a million markers. This is readily achievable with current microarray technology or with next-generation sequencing. In addition, the observation that only 3% of the polymorphic probes assort with population suggests that gene flow among these populations is relatively high. Our data lead us to predict that genome-wide association mapping in *M. truncatula* is likely to be successful in comprehensively localizing the genetic basis of adaptation. Indeed, the present coarse survey has already yielded several plausible candidates for local adaptation to soil salinity.

Acknowledgements

Matilde Cordeiro (MAC) joined this project after the microarray data was generated, when she validated the array, defined the LD blocks assorting by habitat and the candidate genes involved. As specified in the published manuscript: MLF, SVN, MB, TH, and MEA conceived the study. TH, MB, and MEA provided material. MLF performed the microarray experiment and analyzed the data with input from SVN. MAC validated the array data and gathered additional sequence data with input from RVP and DRC. MLF, MAC, DRC, and SVN wrote the paper with input from all authors.

MAC acknowledges Fundação para a Ciência e Tecnologia SFRH/BD/39905/2007. This work was supported by National Science Foundation Plant Genome Research Program No. 0820846 and National Science Foundation Office of International Science and Engineering No. 0751073 awards to SVN. The authors would like to thank the following people: Carine Ameline-Torregrosa (ENSAT) helped with DNA extractions; Thomas L. Turner (UCSB) provided the experimental protocol and provided advice; Matt Rolston (UCD) provided the hybridization recipe and protocol; Betty Schaub (CHLA) performed the hybridizations; Yuri Bendana (USC) implemented a local instance of *M. truncatula* GBrowse.

References

Aranzana MJ, Kim S, Zhao K, Bakker E, Horton M, Jakob K, Lister C, Molitor J, Shindo C, Tang C, et al (2005) Genome-wide association mapping in Arabidopsis identifies previously known flowering time and pathogen resistance genes. *PLoS genetics,* **1**: e60.

Badri M, Ilahi H, Huguet T, Aouani ME (2007) Quantitative and molecular genetic variation in sympatric populations of *Medicago laciniata* and *M. truncatula* (Fabaceae): relationships with eco-geographical factors. *Genetical research,* **89**, 107-122.

Bonnin I, Huguet T, Gherardi M, Prosperi JM, Olivieri I (1996) High level of polymorphism and spatial structure in a selfing plant species, Medicago truncatula (Leguminosae), shown using RAPD markers. *American journal of botany,* **83**, 843-855.

Bonnin I, Ronfort J, Wozniak F, Olivieri I (2001) Spatial effects and rare outcrossing events in *Medicago truncatula* (Fabaceae). *Molecular ecology,* **10**, 1371-1383.

Borevitz JO, Liang D, Plouffe D, Chang HS, Zhu T, Weigel D, Berry CC, Winzeler E, Chory J (2003) Large-scale identification of single-feature polymorphisms in complex genomes. *Genome research,* **13**, 513-523.

Borevitz JO, Hazen SP, Michael TP, Morris GP, Baxter IR, Hu TT, Chen H, Werner JD, Nordborg M, Salt DE, et al (2007) Genome-wide patterns of single-feature polymorphism in *Arabidopsis thaliana*. *Proceedings of the National Academy of Sciences of the U S A,* **104**, 12057-12062.

Burnham KP, Anderson DR (2002) Model selection and multimodel inference: a practical information-theoretic approach. New York: Springer.

Caldwell KS, Russell J, Langridge P, Powell W (2006) Extreme population-dependent linkage disequilibrium detected in an inbreeding plant species, *Hordeum vulgare*. *Genetics,* **172**, 557-567.

Chaulet E, Prosperi JM (1994) Genetic diversity of a collection of *Medicago trunculata* Gaernt from Algeria. EUCARPIA, Evaluation and exploitation of genetic resources pre-breeding, Clermont Ferrand (France), 15–18 Mar 1994.

Chen C, Fan C, Gao M, Zhu H (2009) Antiquity and function of CASTOR and POLLUX, the twin ion channel-encoding genes key to the evolution of root symbioses in plants. *Plant physiology,* **149**, 306-317.

Choi HK, Luckow MA, Doyle J, Cook DR (2006) Development of nuclear gene-derived molecular markers linked to legume genetic maps. *Molecular genetics and genomics,* **276**, 56-70.

De Mita S, Santoni S, Hochu I, Ronfort J, Bataillon T (2006) Molecular evolution and positive selection of the symbiotic gene NORK in *Medicago truncatula*. *Journal of molecular evolution,* **62**, 234-244.

Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M, Dudoit S, Ellis B, Gautier L, Ge Y, Gentry J (2004) Bioconductor: open software development for computational biology and bioinformatics. *Genome biology,* **5**: R80.

Gomez SK, Javot H, Deewatthanawong P, Torres-Jerez I, Tang Y, Blancaflor EB, Udvardi MK, Harrison MJ (2009) *Medicago truncatula* and *Glomus intraradices* gene expression in cortical cells harboring arbuscules in the arbuscular mycorrhizal symbiosis. *BMC plant biology,* **9**: 10.

Heath KD, Tiffin P (2007) Context dependence in the coevolution of plant and rhizobial mutualists. *Proceedings Biological sciences of the Royal Society,* **274**, 1905-1912.

Heath KD, Tiffin P (2009) Stabilizing mechanisms in a legume--rhizobium mutualism. *Evolution; international journal of organic evolution,* **63**, 652-662.

Hyten DL, Choi IY, Song Q, Shoemaker RC, Nelson RL, Costa JM, Specht JE, Cregan PB (2007) Highly variable patterns of linkage disequilibrium in multiple

soybean populations. *Genetics,* **175**, 1937-1944.

Johnson NC, Wilson GW, Bowker MA, Wilson JA, Miller RM (2010) Resource limitation is a driver of local adaptation in mycorrhizal symbioses. *Proceedings of the National Academy of Sciences,* **107**, 2093-2098.

Kim S, Plagnol V, Hu TT, Toomajian C, Clark RM, Ossowski S, Ecker RJ, Weigel D, Nordborg M (2007) Recombination and linkage disequilibrium in *Arabidopsis thaliana*. *Nature Genetics*, **39**, 1151–1155.

Langmead B, Trapnell C, Pop M, Salzberg SL (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome biology,* **10**: R25.

Lazrek F, Roussel V, Ronfort J, Cardinet G, Chardon F, Aouani ME, Huguet T (2009) The use of neutral and non-neutral SSRs to analyse the genetic structure of a Tunisian collection of *Medicago truncatula* lines and to reveal associations with eco-environmental variables. *Genetica,* **135**, 391-402.

Lesins K, Lesins I (1979) Genus Medicago (Leguminosae), A Taxogenetic Study. The Hague, The Netherlands: W. Junk.

Librado P, Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics (Oxford, England),* **25**, 1451-1452.

Mather KA, Caicedo AL, Polato NR, Olsen KM, McCouch S, Purugganan MD (2007) The extent of linkage disequilibrium in rice (*Oryza sativa* L.). *Genetics,* **177**, 2223-2232.

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics,* **155**, 945-959.

Pritchard JK, Wen X, Falush D (2010) Documentation for Structure Software: Version 2.3. Chicago, IL.

R Team (2009) R: A Language and Environment for Statistical Computing. Vienna: R Foundation for Statistical Computing.

Rangin C, Brunel B, Cleyet-Marel JC, Perrineau MM, Béna G (2008) Effects of *Medicago truncatula* genetic diversity, rhizobial competition, and strain effectiveness on the diversity of a natural sinorhizobium species community. *Applied and environmental microbiology,* **74**, 5653-5661.

Ronfort J, Bataillon T, Santoni S, Delalande M, David JL, Prosperi J-M (2006) Microsatellite diversity and broad scale geographic structure in a model legume: building a set of nested core collection for studying naturally occurring variation in *Medicago truncatula*. *BMC plant biology,* **6**: 1.

Rosenberg NA(2004) distruct: a program for the graphical display of population structure. *Molecular Ecology Notes*, **4**, 137–138.

Rosenberg, N.A. and Nordborg, M. (2006) A general population-genetic model for the production by population structure of spurious genotype-phenotype associations in discrete, admixed or spatially distributed populations. Genetics, 173, 1665–1678.

Siol M, Bonnin I, Olivieri I, Prosperi JM, Ronfort J (2007) Effective population size associated with self-fertilization: lessons from temporal changes in allele frequencies in the selfing annual *Medicago truncatula*. *Journal of evolutionary biology,* **20**, 2349-2360.

Siol M, Prosperi JM, Bonnin I, Ronfort J (2008) How multilocus genotypic pattern helps to understand the history of selfing populations: a case study in *Medicago truncatula*. *Heredity,* **100**, 517-525.

Slatkin M (2008) Linkage disequilibrium--understanding the evolutionary past and mapping the medical future. *Nature reviews. Genetics,* **9**, 477-485.

Small E, Jomphe M (1989) A synopsis of the genus Medicago (Leguminosae). *Canadian Journal of Botany,* **67**, 3260-3294.

Tester M, Davenport R (2003) Na+ tolerance and Na+ transport in higher plants. *Annals of botany,* **91**, 503-527.

Turner TL, Hahn MW, Nuzhdin SV (2005) Genomic islands of speciation in *Anopheles gambiae*. *PLoS biology,* **3**: e285.

Turner TL, Levine MT, Eckert ML, Begun DJ (2008a) Genomic analysis of adaptive differentiation in *Drosophila melanogaster*. *Genetics,* **179**, 455-473.

Turner TL, von Wettberg EJ, Nuzhdin SV (2008) Genomic analysis of differentiation between soil types reveals candidate genes for local adaptation in *Arabidopsis lyrata*. *PloS one,* **3**: e3183.

Turner TL, Bourne EC, Von Wettberg EJ, Hu TT, Nuzhdin SV (2010) Population resequencing reveals local adaptation of *Arabidopsis lyrata* to serpentine soils. *Nature genetics,* **42**, 260-263.

Wilczek AM, Roe JL, Knapp MC, Cooper MD, Lopez-Gallego C, Martin LJ, Muir CD, Sim S, Walker A, Anderson J, et al (2009) Effects of genetic perturbation on seasonal life history plasticity. *Science (New York, N.Y.),* **323**, 930-934.

Winzeler EA, Richards DR, Conway AR, Goldstein AL, Kalman S, McCullough MJ, McCusker JH, Stevens DA, Wodicka L, Lockhart DJ, et al (1998) Direct allelic variation scanning of the yeast genome. *Science,* **281**, 1194-1197.

Young ND, Udvardi M (2009) Translating *Medicago truncatula* genomics to crop legumes. *Current opinion in plant biology,* **12**, 193-201.

Zhao K, Aranzana MJ, Kim S, Lister C, Shindo C, Tang C, Toomajian C, Zheng H, Dean C, Marjoram P, et al (2007) An Arabidopsis example of association mapping in structured samples. *PLoS genetics,* **3**: e4.

Zhu H, Cannon SB, Young ND, Cook DR (2002) Phylogeny and genomic organization of the TIR and non-tIR NBS-LRR resistance gene family in *Medicago truncatula*. *Molecular plant-microbe interactions : MPMI,* **15**, 529-539.

Supplemental material

Table S1. All genes, probes, and markers contained in candidate regions, i.e., regions defined by markers that assort with saline source habitat.

Chr	Size (Kbp)	pos (in)	pos (end)	Marker probes (bold dif)	Gene annotation	Gene	Amp
1	64.11	17066007	17067620		Patatin	Medtr1g086350	
		17071626	17072774		hypothetical protein	Medtr1g086360	
		17071654		Mtr.22481.1.S1_s_at:440:1103			
		17071725		Mtr.22481.1.S1_s_at:800:291			
		17071745		Mtr.22481.1.S1_s_at:159:1057			
		17071778		Mtr.22481.1.S1_s_at:1039:523			
		17071793		Mtr.22481.1.S1_s_at:210:449			
		17071809		Mtr.22481.1.S1_s_at:567:827			
		17071920		Mtr.22481.1.S1_s_at:878:1063			
		17071969		Mtr.22481.1.S1_s_at:39:511			
		17075797		Mtr.48364.1.S1_at:141:489			
		17075808		Mtr.48364.1.S1_at:270:521			
		17077054	17078848		Nucleic acid-binding	Medtr1g086370	
		17082922	17083143		hypothetical protein	Medtr1g086380	
		17083072		Mtr.21889.1.S1_at:312:165			
		17083112		Mtr.21889.1.S1_at:431:357			
		17084941	17085222		hypothetical protein	Medtr1g086390	
		17084952		Mtr.21890.1.S1_at:1068:1075			
		17084985		Mtr.21890.1.S1_at:305:947			
		17085157		Mtr.21890.1.S1_at:263:801			
		17085435	17086458		hypothetical protein	Medtr1g086400	
		17086513	17089383		Leucine-rich repeat Leucine-rich repeat	Medtr1g086410	-
		17088280	17088992				Chr1
		17088611		**Mtr.21891.1.S1_at:635:557**			"
		17090244	17092618		NAD-dependent epimerase/dehydratase	Medtr1g086420	
		17090272		Mtr.10441.1.S1_at:188:737			
		17090331		Mtr.10441.1.S1_at:237:931			
		17090334		Mtr.10441.1.S1_at:201:783			
		17090336		Mtr.10441.1.S1_at:214:807			
		17090337		Mtr.10441.1.S1_at:189:133			
		17090338		Mtr.10441.1.S1_at:492:401			
		17090339		Mtr.10441.1.S1_at:22:479			
		17094444	17100084		Dynamin Thiamine monophosphate synthase	Medtr1g086430	
		17099532		Mtr.21892.1.S1_at:621:179			
		17099918		Mtr.2558.1.S1_at:783:695			

Chr	Size (Kbp)	pos (in)	pos (end)	Marker probes (bold dif)	Gene annotation	Gene	Amp
		17099933		Mtr.2558.1.S1_at:1011:465			
		17102829	17111254		Calcium-binding EF-hand Ras small GTPase	Medtr1g086440	
		17114595	17116828		Pentatricopeptide repeat	Medtr1g086450	
		17120714	17121765		hypothetical protein	Medtr1g086470	
		17121651		**Mtr.48008.1.S1_at:674:567**			
		17122681	17130116		Helix-loop-helix DNA-binding	Medtr1g086480	
		17122862		Mtr.22479.1.S1_at:874:853			
2.1	17.33	18976699	18977286		hypothetical protein	Medtr2g083010	
		18982274	18982977		hypothetical protein	Medtr2g083020	
		18984721	18985249		hypothetical protein	Medtr2g083030	
		18986587	18987268		hypothetical protein	Medtr2g083040	
		18991291		**Mtr.23081.1.S1_at:1046:237**			
		18991293		**Mtr.23081.1.S1_at:421:337**			
		18994025	18994201		hypothetical protein	Medtr2g083050	
2.2	43.11	21927619	21927843		hypothetical protein	Medtr2g095720	-
		21934468	21934962				Chr2_1
		21934861		**Mtr.20573.1.S1_at:839:803**			"
		21934959		Mtr.20573.1.S1_at:839:727			"
		21938259	21938603		Calcium-binding EF-hand	Medtr2g095730	
		21941919	21942310		hypothetical protein	Medtr2g095740	-
		21942438	21943217				Chr2_2
		21942971		Mtr.20569.1.S1_at:482:439			"
		21942999		**Mtr.20569.1.S1_at:603:565**			"
		21948857	21964035		hypothetical protein	Medtr2g095750	
		21967045	21967206		hypothetical protein	Medtr2g095760	
		21970380	21970724		Calcium-binding EF-hand	Medtr2g095770	
		21970421		Mtr.51495.1.S1_at:156:885			
		21970423		Mtr.51495.1.S1_at:159:109			
3.1	2.162	21844852	21847013		hypothetical protein	Medtr3g092650	
		21845997	21845997	Mtr.38228.1.S1_at:513:585			
		21846396	21846745				Chr3_1
		21846491	21846491	Mtr.38228.1.S1_at:211:365			"
		21846705	21846705	**Mtr.8358.1.S1_at:427:1115**			"
3.2	49.96	22351489	22357993		hypothetical protein	Medtr3g093800	
		22359951	22361112		hypothetical protein	Medtr3g093810	
		22362293	22362839		hypothetical protein	Medtr3g093820	
		22363888	22368308		C2	Medtr3g093830	

Chr	Size (Kbp)	pos (in)	pos (end)	Marker probes (bold dif)	Gene annotation	Gene	Amp
		22370756	22375722		C2 EF-Hand type	Medtr3g093840	
		22374829		**Mtr.10504.1.S1_at:798:1029**			
		22380112	22383536		C2 Phosphoinositide-specific phospholipase C (PLC)	Medtr3g093850	
		22384032	22385360		hypothetical protein	Medtr3g093860	
		22385657	22386300		hypothetical protein	Medtr3g093870	
		22389259	22391165		Virulence factor	Medtr3g093880	
		22391167	22391684		Virulence factor	Medtr3g093890	
		22392482	22393455		glycine-rich protein	Medtr3g093900	
		22394972	22401448		RNA polymerase	Medtr3g093910	
		22395035		Mtr.11064.1.S1_at:378:447			
4.1	28.86	5884997	5886099		UDP-glucuronosyl/UDP-glucosyltransferase	Medtr4g025030	
		5891017	5891430		hypothetical protein	Medtr4g025040	
		5893033	5897572		hypothetical protein	Medtr4g025050	
		5893033	5895530		hypothetical protein	Medtr4g025060	
		5893321		Mtr.32058.1.S1_at:692:991			
		5893325		Mtr.32058.1.S1_at:221:1001			
		5894772		Mtr.33441.1.S1_s_at:795:1017			
		5894826		Mtr.33441.1.S1_s_at:443:255			
		5895060		Mtr.33441.1.S1_at:332:721			
		5895063		Mtr.33441.1.S1_at:753:765			
		5895092		**Mtr.33441.1.S1_at:147:233**			
		5895697		Mtr.9037.1.S1_at:519:955			
		5896843		Mtr.9037.1.S1_at:771:815			
		5899430	5902006		hypothetical protein	Medtr4g025060	
		5904643	5904967		hypothetical protein	Medtr4g025070	
		5908624	5911978		hypothetical protein	Medtr4g025080	
		5913054	5913284		hypothetical protein	Medtr4g025090	
		5913545	5913860		Integrase	Medtr4g025100	
4.2	52.18	28624487	28628045		hypothetical protein	Medtr4g121710	
		28629177	28632708		Oligopeptide transporter OPT superfamily	Medtr4g121720	
		28649960	28654019		Chaperone DnaK	Medtr4g121730	
		28654808	28655620		hypothetical protein	Medtr4g121740	
		28657385	28660799		Bromodomain	Medtr4g121750	
		28658656		**Mtr.37707.1.S1_at:1145:225**			
		28664821	28665060		hypothetical protein	Medtr4g121760	

Chr	Size (Kbp)	pos (in)	pos (end)	Marker probes (bold dif)	Gene annotation	Gene	Amp
		28668603	28670965		Bromodomain	Medtr4g121770	
		28668603	28671748		Bromodomain	Medtr4g121780	
		28673342	28675695		hypothetical protein	Medtr4g121790	
		28676091	28676670		hypothetical protein	Medtr4g121800	
5.1.1	65.62	8617978	8620419		Armadillo	Medtr5g022170	
		8625089	8626192		hypothetical protein	Medtr5g022180	
		8625103		**Mtr.17919.1.S1_at:414:683**			
		8628391	8629448		hypothetical protein	Medtr5g022190	
		8629387		Mtr.41516.1.S1_at:95:123			
		8629743	8631863		Peptidase C48	Medtr5g022200	
		8630168		Mtr.48963.1.S1_at:67:1119			
		8630171		Mtr.48963.1.S1_at:312:935			
		8632118		Mtr.41516.1.S1_at:1023:229			
		8634233	8635309		hypothetical protein	Medtr5g022210	
		8635208		Mtr.17916.1.S1_at:421:959			
		8635311	8637352		FAR1	Medtr5g022220	
		8647104	8647875		HAT dimerisation	Medtr5g022230	
		8651184	8654622		Oxidoreductase FAD/NAD(P)-binding	Medtr5g022240	
		8659225	8661273		Thaumatin	Medtr5g022250	
		8665015	8665647		hypothetical protein	Medtr5g022260	
		8668657	8668967		Short-chain dehydrogenase/reductase SDR	Medtr5g022270	
		8671722	8673449		Thaumatin	Medtr5g022280	
		8675926	8676380				Chr5_1
		8676291		**Mtr.48956.1.S1_at:865:623**			"
		8676293		**Mtr.48956.1.S1_at:763:453**			"
		8676295		**Mtr.48956.1.S1_at:1070:921**			"
		8676297		**Mtr.48956.1.S1_at:464:815**			"
		8676299		**Mtr.48956.1.S1_at:586:901**			"
		8676301		**Mtr.48956.1.S1_at:846:457**			"
		8676311	8679336		hypothetical protein	Medtr5g022290	-
		8679578	8680575		hypothetical protein	Medtr5g022300	
		8680577	8683602		Nodulin-like	Medtr5g022310	
		8680913		Mtr.48955.1.S1_at:284:709			
Between 5.1.1&5.1.2	93.45	8684964	8686606		Rhodanese-like	8684964	
		8687859	8691349		Mlo-related protein	8687859	

Chr	Size (Kbp)	pos (in)	pos (end)	Marker probes (bold dif)	Gene annotation	Gene	Amp
		8690312		Mtr.17908.1.S1_s_at:703:487		8690312	
		8691138		Mtr.17908.1.S1_s_at:183:61		8691138	
		8694568		Mtr.17907.1.S1_at:944:375		8694568	
		8694582		Mtr.17907.1.S1_at:175:477		8694582	
		8699479	8700621		Cyclin-like F-box; F-box protein interaction domain	8699479	
		8704849	8710879		von Willebrand factor	8704849	
		8710809		Mtr.37014.1.S1_at:136:875		8710809	
		8712687	8718370		Senescence-associated	8712687	
		8720896	8721179		hypothetical protein	8720896	
		8723815	8724521		hypothetical protein	8723815	
		8725953	8732768		Response regulator receiver Heavy metal sensor kinase	8725953	
		8738079	8738336		NB-ARC	8738079	
		8738399	8741638		Leucine-rich repeat	8738399	
		8743465	8746242		Haem peroxidase	8743465	
		8748205	8748756		hypothetical protein	8748205	
		8750392	8755532		Protein of unknown function DUF630	8750392	
		8758684	8759397		hypothetical protein	8758684	
		8760270	8761679		Peptidase C1A	8760270	
		8761529		Mtr.8512.1.S1_at:663:217		8761529	
		8761543		Mtr.8512.1.S1_at:1155:473		8761543	
		8762509	8764112		hypothetical protein	8762509	
		8762849		Mtr.31644.1.S1_at:210:277		8762849	
		8765278	8767416		3-dehydroquinate synthase	8769375	
		8765278	8767497			8769567	
		8769375		Mtr.37570.1.S1_at:1093:837		8769738	
		8769567		Mtr.37570.1.S1_at:1036:197		8775550	
		8769738	8772824		D-galactoside/L-rhamnose binding SUEL lectin	8776982	
		8775550	8778414		Glycoside hydrolase Ras GTPase		
		8776982		Mtr.43489.1.S1_at:981:1159			
5.1.2	65.61	8779956	8791459		Vacuolar protein sorting-associated protein 35	Medtr5g022510	
		8792564	8793130		Heavy meta transport/detoxification protein	Medtr5g022520	
		8793254	8793424		hypothetical protein	Medtr5g022530	

Chr	Size (Kbp)	pos (in)	pos (end)	Marker probes (bold dif)	Gene annotation	Gene	Amp
		8799416	8800433		Lipolytic enzyme	Medtr5g022540	
		8800774	8802715		Lipolytic enzyme	Medtr5g022550	
		8807604	8809496		Lipolytic enzyme	Medtr5g022560	
		8825347	8829995		Uncharacterized Cys-rich domain	Medtr5g022570	
		8827403	8828089				Chr5_2
		8827748		**Mtr.42442.1.S1_at:633:1151**			
		8830696	8831587		hypothetical protein	Medtr5g022580	
		8831951	8834564		Splicing factor motif WD40-like	Medtr5g022590	
		8835936	8840655		Diaminopimelate decarboxylase	Medtr5g022600	
		8843627	8845569		SKP1 component	Medtr5g022610	
5.2	24.66	9272181	9273029		Ribulose bisphosphate carboxylase	Medtr5g023550	
		9273889	9274788		hypothetical protein	Medtr5g023560	
		9275508	9277880		Zinc finger	Medtr5g023570	
		9278305	9278986		hypothetical protein	Medtr5g023580	
		9279288	9280720		hypothetical protein	Medtr5g023590	
		9286141	9291825		AMP-dependent synthetase and ligase	Medtr5g023600	
		9291530		**Mtr.43580.1.S1_at:1008:607**			
		9291545		**Mtr.43580.1.S1_at:1012:429**			
		9291589		**Mtr.43580.1.S1_at:442:351**			
		9291604		**Mtr.43580.1.S1_at:394:669**			
		9292742	9296836		Isocitrate dehydrogenase NAD-dependen	Medtr5g023610	
		9296748		Mtr.39958.1.S1_at:92:241			
5.3	6.07	11059938	11061645		Nuclear protein SET Zinc finger	Medtr5g027240	
		11061076		Mtr.11801.1.S1_at:609:1117			
		11061503		**Mtr.11801.1.S1_at:95:849**			
		11063203	11066007		QLQ	Medtr5g027250	
5.4	4.282	30719205	30723486		DNA-binding WRKY Exo70 exocyst complex subunit	Medtr5g081770	
		30720751		Mtr.18994.1.S1_at:4:159			
		30720995		**Mtr.18994.1.S1_at:596:269**			
		30721009		**Mtr.18994.1.S1_at:283:687**			
		30721118		Mtr.18994.1.S1_at:873:605			
		30722713		Mtr.18994.1.S1_at:1019:233			
		30722776		Mtr.18994.1.S1_at:781:827			

Chr	Size (Kbp)	pos (in)	pos (end)	Marker probes (bold dif)	Gene annotation	Gene	Amp
6.1	51.72	10237963	10238121		hypothetical protein	Medtr6g047000	
		10238488	10238685		hypothetical protein	Medtr6g047010	
		10239934	10240203		hypothetical protein	Medtr6g047020	
		10240936	10241205		hypothetical protein	Medtr6g047030	
		10243456	10253009		Leucine-rich repeat	Medtr6g047040	
		10250977		**Mtr.2455.1.S1_at:223:475**			
		10250988		**Mtr.2455.1.S1_at:369:1055**			
		10251060		**Mtr.2455.1.S1_at:14:455**			
		10251122		**Mtr.2455.1.S1_at:188:447**			
		10251134		**Mtr.2455.1.S1_at:970:139**			
		10251275		**Mtr.2455.1.S1_at:230:469**			
		10251490		**Mtr.2455.1.S1_at:71:559**			
		10252294		Mtr.11624.1.S1_at:1019:215			
		10252408		**Mtr.11624.1.S1_at:590:975**			
		10252560		**Mtr.11624.1.S1_at:763:435**			
		10253746		Mtr.29268.1.S1_x_at:761:757			
		10254957	10257135		Leucine-rich repeat	Medtr6g047050	
		10255010		Mtr.32020.1.S1_at:134:1071			
		10255064		Mtr.32020.1.S1_at:736:791			
		10255158		Mtr.32020.1.S1_at:1158:175			
		10256157		Mtr.29574.1.S1_at:71:27			
		10258019	10259019		hypothetical protein	Medtr6g047060	
		10258071		**Mtr.6977.1.S1_at:901:537**			
		10258080		**Mtr.6977.1.S1_at:97:989**			
		10258082		**Mtr.6977.1.S1_at:690:1091**			
		10258085		**Mtr.6977.1.S1_at:235:655**			
		10258087		**Mtr.6977.1.S1_at:84:1057**			
		10258198		**Mtr.6977.1.S1_at:967:359**			
		10258203		**Mtr.6977.1.S1_at:310:235**			
		10258214		**Mtr.6977.1.S1_at:727:319**			
		10261256	10263099		NB-ARC	Medtr6g047070	
		10265169	10268641		hypothetical protein	Medtr6g047080	
		10276617	10289680		hypothetical protein	Medtr6g047090	
6.2	0.868	10662451	10662798		hypothetical protein	Medtr6g059310	
		10662477		Mtr.48980.1.S1_at:1121:1159			
		10662486		Mtr.48980.1.S1_at:1009:629			
		10662492		Mtr.48980.1.S1_at:734:651			
		10662498		Mtr.48980.1.S1_at:452:1125			
		10662519		**Mtr.48980.1.S1_at:668:433**			
		10662530		Mtr.48980.1.S1_at:1018:835			
		10662536		Mtr.48980.1.S1_at:36:979			

Chr	Size (Kbp)	pos (in)	pos (end)	Marker probes (bold dif)	Gene annotation	Gene	Amp
		10662544		Mtr.48980.1.S1_at:925:1133			
		10662568		Mtr.48980.1.S1_at:81:453			
		10662785		Mtr.48980.1.S1_at:480:829			
		10662916	10663318		Proteinase inhibitor I1	Medtr6g059320	
6.3	0.229	11332555	11332783		hypothetical protein	Medtr6g062310	
		11332594		Mtr.22401.1.S1_at:679:1159			
		11332599		Mtr.22401.1.S1_at:348:1075			
		11332613		**Mtr.22401.1.S1_at:135:569**			
		11332679		Mtr.22401.1.S1_at:592:5			
		11332684		Mtr.22401.1.S1_at:612:139			
6.4	3.401	14923810	14927210		hypothetical protein	Medtr6g076790	
		14924078		Mtr.585.1.S1_at:942:261			
		14924270		Mtr.30229.1.S1_at:1072:241			
		14924280		Mtr.30229.1.S1_at:1084:359			
		14924390		**Mtr.30229.1.S1_at:386:321**			
		14926596		Mtr.25758.1.S1_at:261:187			
		14926625		Mtr.25758.1.S1_at:721:493			
		14926632		Mtr.25758.1.S1_at:48:763			
		14926638		Mtr.25758.1.S1_at:794:543			
		14926673		Mtr.25758.1.S1_at:421:641			
		14926685		Mtr.25758.1.S1_at:1159:513			
		14926700		Mtr.25758.1.S1_at:247:1115			
7.1	3.913	22444707	22446711		Nucleoside phosphatase GDA1 CD39	Medtr7g100020	
		22446609		**Mtr.14073.1.S1_at:689:685**			
		22447583	22448619		non-LTR retrolelement reverse transcriptase-like protein	Medtr7g100030	
7.2	21.31	26351771	26352691		IMP dehydrogenase/GMP reductase	Medtr7g111810	
		26352788	26353093		IMP dehydrogenase/GMP reductase	Medtr7g111820	
		26353153	26354160		IMP dehydrogenase/GMP reductase	Medtr7g111830	
		26355228	26359225		Zinc finger	Medtr7g111840	
		26359962	26360123		hypothetical protein	Medtr7g111850	
		26361404	26362773		Late embryogenesis abundant (LEA) group 1	Medtr7g111860	

Chr	Size (Kbp)	pos (in)	pos (end)	Marker probes (bold dif)	Gene annotation	Gene	Amp
		26365622	26368519		Late embryogenesis abundant (LEA) group 1	Medtr7g111870	
		26371740	26373082		Late embryogenesis abundant (LEA) group 1	Medtr7g111880	
		26371931		**Mtr.12942.1.S1_at:303:1091**			
8	24.03	868185	869155		Polyneuridine-aldehyde esterase precursor	Medtr8g008360	
		872633	872957		hypothetical protein	Medtr8g008370	
		874909	878612		Peptidylprolyl isomerase	Medtr8g008380	
		878660	879211		hypothetical protein	Medtr8g008390	
		880472	883861		Cyclin-like F-box	Medtr8g008400	
		884818	884973		hypothetical protein	Medtr8g008410	
		886145	887776		Alpha/beta hydrolase fold	Medtr8g008420	
		886293		Mtr.35740.1.S1_at:1073:953			
		886308		Mtr.35740.1.S1_at:805:725			
		886321		Mtr.35740.1.S1_at:459:1021			
		886359		Mtr.35740.1.S1_at:931:373			
		886745	887776		Alpha/beta hydrolase fold Lipase	Medtr8g008420	
		886813		**Mtr.32725.1.S1_at:117:409**			
		886827		Mtr.32725.1.S1_at:605:663			
		886900		Mtr.32725.1.S1_at:402:789			
		886906		Mtr.32725.1.S1_at:678:723			
		886915		Mtr.32725.1.S1_at:902:1077			
		891304	892217		hypothetical protein	Medtr8g008430	

Table S2. All probes whose marker state was verified by Sanger re-sequencing.

Amplicon	Differentiated probes	Amplicon position	Other sequenced marker probes	Other sequenced probes
Chr1_1	Mtr.21891.1.S1_at:635:557	Chr1:17088280. .17088992		Mtr.21891.1.S1_at:41:741; 575:625; 425:269; 239:55; 324:795; 1037:113; 426:241; 743:289; 1053:267; 920:615
Chr2_1	Mtr.20573.1.S1_at:839:803	Chr2:21934468. .21934962	Mtr.20573.1.S1_at:839:727	Mtr.20573.1.S1_at:122:559; 440:589; 1098:411; 1031:369; 590:279; 114:185; 561:967; 926:433; 1074:421
Chr2_2	Mtr.20569.1.S1_at:603:565; 482:439	Chr2:21942338. .21942817		Mtr.20569.1.S1_at:309:773; 862:1013; 632:193; 819:641
Chr3_1	Mtr.8358.1.S1_at:427:1115	Chr3:21846396. .21846745	Mtr.38228.1.S1_at:211:365	Mtr.38228.1.S1_at:236:249; 79:181/Mtr.8358.1.S1_at:770:585; 1050:553; 239:657; 268:161; 951:489; 429:719; 17:747
Chr5_1.1	Mtr.48956.1.S1_at:763:453, 1070:921, 464:815, 586:901, 846:457, 865:623	Chr5:8675926.. 8676380		Mtr.48956.1.S1_at:682:7; 96:403; 407:659; 492:559; 524:979
Chr5_1.2	Mtr.42442.1.S1_at:633:1151	Chr5:8827403.. 8828089		Mtr.42442.1.S1_at:180:229; 603:1021

Chapter 3.

Population differentiation for germination and early seedling root growth between saline and non-saline origin Tunisian *Medicago truncatula*

***This work is published as**:*

Cordeiro MA, Moriuchi KS, Fotinos TD, Miller KE, Nuzhdin SV, von Wettberg EJ, Cook DR (2014) Population differentiation for germination and early seedling root growth traits under saline conditions in the annual legume Medicago truncatula (Fabaceae). *American journal of botany,* **101**, 488-498.

Matilde Ataide Cordeiro (MAC) was the major contributor for the elaboration of this chapter. MAC conceptualized, designed and performed experiments, analyzed and interpreted the data, and wrote the manuscript.

Chapter 3. Table of contents

Abstract

Seedling establishment and survival are highly sensitive to soil salinity and plants that evolved in saline environments are likely to express traits that increase fitness in those environments. Such traits are of ecological interest and they may have practical value for improving salt tolerance in cultivated species. Responses to soil salinity were examined and tested potential mechanisms of salt tolerance in *Medicago truncatula*, using genotypes that originated from natural populations occurring on saline and non-saline soils.

Germination and seedling responses were quantified and/or compared between saline and non-saline origin genotypes. Germination treatments included a range of NaCl concentrations in both offspring and parental environments. Seedling treatments included NaCl, ABA, and KCl. Saline origin genotypes displayed greater salinity tolerance for germination and seedling traits relative to non-saline origin genotypes. Moreover, we observed population specific differences for the effects of salinity on time to germination and for the impact of parental environment on germination rates. ABA and NaCl treatments had similar negative effects on root growth, although relative sensitivities differed, with saline population less sensitive to NaCl and more sensitive to ABA compared to their non-saline counterparts. These results demonstrate population differentiation for germination and seedling growth traits under saline conditions among populations derived from saline and non-saline environments. These observations are consistent with a syndrome of adaptations for salinity tolerance during early plant development, including traits that are common among saline environments and those that are idiosyncratic to local populations.

Introduction

Soil salinity is a prominent factor limiting crop yield throughout the world, especially in xeric environments (Flowers and Yeo 1995; FAO 2008). Salt stress

can shape genetic potential and indeed plants display a wide range of responses to soil salinity (Munns and Tester 2008), including adaptations that function at different stages of development (Maas 1986; Ashraf et al. 1987; Bayuelo-Jiménez et al. 2002). Germination and initial root growth are among the earliest traits exposed to natural selection (Weinig 2000; Wilczek et al. 2009) and their component traits are key factors determining fitness (e.g., Donohue 2002; Donohue et al. 2010; Bibee et al. 2011). In *Arabidopsis thaliana*, for example, time to germination explained 72% of the genetic variation in fitness (Donohue et al. 2005). Certain halophytic species possess adaptations for germination and early seedling development under saline conditions (Song et al. 2005), while in glycophytic species germination and seedling development may be among the most saline sensitive stages (Chang et al. 1961). Consistent with the importance of early development in adaptation, population differentiation and local adaptation for germination and early seedling traits are found across a range of plant taxa and environmental conditions (reviewed by Donohue et al. 2010; see also Galloway and Fenster 2000; Bischoff and Müller-Schärer 2010; Bibee et al. 2011). Understanding how early seedling traits are influenced by stressful environments can lead to the identification of responses to selection and, ultimately, to the discovery of the underlying genes and functional pathways.

In addition to the proximal impact of environmental stress on seed and seedlings, abiotic stress can impact the parental environment under which seeds develop. Parental exposure to adverse environmental factors can activate plastic adaptations that increase fitness in offspring prior to the offspring experiencing that environment (Roach and Wulff 1986; Galloway 2005), providing them with an early advantage over seedlings whose parents were either not exposed or were unable to transmit the response between generations. For example, seeds of *Iris hexagona* germinated earlier under saline conditions when parental plants were exposed to salt (van Zandt and Mopper 2004). Such transgenerational plasticity may involve adaptations at both the parental and seedling stages (Agrawal 2001; Sultan et al. 2009; Ballhorn et al. 2011), and studies have demonstrated population

differentiation and divergent selection for related traits (Hereford and Moriuchi 2005; Galloway and Etterson 2007), establishing their roles in local adaptation (Sultan 1996; Munir et al. 2001).

Once a seed germinates, the initial growth of the primary and lateral roots plays an important role in the offspring's ability to detect and potentially respond to the environment (Malamy 2005; Nibau et al. 2008). For example, seedling root development is crucial for seedling survival in highly saline environments, because root development determines water absorption capacity, which in turn buffers against saline-induced osmotic stress (Galvan-Ampudia and Testerink 2011). Development itself responds to environmental cues. Expansion at the root tips and root growth rate are reduced by salt, while lateral roots formation is stimulated under conditions of mild salinity and inhibited under high salt (Schenk and Jackson 2002; Osmont et al. 2007; Zhao et al. 2010; Galvan-Ampudia and Testerink 2011). In addition to facilitating water uptake per se, redirection of root growth can allow plants to explore microenvironments that may be more favorable for plant growth (Nibau et al. 2008).

Salinity impacts plant functions through both osmotic stress and ion toxicity. The most immediate manifestation of salinity is osmotic stress. This reduces the imbibition of water into the embryo resulting in the delay or prevention of germination (Wahid et al. 1999; Farissi et al. 2011; Khalil et al. 2011). Later, decreased water potential between the soil and root reduces water uptake and incites water loss (Boursiac et al. 2005; Hauser and Horie 2010). Osmotic stress can also result from high tissue concentrations of salt, especially in leaves. Ion toxicity results from excessive cellular Na^+, which can inhibit K^+ uptake and interfere with K^+ dependent metabolic and physiological functions (Rains and Epstein 1965; Ullah et al. 1993; Hauser and Horie 2010). For example, during germination, ion toxicity can disrupt the metabolism of carbohydrates resulting in delayed seedling development (Wahid et al. 1999; Witzel et al. 2010; Farissi et al. 2011). However ion balance, namely Na^+:K^+, may be more important to salinity tolerance than is the absolute concentration of Na^+ (Maathuis and Amtmann 1999;

Chartzoulakis et al. 2002; Meloni et al. 2008; Kronzucker and Britto 2011; Hauser and Horie 2010). Salinity stress and water deficit share the osmotic stress component. Indeed, the hormone abscisic acid (ABA) integrates root growth and development in response to both drought and high saline environments (e.g., Zhu 2002; Davies et al. 2005; Vinocur and Altman 2005). In addition to ABA's influence on root growth, exogenous application of ABA onto seeds has been shown to increase salinity tolerance in *Oryza sativa* (Gurmani et al. 2011), while many of the root architectural responses to NaCl can also be stimulated by ABA (DeSmet et al. 2003; Ariel et al. 2010; Zhao et al. 2010; Galvan-Ampudia and Testerink 2011).

Seeds of halophytes can avoid osmotic stress by remaining dormant when salinity is high and germinating during periods of low salinity (e.g., Song et al. 2005; Guo et al. 2012; but see Katembe et al. 1998). Compartmentalizing ions that accumulate from excess soil salinity during seedling development may be particularly important, with some halophytic species even requiring Na^+ uptake into the embryo for seed viability and germination (Li et al. 2011; Galvan-Ampudia and Testerink 2011). By contrast, seeds and seedlings of glycophytes tend to be negatively impacted by both osmotic stress and ion toxicity, and they typically display reduced and delayed germination as well as reduced seedling growth when exposed to salinity (Witzel et al. 2010; Farissi et al. 2011; Khalil et al. 2011; Guo et al. 2012). Glycophytes whose habitats include soils with elevated salinity can exhibit salinity tolerance during seed germination, for example through the production of osmolytes to maintain carbohydrate metabolism (Witzel et al. 2010; Chérifi et al. 2011; Farissi et al. 2011) or via mechanisms that maintain Na^+:K^+ ratios (Kent and Läuchi 1985). In *Medicago sativa*, for example, tolerant genotypes had higher and earlier germination, while at the physiological level seedlings accumulated more Na^+ ions and exhibited enhanced production of soluble sugars and proline (i.e., osmolytes) compared to non-tolerant genotypes (Farissi et al. 2011).

In this study, the influence of salinity on germination and seedling root growth is quantified in a glycophytic annual plant, the legume *Medicago truncatula*.

The study system is four natural populations of *M. truncatula* originating in northern Tunisia that differ in soil salinity levels, representing two saline and two non-saline environments (Lazrek et al. 2009). Having in mind a hypothesis of salinity adaptation, the expectation that saline origin genotypes display greater tolerance to salinity during seed germination and seedling root growth was tested. Responsiveness to ABA, absolute Na^+ or K^+ levels, and Na^+:K^+ ratios among genotypes of saline and non-saline origin were also compared.

Material and methods

Medicago truncatula var. *truncatula* (Fabaceae) is a highly selfing annual legume that occurs in a wide range of environments throughout the Mediterranean region (Lazrek et al. 2009). Genotypes from four northern Tunisian populations were used in this study: two populations from saline soils [Enfidha (TN1) and Soliman (TN8)] and two populations from non-saline soils [El Kef (TN7) and Bulla Regia (TN9)]. The TN7 population was an olive grove and TN9 is the site of a Roman-era bath, and both sites are uniformly non-saline (Lazrek et al. 2009; Arraouadi et al. 2011; Castro et al. 2013). The two saline populations are from coastal salty fields or sebkhas with soil salinity being five to 14 times greater than that of the two non-saline populations (Lazrek et al. 2009; Castro et al. 2013). The original collections of TN1 and TN8 occurred in highly saline sites with soil electrical conductivities (EC) over 4 dS, though the ranking of greater soil salinity differed between method or soils sampled by Lazrek et al. (2009) and Castro et al. (2013). The substrate within 500 m of each site tends to be more variable, with non-saline sandy patches interspersed (Arraouadi et al. 2011). In addition to differing in soil salinity, Castro et al. (2013) found that the two saline populations have lower soil nitrogen concentrations but greater magnesium concentrations than non-saline populations. Thirty-nine individual genotypes were assayed,

including nine genotypes from TN7 and ten genotypes each from TN1, TN8, and TN9. Only nine genotypes from TN7 were used because one of the genotypes was identified as *M. littoralis* and was therefore not included in this study. We also used the reference genotype A17, moderately tolerant to salinity (Limami et al. 2007), and widely used for molecular, developmental and physiological studies of *M. truncatula* and the first genotype of this species to be sequenced (Young et al. 2011). These lines have been reared through multiple generations of selfing since their initial collection (Lazrek et al. 2009). In this study, plants for seed were reared in greenhouse common gardens either under non-saline conditions (default) or in the same soil amended to 100 mM NaCl. Seeds were removed from the pods and scarified prior to sowing to open the hard seed coat and break dormancy, therefore our estimate of germination is considered to be a measure of embryo viability when exposed to saline and non-saline conditions, rather than breaking of dormancy cued by salinity.

Germination experiment

Laboratory experiment: quantifying germination under a range of NaCl concentrations

To quantify the effects of origin, parental and offspring environments on the proportion and timing of germination, seeds were collected from each of the 39 genotypes grown under either 0 or 100 mM NaCl throughout their lifespan. Each experimental replicate was based on five seeds of each genotype placed in Petri dishes filled with sand that had been saturated with 0, 15, 30, 50, 75, 100, or 150 mM NaCl solution, and the experiment was repeated with four-fold replication. Prior to germination, seeds were scarified to break dormancy, cold treated at 4^{o}C for five days, then transferred to dark and maintained at room temperature. Germination was scored every two days starting the fifth day after cold treatment until seeds had ceased to germinate (day 23). Seeds were considered to germinate when cotyledons emerged. Seeds that failed to germinate after 23 days were determined to be macerated and therefore non-viable; thus the proportion of germinated seeds

at 23 days is also the maximal value. Time to germination was calculated using median time to 50% germination (T_{50}) per Petri dish.

Root growth experiments

NaCl and soil origin effects on seedling root growth and development

To quantify the effects of salinity on early root growth, seedlings were grown in aeroponic chambers (Penmetsa and Cook 2000) using sterile nutrient media (Lullien et al. 1987) and maintained in a growth chamber with a 16-h photoperiod (20^{o}C night/22^{o}C day). Twenty seedlings from each of 11 genotypes were used, including three genotypes from each saline population (TN1: 1.13, 1.15, 1.21; TN8: 8.22, 8.4, 8.15) and two or three genotypes from non-saline populations (TN7: 7.17, 7.22; TN9: 9.12, 9.20 and 9.21). At three days post germination, the aeroponic nutrient solution was amended with either 0 or 75 mM NaCl. NaCl concentrations were selected based on pilot work with *M. truncatula* in which 75 mM NaCl was estimated to cause half-maximal (ED_{50}) growth inhibition (Cordeiro, unpublished data). Prior to NaCl treatment (day 0) two seedlings of each genotype were harvested and measured, followed by five seedlings per genotype at each of days 7 and 14. Roots were separated from shoot tissue and imaged using a double-sided scanner and WinRHIZO software (Arsenault et al. 1995) to quantify total root length (0.001 mm) and the number of lateral roots. Plants were dried to constant weight and weighed to the nearest 0.001 mg using a CAHN microbalance. Root:shoot ratios were calculated as dry root biomass divided by dry shoot biomass.

Effects of NaCl and ABA on early root growth

To test the hypothesis that ABA regulates root responses to salinity, replicates of five genotypes (A17, TN7.11, TN7.23, TN8.23 and TN8.28) were grown under fully factorial treatments of ABA (i.e., 0 or 25μM) and NaCl (i.e., 0 or 50 mM). Two-day-old seedlings were transferred to 25×5 cm square Petri dishes containing Lullien medium (as above) solidified with 1% agar and supplemented

with one of the four ABA-NaCl treatments. ABA concentrations are consistent with exogenous ABA levels used for the study of ABA responses in numerous plant species (e.g., Shukla et al. 2006; Ariel et al. 2010; Gurmani et al. 2011). Initially we validated these ABA concentrations in a pilot study to determine the concentration of ABA required to elicit a half-maximal growth response (ED_{50}) in *M. truncatula* (Cordeiro, unpublished data). Each genotype was replicated 12 to 20 times for each ABA-NaCl treatment. Radicles were covered with sterile filter paper and aluminum foil was used to maintain darkness. Plates were placed in a growth chamber with a 16-h photoperiod (20°C night/22°C day). Primary root length was measured daily to the nearest 0.01 mm using a digital caliper, while the number of lateral roots was recorded at the end of the 11-day experiment.

Saline and non-saline origin root growth responses to ABA

Responsiveness to ABA treatment was measured for twelve genotypes, including three from each saline (TN1: 1.1, 1.11, 1.21; TN8: 8.3, 8.23, 8.24) and non-saline (TN7: 7.13, 7.19, 7.23; TN9: 9.5, 9.12, 9.22) origin population. Assays were conducted in 15×15 cm square Petri dishes containing Lullien media solidified with 1% agar and either 0 or 25μM ABA. Each treatment was replicated ten times for all twelve genotypes. After nine days, total primary root length was measured to the nearest 0.01 mm using a digital caliper and the number lateral roots was recorded.

Effects of sodium and potassium ion on root growth and root tip mortality

To test the impact of Na^+ and K^+ concentrations and Na^+:K^+ ratios on root growth, replicates of two genotypes (A17 and TN7.23) were used in a fully factorial experiment with four concentrations of NaCl and KCl (0, 2, 10, 50 mM). Seedlings were grown in 25×25 cm square Petri dishes as described above, except that Lullien medium solidified with 1% agar was supplemented with one of the 16 different combinations of NaCl and KCl concentrations. Primary root length was quantified to the nearest 0.01 mm using a digital caliper over a 13-day period (1, 4, 6, 8, 11 and 13 after treatment). On day 13, the total number of alive and dead

root tips was counted per plant, with a root tip considered dead when the tip was necrotic and the root had ceased to grow. Each census consisted of four to eight replicates of each genotype and treatment combination.

Data analyses

All analyses were performed in SAS v 9.3 (SAS Institute, 2011). For the laboratory germination experiment, total proportion of seeds that germinated was analyzed using generalized linear mixed models (PROC GLIMMIX using events/trials syntax) treating germination as binary (0- did not germinate, 1-germinated) with origin, population nested within origin, parental and offspring environment treated as fixed effects and genotype nested within population and origin as random effects. Because our focus is on similarities among saline and non-saline populations (i.e., soil origin), with only two populations per soil origin, we treated population as a fixed effect rather than a random effect. Thus interpretations of population are restricted to the populations used in this experiment rather than generalizations of a random sampling of populations. Germination proportions were greater than 0.95 for the 0, 15, and 30 mM NaCl offspring salinity treatments and were excluded from the analyses (Fig. S1, Supplemental Material). Age at germination was tested using the same model as germination proportion, except age at germination was treated as a continuous variable with a negative binomial distribution (PROC GLIMMIX, LINK= LOG, DIST=NEGBIN).

For the aeroponic growth experiment, which included genotypes from saline and non-saline origins, the parameters of soil origin, population, and genotype were added to the model and analyzed using a mixed-model ANCOVA (PROC MIXED). Genotype nested within population and soil origin was treated as a random effect, while all other factors were treated as fixed effects, and time was treated as the covariate. Analysis of covariance was used (on root length, root:shoot, and root biomass) because of our interest in differences in growth rates (slopes) rather than mean trait values at specific times. Pairwise comparison of

slopes was used to test the significance of interactions with time. To meet model assumptions of normality and homoscedasticity, root length was natural logarithm transformed, and root biomass and root:shoot ratios were square root transformed. The number of lateral roots at the last harvest date was analyzed using generalized linear models treating the number of lateral roots as Poisson distributed (PROC GENMOD LINK=LOG, DIST=POISSON).

To quantify individual root growth rates in the NaCl-ABA and NaCl-KCl experiments, linear regressions of natural logarithm transformed root length over time (PROC REG) were performed. The slopes from these regressions were then used as the response variable in fixed-effect ANOVAs (PROC GLM) testing the effects of treatments and origin. Primary root mortality due to NaCl-KCl was tested using generalized linear models (PROC GENMOD LINK=LOGIT DIST=BIN) where root tip death was treated as a binary distribution (0 - alive, 1 - dead). Because primary root death only occurred in the 10 and 50 mM NaCl treatments, data from seedlings treated with less than 10 mM NaCl were excluded from the analyses. For the ABA experiment, where root length was measured at harvest, total root length was analyzed with a fixed-effect ANOVA (PROC GLM). Total root length was square root transformed to meet ANOVA assumptions of normality and homoscedasticity.

For all analyses, tests of random effects were conducted using one-tailed Z tests from random estimates in the model (PROC GLIMMIX) or χ^2 values calculated from differences in -2LL scores from models with and without the random effect (PROC MIXED) using 1 degree of freedom (Littell *et al.* 2006). All post hoc comparisons of means were performed using LSMEANS unless stated otherwise. We report significance values from these comparisons in the graphs. Data in supplemental text includes significance values from sequential Bonferroni tests to account for multiple comparisons of means (Table S1, Supplemental Material).

Results

Saline origin genotypes display greater salinity tolerance in germination traits

As a prelude to germination experiments, thirty-nine *Medicago truncatula* genotypes originating from four Tunisian populations (two saline, two non-saline) were grown in a greenhouse under non-saline (0 mM NaCl) or saline (100mM NaCl) conditions. Seeds from these individuals were assayed for proportion and timing of germination under a range of NaCl concentrations. Germination was high (96%) for NaCl concentrations less than 50 mM (Fig. S1, Supplemental Material) and thus these data were excluded from subsequent analysis on germination proportions. When NaCl concentrations were greater than 75 mM, saline origin genotypes displayed 15.7 % greater germination than non-saline origin genotypes, with the greatest difference between soil origins (30.7%) occurring at 150 mM NaCl (Table 1, Fig. 1A). The greater germination proportion of saline origin genotypes relative to non-saline origin genotypes as salinity increases is consistent with greater salinity tolerance of saline compared to non-saline origin populations.

Table 1. Test statistics (χ^2 and *F*) and significance from generalized linear mixed models on proportion seed germinated and time to germination (T_{50}) for seeds from saline and non-saline origin genotypes. Parents were grown in 0 or 100 mM NaCl and offspring environment treatments were 75, 100, and 150 mM NaCl. Time to germination was recorded for seeds grown under 0, 15, 30, 50, 75, 100, and 150 mM NaCl offspring environments.

Source	Germination		T_{50} time to germination (days)	
	f	χ^2	f	*F*
Origin	1	11.66***	1	9.08**
Population (O)	2	1.06	2	3.13t
Parental Environment	1	38.79****	1	6.00*
Offspring Environment	2	659.37****	6	55.20****
PE X OE	2	5.41t	6	0.59
Origin X PE	1	0.37	1	1.57
Origin X OE	2	15.13***	6	1.84
Origin X PE X OE	2	0.66	6	0.19
Population (O) X PE	2	21.18****	2	0.28
Population (O) X OE	4	12.86*	2	1.81*
Population (O) X PE X OE	4	7.35	12	0.10

Notes: t $P < 0.10$, * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, **** $P < 0.0001$; Origin/O- Soil origin of the genotype (i.e., genotype from saline or non-saline soils), PE- parental environment, OE- offspring environment.

To test the influence of parental environment on the proportion of germination, parental plants were grown under either 0 or 100 mM NaCl. As shown in Table 1, seeds from saline-grown parents had reduced germination (by 12.3%) relative to seeds from non-saline-grown parents. This was true irrespective of whether the genotypes originated from saline or non-saline environments in Tunisia, with the exception of seeds from saline population TN1 where germination proportion remained high irrespective of parental environment salinity (Fig. 1B).

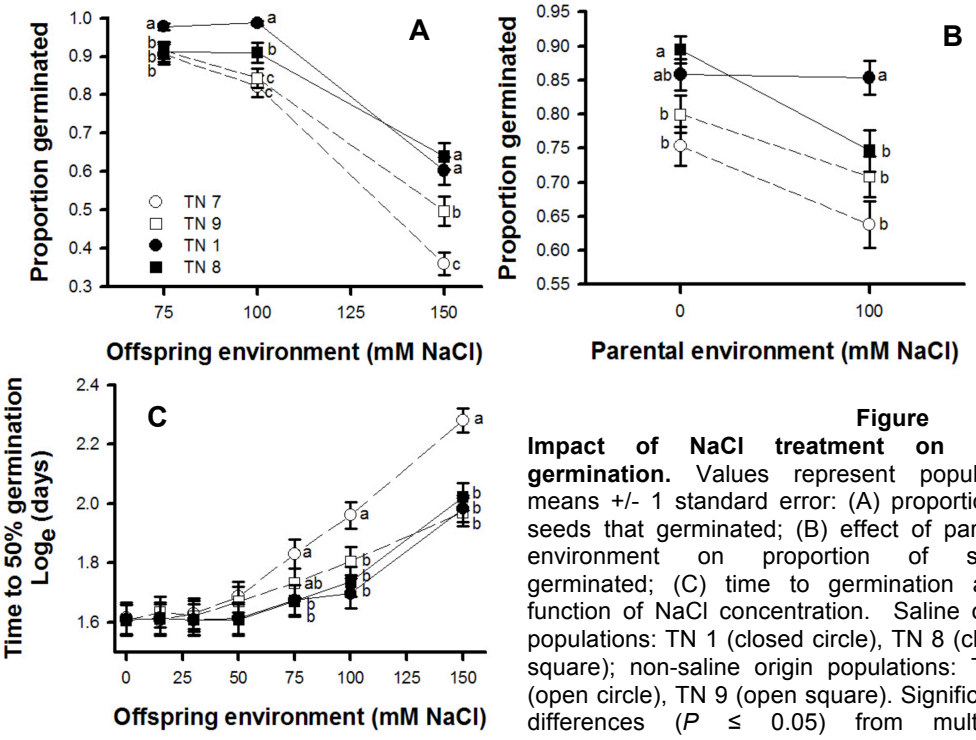


Figure 1. Impact of NaCl treatment on seed germination. Values represent population means +/- 1 standard error: (A) proportion of seeds that germinated; (B) effect of parental environment on proportion of seeds germinated; (C) time to germination as a function of NaCl concentration. Saline origin populations: TN 1 (closed circle), TN 8 (closed square); non-saline origin populations: TN 7 (open circle), TN 9 (open square). Significance differences ($P \leq 0.05$) from multi-test comparison of population by treatment interactions are shown with different letters (Table S1, see Supplemental Material).

Population differentiation was also observed for the effect of salinity on time to germination, with seeds from TN7 (non-saline origin) showing a significantly greater delay in germination at 75, 100 and 150 mM NaCl relative to the other three populations (Table 1, Fig. 1C). As shown in Table 1, parental effects on time to germination were small [4.4% delay in germination when parents were grown in 100 mM (5.8 days) vs 0 mM (5.5 days) NaCl] compared to the effect of offspring

environment [20% delay in germination in the 100 mM NaCl (6.0 days) vs. 0 mM (5.0 days) NaCl offspring environment]. Additionally, within population genetic variation was detected for proportion of seeds germinated (Z = 3.84, *P* < 0.0001) and time to germination (Z = 1.67, *P* = 0.0478).

Saline origin genotypes are less affected by salt during early root growth and development

To quantify the effects of soil origin and offspring environment salinity on seedling root growth patterns, we measured root traits for seedlings growing in saline (75 mM NaCl) or non-saline (0 mM NaCl) conditions. Saline growth conditions were associated with decreased root elongation rates (Table 2, Fig. 2A) and decreased root:shoot ratios (Figure 2B) irrespective of origin environment.

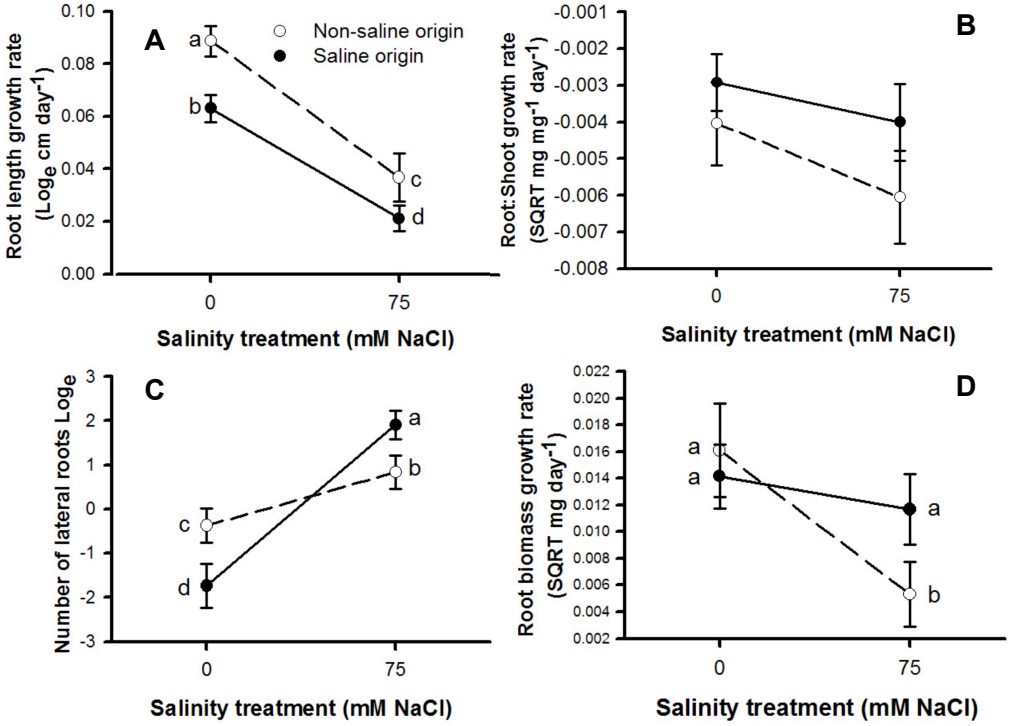


Figure 2. Impact of NaCl treatment on root growth parameters. Values represent means +/- 1 standard error: (A) root length; (B) root:shoot biomass; (C) number of lateral roots 14 days post treatment; (D) root biomass. Symbols specify saline (closed symbols) and non-saline origin (open symbols) genotypes grown in media supplemented with either 0 mM (circle with solid line) or 75 mM (square with dashed line) NaCl. Significance differences ($P \leq 0.05$) from pair-wise comparisons of the means at day 14 are shown with different letters (Table S1, Supplemental Material).

By contrast, we observed origin-specific differences in the relative sensitivity of lateral root formation and biomass accumulation under saline conditions. Thus, under saline conditions, lateral root formation was significantly greater for saline-origin genotypes, and the converse was true of non-saline origin genotypes under non-saline conditions. In particular, saline origin genotypes produced 2.9-fold more lateral roots under saline conditions, and non-saline origin genotypes produced 3.9-fold more lateral roots under non-saline conditions, compared to genotypes of non-saline or saline origins, respectively. In addition to significant differences between saline and non-saline origin populations, we also observed significant within-population variation for the lateral root response (Table 2).

Table 2. Test statistics (χ^2 and *F*) and significance for a mixed-model ANCOVA for root traits [i.e., root length (cm), root biomass (mg), root:shoot (mg/mg)] and a generalized linear fixed-effect model for the number of lateral roots. Saline and non-saline origin genotypes were grown in 0 and 75 mM NaCl.

Source	f	Root length *F*	Root biomass *F*	Root:shoot *F*	Lateral root number χ^2
Salinity	1	15.15****	0.71	0.00	145.76****
Origin	1	8.17*	0.01	0.93	0.08
Salinity X Origin	1	0.47	0.18	0.33	36.55****
Pop (Origin)	2	1.35	0.51	0.62	1.56
Salinity X Pop (O)	2	0.01	1.68	0.26	6.99**
Time	1	366.43****	85.28****	75.47****	NA
Time X Salinity	1	66.03****	8.84**	3.19t	NA
Time X Origin	1	17.86****	0.33	1.93	NA
Time X Salinity X Origin	1	0.49	4.12*	0.42	NA
Time X Pop(O)	2	10.07****	2.04	2.33t	NA
Time X Salinity X Pop(O)	2	1.96	5.62**	0.77	NA
Genotype(Origin (Pop))[1]	1	3.0t	0.0	26.6****	1.75*

Notes: t $P < 0.10$, * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, **** $P < 0.0001$; See Table 1 for abbreviations. Because lateral root production did not occur until late in the aeroponic experiment, lateral root number was analyzed using ANOVA with data from the last census (day 14).
[1]Statistics for random effect (i.e., genotype) are χ^2 from difference in -2LL scores from models with and without the genotype, except for lateral root number, for which one-tailed Z-test was calculated.

Saline and non-saline origin genotypes also differed in their ability to maintain root biomass accumulation under saline growth conditions, with genetic variation most evident among populations (Table 2). In particular, biomass accumulation in saline origin genotypes was not significantly impacted by salinity, while non-saline origin genotypes were highly sensitive to salinity, with 16.8%

lower root biomass accumulation when grown under 75 mM NaCl compared to 0 mM NaCl. Taken together, the observations reveal significant variation between saline and non-saline origin genotypes in seedling growth responses to salinity.

Root elongation rate is affected similarly by ABA and NaCl, with saline origin genotypes more sensitive to ABA treatment than non-saline origin genotypes

ABA is implicated in tolerance to osmotic stress, which is a component of saline stress and also a trigger of growth retardation (Sreenivasulu et al. 2012). In plate assays with seedlings, the main effect of ABA (and NaCl, also tested above) was to reduce root elongation rates, with independent application of 25 µM ABA and 50 mM NaCl decreasing elongation rates by 4.4 % in each treatment relative to control (Table 3, Fig. 3A). Combined ABA and NaCl treatment, further reduced root elongation in an additive manner (8.3% reduction relative to the control treatment; Table 3, Fig. 3A).

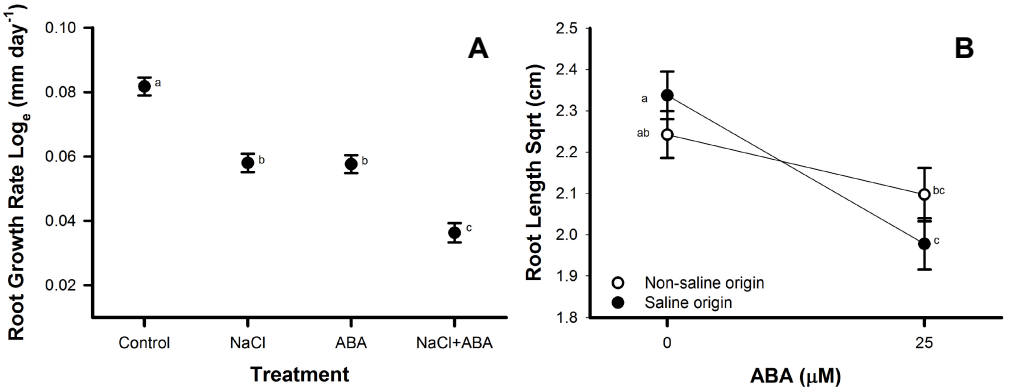


Figure 3. Root growth in response to ABA and NaCl treatment. Values are means +/- 1 standard error: (A) root growth rate, averaged together from both saline and non-saline origin genotypes; (B) root length. Treatments were non-amended media (control), 25 µM of ABA, 50 mM NaCl, and 25 µM of ABA + 50 mM NaCl. In panel B, symbols represent saline (closed circle) and non-saline (open circle) origin genotypes.

Interestingly, the effect of NaCl on growth retardation was effectively reversed by co-treatment with KCl at comparable concentrations. Consistent with results presented above, 50 mM NaCl (but not lower treatment concentrations) significantly reduced root growth rates (Fig. 4A) and also increased root tip mortality (Fig.4B). These negative effects on root growth rate and root tip mortality

were mitigated by KCl in a concentration dependent manner (Table 4), with KCl concentrations as low as 2 mM significantly reversing NaCl-induced root growth retardation and root tip mortality (Fig. 4B; Table 4).

Table 3. Test statistics and significance for primary root growth rate analyzed using fixed-effects ANOVA. Treatments were 0 and 25 µM ABA and 0 and 50 mM NaCl.

Source	Primary root growth rate (cm/day) *F*
ABA/NaCl Experiment	
ABA	64.98****
NaCl	61.87****
ABA X NaCl	0.17
ABA/Origin Experiment	
ABA	17.52****
Origin	0.04
ABA X Origin	3.16t

Notes: t $P < 0.10$, * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, **** $P < 0.0001$;
All factors in the model have 1 degree of freedom.

To determine whether sensitivity to ABA varies among genotypes, we scored root length as a function of ABA in six saline and six non-saline origin genotypes. Interestingly, root length of saline origin genotypes was more affected by ABA relative to non-saline origin genotypes, with ABA resulting in reductions to root length of 31.3 % and 13.9 %, respectively (Table 3, Fig. 3B). The observation that saline origin genotypes are more sensitive to ABA contrasts with results presented above, where saline origin genotypes were observed to be less sensitive to NaCl.

Table 4. Test statistics (χ^2 and F) and significance for the affect of NaCl/KCl on primary root growth rate (two-way fixed effect ANOVA) and root tip mortality (generalized linear model).

Source	f	Primary root growth rate (cm/day) *F*	f	Primary root tip mortality χ^2
NaCl	3	5.88***	1	0.84
KCl	3	1.23	3	6.11
NaCl X KCl	9	2.63**	3	26.42****

Note: t $P < 0.10$, * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, **** $P < 0.0001$

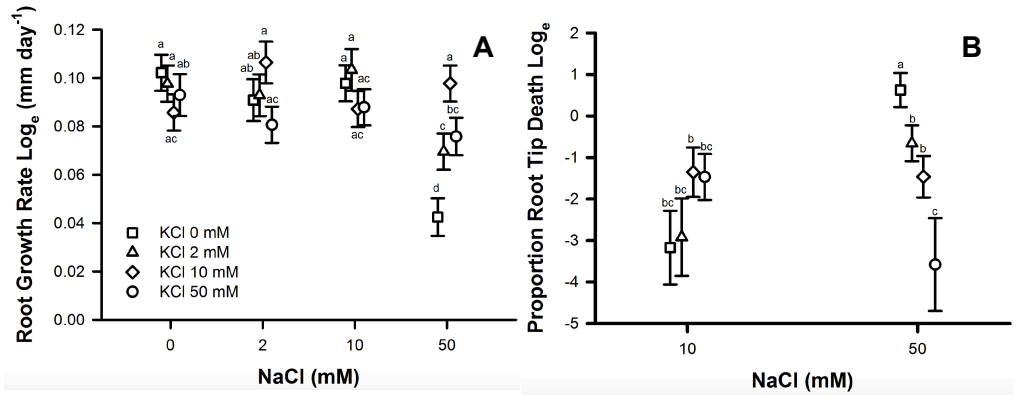


Figure 4. Root growth rate and root tip mortality as a function of NaCl:KCl in two genotypes (the reference A17 and the non-saline origin TN7.23). Values are mean +/- 1 standard error for (A) root growth rate and (B) primary root tip mortality. Seedlings grown at the specified concentration of NaCl and one of four KCl concentrations: 0 mM (square), 2 mM (triangle), 25 mM (diamond), 125 mM (circle).

Discussion

For short-lived annuals, the breadth of environmental conditions under which a seed will germinate and establish are key determinants of species distribution. Here, we examined the influence of historical evolutionary and ecological processes selection and parental environment on salinity tolerance in *Medicago truncatula*, focusing specifically on seed and seedling traits. Salinity is a major abiotic stress that limits agricultural production and shapes plant communities. Our experimental system takes advantage of four increasingly well-studied populations (two saline and two non-saline) of *M. truncatula* that occur in northern Tunisia (e.g., Lazrek et al. 2009; Friesen et al. 2010; Castro et al. 2013). Consistent with the expectation that natural selection acts to increase salinity tolerance, especially on early seedling traits, saline origin genotypes differed in numerous responses to NaCl compared their non-saline origin counterparts (Figs. 1 and 2). Although the current study focused primarily on responses to NaCl, we also observed population differentiation with respect to ABA sensitivity (Fig. 3), suggesting a link between responsiveness to this important stress hormone and salinity adaptation in these *M. truncatula* populations. We discuss these results in

the context of known effects of salinity on germination and seedling establishment, adaptation to salinity, and mechanisms of salinity tolerance.

Germination is the plant's first action and can influence how individuals perform in subsequent developmental phases. We observed population differentiation for both germination proportion (Table 1, Fig. 1A) and time to germination (Fig. 1C) under saline conditions. In particular, germination of seeds from saline origin genotypes was less sensitive to salinity than was germination of seeds from non-saline origin populations (Figure 1A). We also observed differences that were specific to individual populations; thus seeds from TN7, a non-saline origin population, exhibited significantly delayed germination in relation to other populations (Figure 1C). In all cases seeds that failed to germinate were macerated and therefore non-viable, making germination rates a proxy for viability on exposure to salinity, as has been observed in other glycophytes including closely related *Medicago* species (Chérifi et al. 2011; Farissi et al. 2011). These effects of salinity on germination in *M. truncatula* are in agreement with results obtained with other *Medicago* species (Chérifi et al. 2011; Farissi et al. 2011; Khalil et al. 2011) and species of other genera (Smith and McComb 1983; Carlson et al. 1983; Petkova et al. 1995; Song et al. 2008; but see Melonie et al. 2008), including those demonstrating greater salinity tolerance of saline origin populations (e.g., DiTommaso 2004; van Zandt and Mopper 2004).

In addition to saline environments directly impacting seed and seedling traits, the parental environment can influence seed in the next generation by altering physiological and developmental set points, and potentially through epigenetic mechanisms. We did not observe increased salinity tolerance for germination when parent plants were exposed to stressful saline levels. Rather the main effect of parental exposure to salinity was decreased germination proportions in three of the four populations (Fig. 1C). In the fourth population, namely in TN1, a saline-origin population, germination was not impacted by parental exposure to NaCl. These results suggest that parental environmental effects on germination may contribute to population differentiation (Table 1, Fig. 1B), as documented in

other systems (Sultan 1996; Munir et al. 2001; Galloway, 2005; Galloway and Etterson 2007). Moreover, the observation of phenotypes specific to the saline-origin population TN1 strengthens the idea, raised above, that certain salinity phenotypes are locally idiosyncratic among these *M. truncatula* populations. We speculate that the ability to maintain high seed germination rates in seeds that are reared in saline environments may provide fitness advantage against a subset of saline-related factors that are specific to the TN1 environment. Taken together, these results indicate that parental and offspring environmental effects contribute differentially to salinity-related germination phenotypes among the four surveyed populations.

In agreement with prior work in other systems (e.g., Ariel et al. 2010; Zolla et al. 2010; Galvan-Ampudia and Testerink 2011, Rahnama et al. 2011; Farissi et al. 2011; Guo et al. 2012), a main effect of salinity on seedling growth was to reduce root elongation rates. It is noteworthy that although non-saline origin genotypes had longer roots independent of salinity (Fig. 2A), they tended to experience a greater reduction in root elongation rates in the presence of salt than did saline origin genotypes (Table 2). Salinity can also suppress (e.g., Rubinigg et al. 2004; Xiong et al. 2006) or stimulate (e.g., Zolla et al. 2010; Galvan-Ampudia and Testerink 2011) lateral root formation, depending on concentration and the system under study (Liang and Harris, 2005). Our results document increased lateral root formation at 75 mM NaCl (Fig. 2C). This pattern of increased lateral root development under saline conditions is consistent with Zahaf et al. (2012), wherein increased lateral root formation was observed in the salinity-tolerant *M. truncatula* genotype TN1.11 (TN1- saline origin population) when exposed to salt. Rahnama et al. (2011) proposed that increased lateral root production under saline conditions allows plants to sustain soil nutrient and water intake while restructuring root architecture to locate optimal microenvironments. Thus maintenance of greater primary and lateral root growth under saline conditions may allow saline tolerant plants to prospect for relatively lower salinity micro-patches to which root functions are preferentially allocated (e.g., Flores et al. 2002). We also observed

that saline origin genotypes were better able to maintain root biomass accumulation under saline conditions (Fig. 2D). Considering the overlapping distributions of root:shoot biomass response to salinity (Fig. 2B), saline origin genotypes more effectively maintain biomass allocation to both root and shoot systems. However, the ability of saline origin genotypes to maintain root biomass accumulation is not due to primary root growth (Fig. 2A), but likely due to increased root branching (Fig. 2C) and root diameter (data not shown). It has been suggested that investment in roots rather than shoots may be adaptive given that it improves water relations (Bayuelo-Jiménez et al. 2002; Munns and Tester 2008; Galvan-Ampudia and Testerink 2011).

Considerable data indicates a link between growth regulation, salt stress and the hormone ABA (Sreenivasulu et al. 2012). Gurmani et al. (2011) demonstrate that seeds of *Oryza sativa* soaked in ABA and then sown into saline environments were able to maintain Na^+:K^+ homeostasis, whereas control seeds developed ionic imbalance due to high intracellular Na^+. Interestingly, transgenic tobacco expressing a chickpea gene (CAP2, a protein coding gene with an AP2/ERF domain - APETALA2/ethylene-responsive factor) was found to have enhanced osmotic and salinity stress tolerance and one of the associated phenotypes was an increase in ABA-dependent lateral root formation (Shukla et al. 2006). Here we report that saline and non-saline genotypes of *M. truncatula* have contrasting responses to NaCl and ABA (Compare Figs. 2C,D and 3B), as root growth of saline origin genotypes was more affected by ABA and less affected by salt relative to non-saline origin genotypes. Nevertheless, ABA was generally suppressive of root growth at the concentration tested, and the effect was additive with that of NaCl. In their study of *M. truncatula*, Ariel *et al* (2010) observed that osmotic stress decreased lateral root formation in a moderately tolerant genotype given that ABA mediates growth responses to osmotic stress and that osmotic stress is a component of salinity stress, our results combined with those of Ariel *et al* indicate a potentially complex interaction among the underlying pathways. In any case, these initial studies, and in particular the observation that saline and non-

saline origin genotypes differ in their relative ABA responsiveness, highlight the need for further investigation into the role of ABA in salinity adaptation in these *M. truncatula* populations. Of relevance to this issue, Friesen et al. (in review) identified candidate genes for salinity tolerance in these same Tunisian populations, with inferred roles in ABA signaling. Candidate genes were identified based on allele assortment with soil origin and include Medtr3g098090, an ortholog of AtCPK12 which is a negative regulator of ABA signaling in germination and early growth of Arabidopsis (Zhao et al. 2011), and Medtr4g128820, a CBL interacting protein kinase (CIPK, an ortholog of AtCIPK21) which belongs to a gene family that has been associated not only with ABA signaling, but also with Na^+ homeostasis and K^+ uptake (Weinl and Kudla 2009).

Recent studies suggest that ion balance of Na^+ and K^+ may be more important in the maintenance of physiological processes under saline conditions than is the absolute concentration of Na^+ (Maathuis and Amtmann 1999; Chartzoulakis et al. 2002; Meloni et al. 2008). In the current study, exogenous application of KCl counteracted the negative effects of NaCl on primary root growth and root mortality (Fig. 4A,B). Moreover, when KCl was greater than NaCl concentration, KCl had small negative effects on root growth and root survival (Fig. 4A,B). These patterns support the generalization that maintenance of Na^+:K^+ ratios within the plant is important for salinity tolerance. Measurements of Na^+ and K^+ concentrations in whole tissue samples are part of a recent reciprocal field study focusing on *M. truncatula* populations originating from saline and non-saline environments in Portugal (Cordeiro, submitted).

Conclusion

While direct evidence for adaptive evolution requires quantifying fitness (e.g., via a reciprocal transplant experiment), our observations that saline-origin genotypes are more able to sustain germination and growth under saline conditions compared to non-saline-origin genotypes are consistent with early life

cycle adaptations for salinity tolerance in saline-origin populations. Others have observed a correlation between seed and seedling salinity tolerance (Redondo-Gomez et al. 2008; Song et al. 2008; but see Bayuelo-Jiménez et al. 2002; Reviewed by Donohue et al. 2010) and our results document a similar situation in Tunisian populations of *M. truncatula*. Of potential practical importance, these data suggest that rapid screens for germination and/or early root growth traits under saline conditions may offer a simple means to identify genotypes likely to possess adaptations to saline conditions. Although not addressed here, a factor unique to legumes and certain allied taxa (e.g., *Alnus* spp.) is the sensitivity of symbiotic nitrogen fixation to abiotic stress, including soil salinity and drought. Moreover, rhizobia may influence salinity tolerance (Zahran 1999; Carelli et al. 2000; Bianco and Defez 2009). Interestingly, in addition to its role in abiotic stress responses, ABA also affects nodulation (Ding et al. 2005; Liang and Harris, 2005) and is increasingly recognized as a modulator of plant-microbe interactions (Robert-Seilaniantz et al. 2011; Choi et al. 2013). Thus, salinity tolerance may be more complex in legumes than other plant systems and these are areas ripe for further more detailed investigation, potentially using these Tunisian populations for *M. truncatula* as a study system.

Acknowledgements

Matilde Cordeiro (MAC) conceptualized, designed and performed experiments, analyzed and interpreted the data, and wrote the manuscript. As mentioned in the published manuscript: MAC, KSM, KEM and TDF designed and performed experiments; SVN, EVW and DRC contributed with reagents and rationale; MAC, KSM, EVW and DRC wrote the paper with input from the other authors.

The authors thank V. Penmetsa for advice and help setting up and interpreting the laboratory experiments; L. Vance, B. Castro, K. Rowley, A. Waggoner, and H. Yokoyama-Hatch for help with the various experiments; and

M.L. Friesen, S.Y. Strauss, B.K. Riely, E. Larrainzar, and W. Vu for comments on the manuscript and interpretation of results. This work was supported by NSF PGRP 08-20846. M. Cordeiro was supported by FCT fellowship SFRH/BD/39905/2007. T. Fotinos was supported by NIH/NIGMS R25 GM061347. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

References

Agrawal AA (2001) Transgenerational consequences of plant responses to herbivory: an adaptive maternal effect? *American Naturalist*, **157**, 555–569.

Ariel F, Diet A, Verdenaud M, Gruber V, Frugier F, Chan R, Crespi M (2010) Environmental regulation of lateral root emergence in *Medicago truncatula* requires the HD-Zip I transcription factor HB1. *Plant Cell*, **22**, 2171–2183.

Arraouadi S, Chardon F, Huguet T, Aouani ME, Badri M (2011) QTLs mapping of morphological traits related to salt tolerance in *Medicago truncatula*. *Acta physiologiae plantarum,* **33**, 917-926.

Arsenault J-L, Poulcur S, Messier C, Guay R (1995) WinRHlZO™, a root-measuring system with a unique overlap correction method. *HortScience,* **30**, 906-906.

Ashraf M, McNeilly T, Bradshaw AD (1987) Selection and heritability of tolerance to sodium chloride in four forage species. *Crop Science,* **27**, 232-234.

Ballhorn DJ, Kautz S, Jensen M, Schmitt I, Heil M, Hegeman AD (2011) Genetic and environmental effects on plant defences against herbivores. *Journal of Ecology*, **99**, 313–326.

Bayuelo-Jimenez JS, Craig R, Lynch JP (2002) Salinity tolerance of species during germination and early seedling growth. *Crop Science,* **42**, 1584-1594.

Bianco C, Defez R (2009) *Medicago truncatula* improves salt tolerance when nodulated by an indole-3-acetic acid-overproducing *Sinorhizobium meliloti* strain. *Journal of experimental botany,* **60**, 3097-3107.

Bibee K, Shishido K, Hathaway RP, Heschel MS (2011) Population differentiation of *Impatiens capensis* (Balsaminaceae) at the range limit. *International Journal of Plant Sciences,* **172**, 211-219.

Bischoff A, Müller-Schärer H (2010) Testing population differentiation in plant species--how important are environmental maternal effects. *Oikos,* **119**, 445-454.

Boursiac Y, Chen S, Luu DT, Sorieul M, van den Dries N, Maurel C (2005) Early effects of salinity on water transport in Arabidopsis roots. Molecular and cellular features of aquaporin expression. *Plant physiology,* **139**, 790-805.

Carelli, M., S. Gnocchi, S. Fancelli, A. Mengoni, D. Paffetti, C. Scotti, and M. Bazzicalupo. 2000. Genetic diversity and dynamics of *Sinorhizobium meliloti* populations nodulating different alfalfa cultivars in Italian soils. *Applied and Environmental Microbiology* 66: 4785–4789.

Carlson JR, Ditterline RL, Martin JM, Sands DC, Lund RE (1983) Alfalfa seed germination in antiobiotic agar containing NaCl. *Crop science,* **23**, 882-885.

Castro BM, Moriuchi KS, Friesen ML, Badri M, Nuzhdin SV, Strauss SY, Cook DR, Wettberg E (2013) Parental environments and interactions with conspecifics alter salinity tolerance of offspring in the annual *Medicago truncatula*. *Journal of Ecology,* **101**, 1281-1287.

Chang CW (1961) Effect of irrigation water and exchangeable sodium on soil properties and growth of alfalfa. *Soil Science*, **91**, 29–37.

Chartzoulakis K, Loupassaki M, Bertaki M, Androulakis I (2002) Effects of NaCl salinity on growth, ion content and CO^2 assimilation rate of six olive cultivars. *Scientia Horticulturae,* **96**, 235-247.

Chérifi K, Boubaker H, Msanda F, Saadi B, Boufous E, Mousadik AE (2011) Variability for salt tolerance during germination in *Medicago ciliaris* (L.) and *Medicago polymorpha* (L.). *International Research Journal of Plant Science,* **2**, 201-208.

Choi HK, Iandolino A, da Silva FG, Cook DR (2013) Water deficit modulates the response of *Vitis vinifera* to the Pierce's disease pathogen *Xylella fastidiosa*. *Molecular plant-microbe interactions : MPMI,* **26**, 643-657.

Davies WJ, Kudoyarova G, Hartung W (2005) Long-distance ABA signaling and its relation to other signaling pathways in the detection of soil drying and the mediation of the plant' s response to drought. *Journal of Plant Growth*

Regulation, **24**, 285-295.

De Smet I, Signora L, Beeckman T, Inzé D, Foyer CH, Zhang H (2003) An abscisic acid-sensitive checkpoint in lateral root development of Arabidopsis. *The Plant journal,* **33**, 543-555.

Ding Y, Kalo P, Yendrek C, Sun J, Liang Y, Marsh JF, Harris JM, Oldroyd GE (2008) Abscisic acid coordinates nod factor and cytokinin signaling during the regulation of nodulation in *Medicago truncatula*. *The Plant cell,* **20**, 2681-2695.

DiTommaso A (2004) Germination behavior of common ragweed (*Ambrosia artemisiifolia*) populations across a range of salinities. *Weed Science*, **52**, 1002–1009.

Donohue K (2002) Germination timing influences natural selection on life-history characters in *Arabidopsis thaliana*. *Ecology*, **83**, 1006–1016.

Donohue K, Dorn L, Griffith C, Kim E, Aguilera A, Polisetty CR, Schmitt J (2005) The evolutionary ecology of seed germination of *Arabidopsis thaliana*: variable natural selection on germination timing. *Evolution,* **59**, 758-770.

Donohue K, de Casas RR, Burghardt L, Kovach K, Willis CG (2010) Germination, postgermination adaptation, and species ecological ranges. *Annual Review of Ecology, Evolution and Systematics,* **41**, 293-319.

FAO (2008) FAO Land and Plant Nutrition Management Service. Available at http://www.fao.org/ag/agl/agll/spush.

Farissi M, Bouizgaren A, Faghire M, Bargaz A, Ghoulam C (2011) Agro-physiological responses of Moroccan alfalfa (*Medicago sativa* L.) populations to salt stress during germination and early seedling stages. *Seed Science and Technology,* **39**, 389-401.

Flores P, Angeles Botella M, Martínez V, Cerdá A (2002) Response to salinity of tomato seedlings with a split-root system: nitrate uptake and reduction. *Journal of Plant Nutrition,* **25**, 177-187.

Flowers TJ, Yeo AR (1995) Breeding for salinity resistance in crop plants: where next? *Functional Plant Biology,* **22**, 875-884.

Friesen ML, Cordeiro MA, Penmetsa RV, Badri M, Huguet T, Aouani ME, Cook DR, Nuzhdin SV (2010) Population genomic analysis of Tunisian *Medicago truncatula* reveals candidates for local adaptation. *The Plant journal,* **63**, 623-635.

Galloway LF (2005) Maternal effects provide phenotypic adaptation to local environmental conditions. *New Phytologist*, **166**, 93–100.

Galloway LF, Etterson JR (2007) Transgenerational plasticity is adaptive in the wild. *Science*, **318**, 1134–1136.

Galloway LF, Fenster CB (2000) Population differentiation in an annual legume: local adaptation. *Evolution*, **54**, 1173–1181.

Galvan-Ampudia CS, Testerink C (2011) Salt stress signals shape the plant root. *Current Opinion in Plant Biology*, **14**, 296–302.

Guo Y, Jia W, Song J, Wang D, Chen M, Wang B (2012) *Thellungilla halophila* is more adaptive to salinity than *Arabidopsis thaliana* at stages of seed germination and seedling establishment. *Acta physiologiae plantarum,* **34**, 1287-1294.

Gurmani AR, Bano A, Khan SU, Din J, Zhang JL (2011) Alleviation of salt stress by seed treatment with abscisic acid (ABA), 6-benzylaminopurine (BA) and chlormequat chloride (CCC) optimizes ion and organic matter accumulation and increases yield of rice (*Oryza sativa* L.). *Australian Journal of Crop Science,* **5**: 1278.

Hauser F, Horie T (2010) A conserved primary salt tolerance mechanism mediated by HKT transporters: a mechanism for sodium exclusion and maintenance of high K^+/Na^+ ratio in leaves during salinity stress. *Plant, Cell and Environment*, **33**, 552–565.

Hereford J, Moriuchi KS (2005) Variation among populations of *Diodia teres* (Rubiaceae) in environmental maternal effects. *Journal of Evolutionary Biology*, **18**, 124–131.

Katembe WJ, Ungar IA, Mitchell JP (1998) Effect of salinity on germination and seedling growth of two Atriplex species (Chenopodiaceae). *Annals of botany,* **82**, 167-175.

Kent LM, Läuchli A (1985) Germination and seedling growth of cotton: salinity-calcium interactions. *Plant, cell & environment,* **8**, 155-159.

Khalil C, Boufous E, Mousadik AE (2011) Diversity of salt tolerance during germination in *Medicago ciliaris* (L.) and *Medicago polymorpha* (L.). *Atlas Journal of Plant Biology,* **1**, 6-12.

Kronzucker HJ, Britto DT (2011) Sodium transport in plants: a critical review. *New phytologist,* **189**, 54-81.

Lazrek F, Roussel V, Ronfort J, Cardinet G, Chardon F, Aouani ME, Huguet T (2009) The use of neutral and non-neutral SSRs to analyse the genetic structure of a Tunisian collection of *Medicago truncatula* lines and to reveal associations with eco-environmental variables. *Genetica,* **135**, 391-402.

Li X, Zhang X, Song J, Fan H, Feng G, Wang B (2011) Accumulation of ions during seed development under controlled saline conditions of two Suaeda salsa populations is related to their adaptation to saline environments. *Plant and soil,* **341**, 99-107.

Liang Y, Harris JM (2005) Response of root branching to abscisic acid is correlated with nodule formation both in legumes and nonlegumes. *American journal of botany,* **92**, 1675-1683.

Limami AM, Ricoult C, Planchet E, González EM, Ladrera R, Larrainzar E, Arrese-Igor C, Merchan F, Crespi M, Frugier F (2007) Response of Medicago truncatula to abiotic stress. *The Medicago truncatula handbook,* The Samuel

Noble Foundation, Ardmore, Oklahoma, USA.

Littell RC, Milliken GA, Stroup WW, Wolfinger RD, Schabenberger O (2006) SAS for mixed models. SAS Institute. *Inc.,* Cary, North Carolina, USA.

Lullien V, Barker DG, De Lajudie P, Huguet T (1987) Plant gene expression in effective and ineffective root nodules of alfalfa (*Medicago sativa*). *Plant molecular biology,* **9**, 469-478.

Maas EV (1986) Salt tolerance of plants. *Applied Agricultural Research*, **1**, 12–26.

Maathuis FJM, Amtmann A (2004) K Nutrition and Na toxicity: The basic of cellular K/Na ratios. *Annals Botany,* **84**, 123-133.

Malamy JE (2005) Intrinsic and environmental response pathways that regulate root system architecture. *Plant Cell and Environment*, **28**, 67–77.

Meloni DA, Gulotta MR, Martinez CA (2008) Salinity tolerance in *Schinopsis quebracho* colorado: Seed germination, growth, ion relations and metabolic responses. *Journal of Arid Environments,* **72**, 1785-1792.

Munir J, Dorn LA, Donohue K, Schmitt J (2001) The effect of maternal photoperiod on seasonal dormancy in *Arabidopsis thaliana* (Brassicaceae). *American journal of botany,* **88**, 1240-1249.

Munns R, Tester M (2008) Mechanisms of salinity tolerance. *Annual review of plant biology* **59**, 651-681.

Nibau C, Gibbs DJ, Coates JC (2008) Branching out in new directions: the control of root architecture by lateral root formation. *New phytologist,* **179**, 595-614.

Osmont KS, Sibout R, Hardtke CS (2007) Hidden branches: developments in root system architecture. *Annual review of plant biology,* **58**, 93-113.

Penmetsa RV, Cook DR (2000) Production and characterization of diverse developmental mutants of Medicago truncatula. *Plant physiology,* **123**, 1387-1398.

Petkova D, Nedjialkov D, Djilianov D (1995) Early screening for drought tolerance in cultivated alfalfa. *Bulgarian Journal of Agricultural Science,* **1**, 429-432.

Rains DW, Epstein E (1965) Transport of sodium in plant tissue. *Science*, **148**: 1611.

Rahnama A, Munns R, Poustini K, Watt M (2011) A screening method to identify genetic variation in root growth response to a salinity gradient. *Journal of experimental botany,* **62**, 69-77.

Redondo-Gómez S, Mateos-Naranjo E, Cambrollé J, Luque T, Figueroa ME, Davy AJ (2008) Carry-over of differential salt tolerance in plants grown from dimorphic seeds of *Suaeda splendens*. *Annals of botany,* **102**, 103-112.

Roach DA, Wulff RD (1987) Maternal effects in plants. *Annual review of ecology and systematics,* **18**, 209-235.

Robert-Seilaniantz A, Grant M, Jones JD (2011) Hormone crosstalk in plant disease and defense: more than just jasmonate-salicylate antagonism. *Annual review of phytopathology,* **49**, 317-343.

Rubinigg M, Wenisch J, Elzenga JTM, Stulen I (2004) NaCl salinity affects lateral root development in *Plantago maritima*. *Functional plant biology,* **31**, 775-780.

SAS Institute (2011) SAS/STAT user's guide, version 9.33. SAS Institute, Cary, North Carolina, USA.

Schenk HJ, Jackson RB (2002) Rooting depths, lateral root spreads and below-ground/above-ground allometries of plants in water-limited ecosystems. *Journal of Ecology,* **90**, 480-494.

Shukla RK, Raha S, Tripathi V, Chattopadhyay D (2006) Expression of CAP2, an APETALA2-family transcription factor from chickpea, enhances growth and tolerance to dehydration and salt stress in transgenic tobacco. *Plant physiology,* **142**, 113-123.

Smith MK, McComb JA (1983) Selection for NaCl tolerance in cell cultures of

Medicago sativa and recovery of plants from a NaCl-tolerant cell line. *Plant cell reports,* **2**, 126-128.

Song J, Feng G, Tian C, Zhang F (2005) Strategies for adaptation of Suaeda physophora, Haloxylon ammodendron and Haloxylon persicum to a saline environment during seed-germination stage. *Annals of botany,* **96**, 399-405.

Song J, Fan H, Zhao Y, Jia Y, Du X, Wang B (2008) Effect of salinity on germination, seedling emergence, seedling growth and ion accumulation of a euhalophyte *Suaeda salsa* in an intertidal zone and on saline inland. *Aquatic Botany,* **88**, 331-337.

Sreenivasulu N, Harshavardhan VT, Govind G, Seiler C, Kohli A (2012) Contrapuntal role of ABA: does it mediate stress tolerance or plant growth retardation under long-term drought stress? *Gene,* **506**, 265-273.

Sultan SE (1996) Phenotypic plasticity for offspring traits in *Polygonum persicaria*. *Ecology*, **77**, 1791–1807.

Sultan SE, Barton K, Wilczek AM (2009) Contrasting patterns of transgenerational plasticity in ecologically distinct congeners. *Ecology,* **90**, 1831-1839.

Ullah SM, Soja G, Gerzabek MH (1994) Ion uptake, osmoregolation and plant-water relations in faba beans (*Vicia faba* L.) under salt stress. .

Van Zandt PA, Mopper S (2004) The effects of maternal salinity and seed environment on germination and growth in Iris hexagona. *Evolutionary Ecology Research,* **6**, 813-832.

Vinocur B, Altman A (2005) Recent advances in engineering plant tolerance to abiotic stress: achievements and limitations. *Current opinion in biotechnology,* **16**, 123-132.

Wahid A, Rasul E, Rao A-U (1999) Germination of seeds and propagules under salt stress. *Handbook of plant and crop stress,* **2**, 153-167.

Weinig C (2000) Differing selection in alternative competitive environments: shade-

avoidance responses and germination timing. *Evolution; international journal of organic evolution,* **54**, 124-136.

Weinl S, Kudla J (2009) The CBL-CIPK Ca^{2+}-decoding signaling network: function and perspectives. *The New phytologist,* **184**, 517-528.

Wilczek AM, Roe JL, Knapp MC, Cooper MD, Lopez-Gallego C, Martin LJ, Muir CD, Sim S, Walker A, Anderson J, et al (2009) Effects of genetic perturbation on seasonal life history plasticity. *Science,* **323**, 930-934.

Witzel K, Weidner A, Surabhi GK, Varshney RK, Kunze G, Buck-Sorlin GH, Börner A, Mock HP (2010) Comparative analysis of the grain proteome fraction in barley genotypes with contrasting salinity tolerance during germination. *Plant, cell and environment,* **33**, 211-222.

Xiong L, Wang RG, Mao G, Koczan JM (2006) Identification of drought tolerance determinants by genetic analysis of root response to drought stress and abscisic acid. *Plant physiology,* **142**, 1065-1074.

Young ND, Debellé F, Oldroyd GE, Geurts R, Cannon SB, Udvardi MK, Benedito VA, Mayer KF, Gouzy J, Schoof H, et al (2011) The Medicago genome provides insight into the evolution of rhizobial symbioses. *Nature,* **480**, 520-524.

Zahran HH (1999) Rhizobium-legume symbiosis and nitrogen fixation under severe conditions and in an arid climate. *Microbiology and molecular biology reviews,* **63**, 968-989.

Zahaf O, Blanchet S, de Zélicourt A, Alunni B, Plet J, Laffont C, de Lorenzo L, Imbeaud S, Ichanté JL, Diet A, et al (2012) Comparative transcriptomic analysis of salt adaptation in roots of contrasting *Medicago truncatula* genotypes. *Molecular plant,* **5**, 1068-1081.

Zhao Y, Wang T, Zhang W, Li X (2011) SOS3 mediates lateral root development under low salt stress through regulation of auxin redistribution and maxima in

Arabidopsis. *New phytologist,* **189**, 1122-1134.

Zhao R, Sun HL, Mei C, Wang XJ, Yan L, Liu R, Zhang XF, Wang XF, Zhang DP (2011) The Arabidopsis Ca^{2+}-dependent protein kinase CPK12 negatively regulates abscisic acid signaling in seed germination and post-germination growth. *New phytologist,* **192**, 61-73.

Zolla G, Heimer YM, Barak S (2010) Mild salinity stimulates a stress-induced morphogenic response in *Arabidopsis thaliana* roots. *Journal of experimental botany,* **61**, 211-224.

Zhu J-K (2002) Salt and drought stress signal transduction in plants. *Annual review of plant biology,* **53**: 247.

Supplemental material

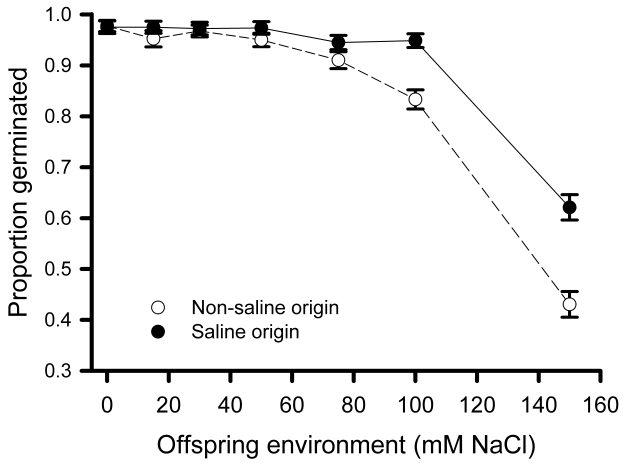


Figure S1. Mean (+/- 1 standard error) germination proportion from the laboratory germination experiment of saline origin (closed circles) and non-saline origin (open circle) genotypes. Seeds from parental plants were grown in saline and non-saline environments and sown in either 0, 15, 30, 50, 75, 100 or 150 mM NaCl.

Table S1. Significance values from sequential Bonferroni tests for all the experiments where multiple comparisons of means were performed.

Trait	Comparison	statistic	P-value
Experiment: Germination study			
Trait: Germination proportion comparisons of means between populations were made within each offspring environment salinity concentration.	**75 mM NaCl OE**	**t-value**	
	TN7 vs. TN1	2.57	**0.0105**
	TN7 vs. TN8	0.38	0.7069
	TN7 vs. TN9	0.56	0.5725
	TN1 vs. TN8	3.43	**0.0006**
	TN1 vs. TN9	3.60	**0.0003**
	TN8 vs. TN9	0.16	0.8727
	100 mM NaCl OE	**t-value**	
	TN7 vs. TN1	5.42	**<0.0001**
	TN7 vs. TN8	3.32	**0.0009**
	TN7 vs. TN9	0.11	0.9140
	TN1 vs. TN8	2.20	**0.0277**
	TN1 vs. TN9	5.36	**<0.0001**
	TN8 vs. TN9	3.26	**0.0012**
	150 mM NaCl OE	**t-value**	
	TN7 vs. TN1	6.43	**<0.0001**
	TN7 vs. TN8	7.49	**<0.0001**
	TN7 vs. TN9	3.34	**0.0009**
	TN1 vs. TN8	1.21	0.2255
	TN1 vs. TN9	3.21	**0.0014**
	TN8 vs. TN9	4.37	**<0.0001**

Trait	Comparison	statistic	P-value
Experiment: Germination study	**0 mM NaCl PE**	**t-value**	
Trait: Germination proportion	TN7 vs. TN1	2.03	0.0423
comparisons of means between	TN7 vs. TN8	3.10	**0.0020**
populations were made within each	TN7 vs. TN9	0.51	0.6135
parental environment salinity	TN1 vs. TN8	1.07	0.2835
concentration.	TN1 vs. TN9	1.56	0.1186
	TN8 vs. TN9	2.64	**0.0084**
	100 mM NaCl PE	**t-value**	
	TN7 vs. TN1	3.66	**0.0003**
	TN7 vs. TN8	0.95	0.3420
	TN7 vs. TN9	0.50	0.6188
	TN1 vs. TN8	2.82	**0.0049**
	TN1 vs. TN9	3.24	**0.0012**
	TN8 vs. TN9	0.46	0.6473
Experiment: Germination study			
Trait: Germination proportion	**Comparison of PE slopes**	**F**	
comparisons of slopes between	TN7 vs. TN1	4.79	0.0287
populations were made within each	TN7 vs. TN8	10.63	**0.0049**
parental environment salinity	TN7 vs. TN9	0.00	0.9558
concentration.	TN1 vs. TN8	21.23	**<0.0001**
	TN1 vs. TN9	16.88	**0.0002**
	TN8 vs. TN9	13.05	**0.0003**
Experiment: Germination study			
Trait: Time to 50% germination	**75 mM NaCl**	**t-value**	
comparisons of means between	TN7 vs. TN1	2.24	0.0251
populations were made within each	TN7 vs. TN8	2.24	0.0255
offspring environment salinity	TN7 vs. TN9	1.43	0.1528
concentration.	TN1 vs. TN8	0.02	0.9806
	TN1 vs. TN9	0.84	0.4020
	TN8 vs. TN9	0.85	0.3944
	100 mM NaCl	**t-value**	
	TN7 vs. TN1	2.37	0.0177
	TN7 vs. TN8	3.89	**0.0001**
	TN7 vs. TN9	3.32	**0.0009**
	TN1 vs. TN8	0.53	0.5942
	TN1 vs. TN9	1.58	0.1151
	TN8 vs. TN9	1.02	0.3074
	150 mM NaCl	**t-value**	
	TN7 vs. TN1	4.91	**<0.0001**
	TN7 vs. TN8	4.87	**<0.0001**
	TN7 vs. TN9	4.22	**<0.0001**
	TN1 vs. TN8	0.62	0.5344
	TN1 vs. TN9	0.18	0.8544
	TN8 vs. TN9	0.79	0.4315

Trait	Comparison	statistic	P-value
Experiment: Caisson Exp.			
Trait: Root length		**F**	
comparisons of slopes between populations were made within each salinity concentration.	NS 0mM vs. NS 75mM	30.24	**<0.0001**
	NS 0mM vs. S 0mM	13.33	**<0.0001**
	NS 0mM vs. S 75mM	89.32	**<0.0001**
	NS 75mM vs. S 0mM	6.82	**0.0102**
	NS 75mM vs. S 75mM	5.76	**0.0179**
	S 0mM vs. S 75mM	35.69	**<0.0001**
Experiment: Caisson Exp.			
Trait: Root-Shoot		**F**	
comparisons of slopes between populations were made within each salinity concentration.	NS 0mM vs. NS 75mM	2.60	0.1096
	NS 0mM vs. S 0mM	0.33	0.5692
	NS 0mM vs. S 75mM	0.07	0.7894
	NS 75mM vs. S 0mM	6.54	0.0119
	NS 75mM vs. S 75mM	2.07	0.1526
	S 0mM vs. S 75mM	0.74	0.3897
Experiment: Caisson Exp.			
Trait: Root biomass		**F**	
comparisons of slopes between populations were made within each salinity concentration.	NS 0mM vs. NS 75mM	11.75	**0.0009**
	NS 0mM vs. S 0mM	1.04	0.3110
	NS 0mM vs. S 75mM	2.61	0.1086
	NS 75mM vs. S 0mM	7.18	**0.0085**
	NS 75mM vs. S 75mM	4.39	0.0383
	S 0mM vs. S 75mM	0.49	0.4862
Experiment: Caisson Exp.		**t-value**	
Trait: Lateral root number	NS 0mM vs. NS 75mM	8.00	**<0.0001**
comparisons of means between populations were made within each salinity concentration.	NS 0mM vs. S 0mM	2.22	0.0288
	NS 0mM vs. S 75mM	4.52	**<0.0001**
	NS 75mM vs. S 0mM	4.25	**<0.0001**
	NS 75mM vs. S 75mM	2.14	0.0347
	S 0mM vs. S 75mM	9.79	**<0.0001**
Experiment: ABA-NaCl experiment			
Trait: Root length		**F**	
comparisons of slopes between populations were made within each treatment.	Control vs. NaCl	5.94	**<.0001**
	Control vs. ABA	6.10	**<.0001**
	Control vs. ABA+NaCl	11.07	**<.0001**
	NaCl vs. ABA	0.09	0.9244
	NaCl vs. ABA+NaCl	5.24	**<.0001**
	ABA vs. ABA+NaCl	5.20	**<.0001**
Experiment: ABA experiment		**F**	
Trait: Root length	0ABA-NSorg vs. 0ABA-Sorg	1.18	0.2403
comparisons of slopes between treatments and soil origin combinations.	0ABA-NSorg vs. 25ABA-NSorg	1.69	0.0926
	0ABA-NSorg vs. 25ABA-Sorg	3.15	**0.0019**
	0ABA-Sorg vs. 25ABA-NSorg	2.77	**0.0060**
	0ABA-Sorg vs. 25ABA-Sorg	4.25	**<.0001**
	25ABA-NSorg vs. 25ABA-Sorg	1.33	0.1841

Trait	Comparison	statistic	P-value
Experiment: NaCl – KCl experiment			
Trait: Root growth rate	**0 mM NaCl**	**t-value**	
comparisons of means between KCl treatment within each NaCl concentration.	0 mM KCl vs. 2 mM KCl	0.42	0.6732
	0 mM KCl vs. 10 mM KCl	1.55	0.1214
	0 mM KCl vs. 50 mM KCl	0.81	0.4210
	2 mM KCl vs. 10 mM KCl	1.13	0.2586
	2 mM KCl vs. 50 mM KCl	0.41	0.6785
	10 mM KCl vs. 50 mM KCl	0.63	0.5270
	2 mM NaCl	**t-value**	
	0 mM KCl vs. 2 mM KCl	0.17	0.8680
	0 mM KCl vs. 10 mM KCl	1.27	0.2039
	0 mM KCl vs. 50 mM KCl	0.89	0.3734
	2 mM KCl vs. 10 mM KCl	1.11	0.2691
	2 mM KCl vs. 50 mM KCl	1.07	0.2859
	10 mM KCl vs. 50 mM KCl	2.25	0.0252
	10 mM NaCl	**t-value**	
	0 mM KCl vs. 2 mM KCl	0.49	0.6256
	0 mM KCl vs. 10 mM KCl	1.00	0.3160
	0 mM KCl vs. 50 mM KCl	0.94	0.3498
	2 mM KCl vs. 10 mM KCl	1.42	0.1574
	2 mM KCl vs. 50 mM KCl	1.36	0.1765
	10 mM KCl vs. 50 mM KCl	0.07	0.9458
	50 mM NaCl	**t-value**	
	0 mM KCl vs. 2 mM KCl	2.52	**0.0126**
	0 mM KCl vs. 10 mM KCl	5.13	**<.0001**
	0 mM KCl vs. 50 mM KCl	3.05	**0.0026**
	2 mM KCl vs. 10 mM KCl	2.66	**0.0084**
	2 mM KCl vs. 50 mM KCl	0.58	0.5626
	10 mM KCl vs. 50 mM KCl	2.04	0.0433
Experiment: NaCl – KCl experiment			
Trait:Primary root tip death	**10 mM NaCl**	**z-value**	
comparisons of means between KCl treatment within the 10 mM and 50 mM NaCl concentrations.	0 mM KCl vs. 2 mM KCl	0.69	0.4887
	0 mM KCl vs. 10 mM KCl	0.49	0.0866
	0 mM KCl vs. 50 mM KCl	1.6	0.1095
	2 mM KCl vs. 10 mM KCl	1.02	0.3099
	2 mM KCl vs. 50 mM KCl	0.89	0.3726
	10 mM KCl vs. 50 mM KCl	0.15	0.8809
	50 mM NaCl	**z-value**	
	0 mM KCl vs. 2 mM KCl	2.01	0.0445
	0 mM KCl vs. 10 mM KCl	3.18	**0.0014**
	0 mM KCl vs. 50 mM KCl	3.57	**0.0004**
	2 mM KCl vs. 10 mM KCl	1.36	0.1753
	2 mM KCl vs. 50 mM KCl	2.61	**0.0091**
	10 mM KCl vs. 50 mM KCl	1.87	0.0621

Chapter 4.

Independent evolution of Portuguese and Tunisian *Medicago truncatula* that occur in environments that contrast in salinity

***This work was submitted for publication as**:*

Cordeiro MA, Moriuchi KS, Chang PL, Friesen ML, Vu WT, Porter S, Scharnagl K, Mukherjee J, Lee C, Kim J-H, Oh M-J, Lim W-J, Fevereiro PS, Strauss SY, Kim N, Mun J-H, vonWettberg EJ, Nuzhdin SV, Choi H-K, Cook DR (2016) Independent evolution of Portuguese and Tunisian *Medicago truncatula* that occur in environments that contrast in salinity. Submitted to *Molecular ecology*.

Matilde Ataide Cordeiro (MAC) was the major contributor for the elaboration of this chapter. MAC conceptualized, designed and performed experiments, analyzed and interpreted the data, and wrote the manuscript. The exception relates mainly with the execution of the libraries, whole-genome sequencing and single nucleotide polymorphism (SNP) calls for the Portuguese *Medicago truncatula* genotypes.

Chapter 4. Table of contents

Abstract

Salinity adaptation is a dynamic and complex process that is influenced by the population's biogeographic, edapho-climatic, genetic and demographic context. An ecological genomics approach was used to test hypotheses of adaptation among geographically isolated populations of the model legume *Medicago truncatula* (Mtr) collected from environments contrasting in salinity in Tunisia and Portugal. Phenotypic and genome analyses were used to infer the mechanisms and the genetic basis of salinity adaptation in Tunisian and Portuguese populations, respectively. Portuguese genotypes had both higher genetic diversity and a more complex population structure relative to Tunisian genotypes, suggestive of higher dispersal rates and adaptive potential in Portuguese populations. A field reciprocal transplant experiment conducted in Portugal showed evidence of salinity adaptation but also revealed that salinity is not the only factor driving performance at the Portuguese planting sites. Trait selection in Portuguese and Tunisian populations across saline and non-saline habitats indicates that common developmental and physiological processes are critical under salinity. But the geographic separation and genetic differentiation suggest that Tunisian and Portuguese populations acquired salinity adaptation independently. Specifically, parallel phenotypic divergence is apparent for tolerance related traits with common developmental and physiological processes targeted by selection in Tunisia and Portugal, while a key avoidance related trait for Tunisian populations is decoupled from salinity in Portuguese populations. Furthermore, genomic data indicate that distinct genome regions were under selection in Tunisian and Portuguese saline populations, though some target genes are predicted to impact common pathways.

Introduction

Divergent selection across habitats can result in multiple, non-exclusive, adaptive responses that increase individual fitness in local environments. Local adaptation is one possible response and is a special case of a genotype by environment interaction (Kawecki & Ebert 2004). Locally adapted genotypes have higher fitness in their home environment relative to genotypes from other environments (e.g., Turesson 1922; Clausen *et al.* 1940; Busoms *et al.* 2015). Although examples of local adaptation are common, not all populations evolve to have relatively higher performance at their home environment (e.g., Galloway & Fenster 2000; Hansen *et al.* 2006; Hereford & Winn 2008; reviews Kawecki & Ebert 2004; Hereford 2009).

For local adaptation to evolve selection must favor different traits in populations inhabiting different environments (i.e., divergent selection) and relevant genetic variation must exist within populations (Reviewed by Kawecki & Ebert 2004; Hereford 2009; Sanford & Kelly 2011). Local adaptation can be inferred when the reaction norms for fitness measures among populations cross in the comparison between home and away environments, however such patterns are not always detected nor expected (Fry 1996; Kawecki & Ebert 2004; Hereford 2009). Populations may still display local adaptation without crossing reaction norms, as might occur when alleles that confer fitness benefits in the home environment are neutral in the foreign environment. Additionally, biogeographic and demographic factors can constrain population-level responses and confound interpretation, for example when gene flow among populations has a homogenizing effect on alleles and outpaces natural selection (Lenormand 2002), or when alleles are driven towards fixation independent of selection as can occur for populations of small sizes (Postma & van Noordwijk 2005).

Deciphering the genetic bases of adaptation represents a current challenge in ecological genomics. In the simplest case, local adaptation involves major genes with strong contributions to fitness and the candidate genes inform us about ecologically relevant mechanisms (MacNair 1983; Courbot *et al.* 2007; Baxter *et al.*

2010; Huang *et al.* 2010; Friesen *et al.* 2014). Such situations are amenable to direct genetic analysis using the logic of selective sweeps or quantitative trait loci (QTL) analyses. More commonly, evolution involves multiple genes of small effect, each influencing distinct traits that make incremental contributions to fitness (Pritchard & Di Rienszo 2010; Le Corre & Kremer 2012; Gould *et al.* 2014). Moreover, depending on the frequency and distribution of causal alleles within a population, different individuals may contain different combinations of alleles and thus possess different adaptive capacities. This combination of genetic heterogeneity and incremental effects is intractable to standard genetic tests. However, the recent advent of population-level genomics to identify genome variation, combined with computational analysis to correlate such variation with traits and environments, provides increasing power to nominate candidate genes and genomic intervals associated with adaptive phenotypes (Galloway & Fenster 2000; Hall *et al.* 2010; Huang *et al.* 2010; Turner *et al*. 2010; Fournier-Level *et al.* 2011; Ingvarssom & Street 2011).

Here, we test for adaptation to salinity using the model legume *Medicago truncatula* var. *truncatula* (Mtr) (Cook 1999; Young *et al.* 2011). Mtr is a predominantly selfing annual legume native to the Mediterranean Basin where it occurs in a range of environments including saline and non-saline habitats (Ronfort *et al*. 2006; Lazrek *et al*. 2009; Badri *et al*. 2007; Friesen *et al*. 2010; Friesen *et al.* 2014). Plants have evolved distinct strategies to cope with salinity, including tolerance, avoidance and exclusion mechanisms (Flowers *et al.* 1977; Munns & Tester 2008). Glycophytes such as Mtr and most other legumes can occur in both saline and non-saline environments, with the corollary that they often possess standing variation for adaptations to evolve in their particular habitats. The evolution of salinity tolerance in glycophyte populations depends on multiple factors, including soil salinity levels, related soil characteristics (i.e., soil water content, nutrient levels, cation exchange capacity), seed dispersal rates and distances (i.e., gene flow), life cycle stages sensitive to salinity, and the amount of genetic variation within populations. Irrespective of such complexity, a common

expectation of saline-adapted populations is that they will survive and reproduce to a greater extent under higher salt conditions than populations that evolved in non-saline environments (Lowry *et al.* 2008, 2009; Munns & Tester 2008; Friesen *et al.* 2014; Busoms *et al.* 2015).

Previous studies using Tunisian genotypes of Mtr demonstrate that populations from saline and non-saline habitats are genetically differentiated by relatively small genome regions (e.g., Friesen *et al.* 2010), while greenhouse experiments indicate that the same populations are adapted to their home soil salinity levels (e.g., Arraouadi *et al.* 2012; Castro *et al.* 2013; Cordeiro *et al.* 2014; Friesen *et al.* 2014). In particular, salt adapted genotypes from northern Tunisia populations avoid salinity stress by germinating and flowering earlier relative to non-salt adapted genotypes, when soil salinity is transiently low in rainy Mediterranean winters. Moreover, the genomes of saline populations display hallmarks of physiological tolerance, being specifically enriched in alleles of candidate abiotic stress regulatory genes, including calcium-dependent protein kinase (CPK) and CBL interacting protein kinase (CIPK) paralogs, and genes implicated in stress hormone metabolism (i.e., abscisic acid; Friesen *et al.* 2014).

Here we take an ecological genomics approach to test and compare salinity adaptations in two geographically isolated populations of Mtr derived from northern Tunisia and southern Portugal, each represented by a set of local meta-populations whose local environments contrast in salinity levels. In the case of Tunisia, these populations are linked by moderate gene flow (Friesen *et al*., 2014). Whole genome re-sequencing and population genetics were used to address questions of genetic differentiation and selection history, while a field reciprocal transplant experiment conducted at the original Portuguese collection sites was used to quantify the environmental dependence of fitness and its associated traits. We test: i) whether Tunisian populations that have evolved salinity adaptation in Tunisia are also adapted when grown in Portuguese saline environments; ii) whether Portuguese populations exhibit local and/or salinity-associated adaptation; and iii) if we find parallel patterns of phenotypic divergence in Tunisian and

Portuguese populations, i.e., if phenotypic responses and genomic data suggest common or distinct mechanisms of salinity adaptation. Towards these ends, fitness components and traits associated with plant growth were quantified to identify the stage of the life cycle that is most negatively influenced by soil salinity, and traits associated with salinity tolerance and avoidance were quantified to identify the potential mechanism of salinity adaptation.

Materials and methods

Original collection sites and Portuguese transplant locations

Medicago truncatula (Mtr) populations were collected at two locations in southern Portugal (PT) in June of 2010: Castro Marim (CM_{loc}) and Gilberto (Gil_{loc}) separated by ~35 km. At each location, two field sites were selected based on the presence or absence of soil salinity totaling four sites: CM saline (CMS) and CM non-saline (CMNS) at CM_{loc}; and Gil saline (GilS) and Gil non-saline (GilNS) at Gil_{loc} (Fig.1A). *In situ* measurements of salinity via electro conductivity (EC_{field}) were recorded at each field site when seeds were initially collected (May 2010) and again during the course of the field reciprocal transplant experiment (January to June 2011; Slavich & Petterson 1993). At each site, three to four 0.5 kg soil samples were collected from Mtr root zones and analyzed for a range of physical-chemical properties of soil at the A&L Western Agricultural Laboratories (Modesto, California) including several soil salinity parameters, and organic and inorganic nutrients (see Supplemental Material). To quantify seasonal variation in soil water content, additional soil samples were collected during late March (3/25/2011), mid-April (4/17/2011), and at time of harvest (CM_{loc}: 5/25/2011; Gil_{loc}: 6/2/2011). Soil samples were weighed to the nearest 0.1 mg, dried and subsequently reweighed. Soil water content was calculated as: Soil Water Content =*100*×[(*fresh weight – dry weight) / dry weight*].

Genotypes from widely studied Mtr populations with origin in saline and non-saline soils in northern Tunisia (saline: TN1 and TN8; non-saline: TN7 and

TN9; Lazrek *et al.* 2009) were selected for having contrasting fitness under saline conditions (highest and lowest reproduction within each population; Table S1, Supplemental Material). Previous studies have shown that TN populations are adapted to their local soil salinity levels (Friesen *et al.* 2014). The main mechanism involved in adaptation of saline origin populations is avoidance by expressing increased growth rates and earlier flowering to escape salt accumulation in the soil as water becomes scarce, but some level of tolerance has also been detected (Castro *et al.* 2013; Cordeiro *et al.* 2014; Friesen *et al.* 2014; Moriuchi *et al.* 2016).

NGS-based re-sequencing and analysis of population structure

Whole genome re-sequencing was conducted for a total of 39 Portuguese genotypes (Table S1, Supplemental Material). Barcoded libraries, composed of 500bp insert fragments, were constructed and sequenced in multiplex across 14 lanes of an Illumina HiSeq 1000 using 101bp paired end reads. The resulting reads were mapped to the Mtr v4.0 reference (Tang *et al.* 2014) using BWA algorithms (Li & Durbin 2010), allowing 8 mismatches per read. Polymorphisms were called using the GATK pipeline (McKenna *et al.* 2010), which considers duplicate removal, indel realignment, and base quality score recalibration, while simultaneously calling variants through the HaplotypeCaller program. Variants were filtered using standard hard filtering parameters according to GATK Best Practices recommendation (DePristo *et al.* 2011; van der Auwera *et al.* 2013).

Genetic structure of Portuguese populations was deduced with STRUCTURE (Pritchard *et al.* 2000) using a subset of 185,446 loci that were called in all 39 genotypes. Linkage effects were minimized by requiring that all loci were separated by at least 1,000 bases. STRUCTURE was run assuming admixture and using correlated allele frequencies. Ten independent runs of 10,000 burn-in MCMC iterations followed by 50,000 iterations were performed for two to nine subdivisions (K=2 to K=9). STRUCTURE HARVESTER was used to calculate optimal K using the mean log probability (LnP[D], Evanno *et al.* 2005). Neighbor joining phylogenetic trees were built using DARwin (Perrier & Jacquemoud-Collet

2006). AMOVA, F-statistics and pairwise differentiation analyses were conducted using GenoDive 2.0 (Meirmans & van Tienderen 2004) and VCFtools (Petr *et al.* 2011). Comparisons involving both Portuguese and Tunisian genotypes (Friesen *et al.* 2014) used a set of 28,255 SNPs for which the loci were called in all Portuguese and Tunisian genotypes.

Seed Germination and Reciprocal Transplantation

At each of the four Portuguese planting sites, two transects were established following the distribution of naturally occurring Mtr plants. During mid-January 2011, a minimum of three replicates from each of the 60 genotypes (Table S1, Supplemental Material) were used for each site; a total of 1,526 seed (348 for each CM_{loc} site and 420 for each Gil_{loc} site) were scarified, randomized along trays for each transect, and germinated in ~15 cc of soil collected from both transects at each planting site (Table S1, Supplemental Material). Trays were stored outside in a protected area and were watered daily to soil saturation. Seeds were considered germinated when cotyledons were observed. After two weeks, plugs of the trays with no germinant were transplanted with seedlings that were forced to germinate under non-saline conditions. One week later seedlings were transplanted into the field sites. Seedlings were randomly transplanted into each transect at least 10 cm apart, were watered daily for the first week to minimize transplant shock, every other day for the second week, and whenever dried after that. Seedlings that died within two weeks of transplanting were considered dead due to transplant shock and excluded from subsequent analyses.

From the start of flowering, plants were censused weekly and mature pods were collected. The number of leaves at first flower was recorded as a proxy for size at reproduction. At time of harvest (CM_{loc}: 5/25/2011; Gil_{loc}: 6/2/2011) all plants, including full root systems, were collected. For transplant sites with no visible plants, soil was inspected to confirm plant death. Soil was washed from roots and the number of nodules was counted. Fresh weight of above ground and root tissue were weighed to the nearest 0.01mg and reweighed after plants had

dried to a constant weight. Tissue water content was calculated as: *100*×[(*fresh weight – dry weight) / dry weight*].

Quantification of leaf ion content

To correlate Na^+ and other ions (i.e., K, Ca, Mg, Mn, Zn and P) with plant growth and performance, ion content was determined for leaves collected at harvest using a subset of Portuguese genotypes (Table S2, Supplemental Material) according to Lahner *et al.* (2003). Briefly, three dried leaves from each plant were ground using a Qiagen TissueLyser with tungsten carbide balls and weighed on a CAHN microbalance. Among successively analyzed samples, the 16^{th} and 17^{th} samples were controls: one blank and one NIST (National Institute of Standards and Technology, SMR 1570a trace elements in spinach leaves), respectively. Ion content was analyzed using an Agilent 7500CE ICP-MS (Agilent Technologies, Palo Alto, CA) at the Interdisciplinary Center for Plasma Mass Spectrometry at UC Davis (ICPMS.UCDavis.edu) for Na, K, Ca, Mg, Mn, Zn and P. Tunisian genotypes were analyzed for Na^+ and K^+ content using greenhouse grown plants treated with 0mM and 100mM NaCl (from greenhouse experiments in Friesen *et al.* 2014) and analyzed using an Agilent 7500CE ICP-MS (Agilent Technologies, Palo Alto, CA) at the USDA ARS at Chapman Field, Coral Gables, FL (see Supplemental Material).

Data analyses

Plant performance and adaptation

Fitness was estimated as the product of germination (0, 1), survival to reproduction or to the end of study (0, 1), and number of pods (Conner 1996). The probability of a genotype increasing in the seed bank over time (here called "genotype growth rate" but also represented by "λ"; Molles & Cahill 1999) was estimated using the 'Matrix' and 'lattice' R packages (R Development Core Team, 2008) based on fitness values calculated from field phenotypes, combined with the observation that each pod contains on average six seed and the assumption that

1-2 seeds germinate from a pod in a given season, which agrees with our previous field observations (Friesen *et al.* 2014). Genotype growth rate is an alternative estimate of fitness, representing the means for each genotype in each transect and thus minimizing micro-spatial variation within a planting site.

Mixed-model ANOVAs were used to test whether populations are adapted to their home soil salinity levels, and whether populations from different countries differ in their response to soil salinity. Mixed-model generalized linear models (PROC GLIMMIX for binary response variables and PROC MIXED for continuous response variables) were used to analyze germination, survival to reproduction, number of pods, fitness, total biomass, and genotype growth rate. All four continuous variables were natural logarithm transformed to meet assumptions of ANOVA. For these traits, planting sites and their interactions (i.e., country of origin [TN and PT], population nested within country of origin [PT: CMNS, CMS, GilNS and GilS; TN: TN1, TN7, TN8, TN9], planting location [CM_{loc} and Gil_{loc}], planting soil type [NS and S] nested within location) were treated as fixed effects, while genotype nested within country of origin and population was treated as a random effect. Transect was included as a covariate to remove spatial variation within each planting site in all analyses except germination, the data for which was collected *ex situ*. Least-square means comparisons (LSMEANS) were performed to test significance in differences between factors and were used to test hypotheses of local adaptation.

Genetic and environmental influences on morphology

Interactions between population and country of origin responses to planting sites were tested using mixed-model ANOVAs (PROC MIXED) performed on five morphological traits, seven ions and two ion ratios: i) age at first flowering, number of leaves at first flower, root and shoot water content, and number of root nodules; ii) Na, K, Ca, Mg, Mn, Zn and P (Table S2, Supplemental Material); and iii) Na/K and Ca/Mg, respectively. Effects were treated as described above. Data were either square root transformed (root and shoot water content, and Mg), natural

logarithm transformed (age at first flowering, number of leaves at first flowering, number of nodules, Na, Ca, Mn and Zn), or not transformed (K, P, Na/K and Ca/Mg) to meet the assumptions of normality and homoscedasticity of the ANOVA. Least-square means comparisons were performed for significant terms in the model to test for differences in trait values between factors.

Relationship between traits and performance

To quantify the relationship between traits and performance, phenotypic and genotypic selection analyses, which estimate both total selection (s) and direct selection (β), were performed separately for each trait and each planting site. Direct and indirect selection were partitioned using selection gradient analyses (Lande & Arnold 1983; Mitchell-Olds & Shaw 1987). Prior to selection analyses, traits were standardized to have a mean and a standard deviation of one, and performance was relativized to the mean performance within each planting site (Lande & Arnold 1983). Phenotypic and genotypic selection analyses were conducted using ANCOVAs (PROC GLM), with population nested within country of origin and transect as fixed-effects, and the standardized trait as a continuous effect on relative performance (Donohue et al. 2000).

In order to test whether estimated selection differentials and gradients differed between planting sites, ANCOVA's were performed including the planting site by trait interaction. When this interaction term was significant, pairwise comparisons were performed between each planting site to identify which planting sites differed. All analyses, unless otherwise stated, were conducted in SAS version 9.3 (SAS Inc. 2011).

Results

Planting sites are differentiated by soil characteristics, including salinity

Soil from all planting sites was characterized for nutrient composition, electro-conductivity (EC), and water content (Tables S3-6 and Fig. S1,

Supplemental Material). Consistent with expectations, saline (S) sites (CMS and GilS) showed at least 4.5 times higher salinity content than non-saline (NS) sites (CMNS and GilNS), differing with respect to all major indices of soil salinity, i.e., sodium absorption ratio [SAR], exchangeable sodium percentage [ESP], soil sodium ion content and EC values. Phosphorous, calcium and zinc concentrations and pH co-varied with soil salinity, with saline sites having lower concentrations of calcium and greater concentrations of phosphorous and zinc, and elevated pH compared to non saline sites. Other soil nutrients differed among planting sites and were not correlated with salinity. Of particular note, GilNS had unusually low P and high Ca-to-Mg ratio, while the GilS had unusually high K, and therefore balanced Na-to-K ratio (Tables S3-5, Supplemental Material; Peverill et al. 1999).

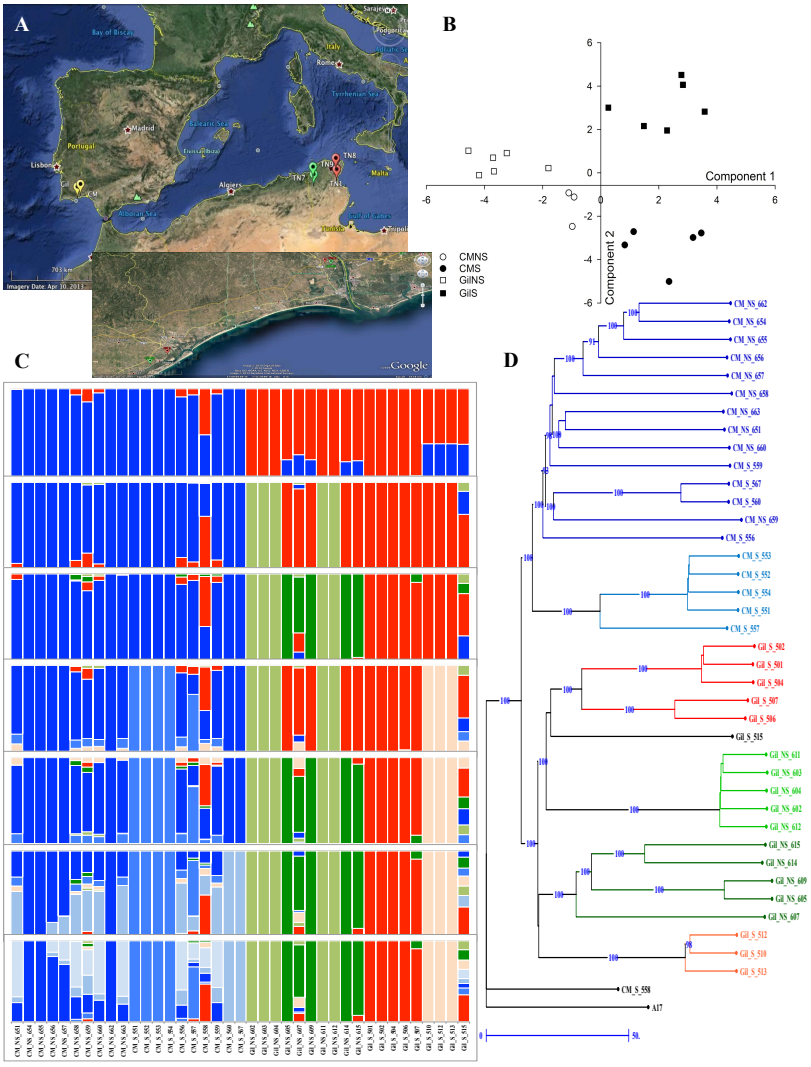


Figure 1. Distribution, environment and genetic relationships among the 39 sequenced Portuguese genotypes. A. Map of the origin of the study populations in Southern Portugal and Northern Tunisia; Portuguese Gil and CM locations as squares and circles, respectively; Tunisian TN1 and TN7 as triangles, TN8 and TN9 as diamonds; non-saline populations in green, saline populations in red. B. Two vector representation of the principal coordinate analysis (PCA) using soil analysis from the four Portuguese saline and non-saline site (CMS, CMNS, GilS, GilNS), which represent 67.27% of the variation of soil characteristics: coordinate 1 represents 35.41% of the variation and coordinate 2 represents 31.87% of the variation. C. Allele-frequency based population assignment using STRUCTURE (Pritchard et al., 2000) assuming subgroups from K=2 to K=8. D. Neighbor joining tree together with the reference genotype A17: Genotypes are color coded based on the K=6 STRUCTURE assignment (Table S1, Supplemental Material); dark blue: CMNS; light blue: CMS; light green: GilNSA; dark green: GilNSB; dark red: GilSA; light red: GilSB; black: unassigned genotypes.

Figure 1 depicts these trends in a principal component (PC) analysis, in which saline and non saline sites readily resolve (Fig. 1B). PC1 explains 35.4% of the variation in soil nutrients among sites and is positively correlated with all sodium related measurements. PC2 further resolves planting sites by location (Gil_{loc} from CM_{loc}), explaining 31.9% of the variation in soil characteristics. Cation exchange capacity (cmol kg^{-1}), pH, potassium, and iron are positively correlated with PC2, while soluble calcium (meq/L) and EC are negatively correlated with PC2 (Tables S3-5, Supplemental Material).

Genetic differentiation of Portuguese and Tunisian populations

With the goal of deducing population genomic features within and among sites, 39 Portuguese Mtr accessions were sequenced to an average depth of 26X. Among these accessions we identified 964,183 SNPs, substantially higher than ~28K identified in the 39 Tunisian genotypes using the same SNP calling criteria. To minimize bias due to differences in coverage depth (Tunisian genotypes have an average of 8X coverage) we analyzed a subset of 28,255 collection-wide SNPs and again observed significantly higher genetic diversity in the Portuguese populations compared to Tunisian populations, with nucleotide diversity (π) among Portuguese genotypes twice that observed within Tunisian genotypes ($6.8x10^{-05}$ and $3.4x10^{-05}$, respectively; $F = 1655$, $P \leq 0.0001$). Multi-dimensional PCA readily resolved genotypes based on country of origin and further into component groups (Figs. S2-4, Supplemental Material). Most variation was among genotypes (F_{IS} higher than 40%; Table 1), while country of origin (F_{RT} of 27.4%; Table 1) explains the majority of genetic differentiation (F_{ST} between [0.392 to 0.564] vs among [0.184 to 0.366 and 0.188 to 0.267] Portuguese and Tunisian origin populations; Table 2). LD was significantly extended among Portuguese compared to Tunisian populations (Portuguese $r^2<0.3$ at ~65kb, this work; Tunisian $r^2<0.3$ at ~10kb [Friesen et al. 2014]), both of which differ from range-wide estimates of Branca et al. (2011; $r^2<0.3$ at ~5kb). These observations are consistent with a more recent

impact of demographic factors in Portugal, such as migration-related bottlenecks or population structure, which may constrain the potential for local adaptation.

Table 1: Results from hierarchical analysis of molecular variance (AMOVA). Proportion of variation explained by among region (FRT), population (FST) or individual (FIS), or by within individual (FIT) genetic variation. Portuguese (PT) and Tunisian (TN) genotypes are grouped by site of origin. PT genotypes are grouped by the most likely allele-frequency based genotype clustering using STRUCTURE (K=6).

Source	Site of origin based (PT and TN)			STRUCTURE based population (PT)		
	% Tot. Var.	F-value	Std Dev	% Tot. Var.	F-value	Std Dev
Among Regions (F_{RT})	27.4%	0.274*	0.002	10.3%	0.103	0.003
Among Populations (F_{ST})	17.6%	0.242***	0.001	33.4%	0.373***	0.002
Among Individuals (F_{IS})	43.4%	0.789***	0.002	43.7%	0.776***	0.004
Within Individuals (F_{IT})	11.6%	0.884	0.001	12.6%	0.874	0.002

t $P < 0.10$, * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, **** $P < 0.0001$

Among Portuguese genotypes, STRUCTURE analysis (Pritchard et al. 2000; Falush et al. 2003) revealed genetic relationships reflecting site of origin, though not exclusively (Figs. 1C). Portuguese genotypes are best described as six genetic groups (LnP[D], Evanno et al. 2005), with apparent admixture. Gil_{loc} groups were more diverse than CM_{loc} groups, and at both GilS and GilNS we observed two distinct populations within each site (GilSA/GilSB and GilNSA/GilNSB at K=6, respectively; Figs. 1C, D). GilNSA was the most homogenous and most differentiated from all other genetic groups (Table 3). In addition to multiple groups occurring within individual sites (i.e., Gil_{loc}), we observed cases of individuals from the same genetic group at both saline and non-saline sites within a location (i.e., CM_{loc}). These patterns of genetic structure nested within locations and sites of origin were also revealed by neighbor joining analysis (Fig. 1D).

Of the approximately 1 million variants among Portuguese accessions, 136,331 (14.1%) were found in coding regions, 59,823 of which were non-synonymous. About 7.4% (10,105) of these coding SNPs represent alleles that are private to one of the six genetically defined groups, with the greatest contribution from GilNSA at 4,057 private alleles. By contrast only 852 coding SNPs were specific to individual sites, reflecting the preponderance of differentiation among genetic groups rather than physical sites. The 53 non-synonymous variants that

assort with a minimum of 80% enrichment between saline and non-saline genetic groups represent 38 genes distributed among 23 haplotype blocks and include several candidate genes implicated in salinity and/or soil-type adaptation in Arabidopsis (Table S7, Supplemental Material; see Discussion). We also identified 130 non-synonymous mutations assorting between CM and Gil locations (Table S8, Supplemental Material), representing 84 genes within 36 haplotype blocks. Many of these genes have functions consistent with patterns of location-specific trait selection, as determined below. Although not accounted for in our phenotypic assays, comparison of allele enrichment among locations revealed significant enrichment for NB-ARC genes likely to function in disease resistance, conspicuously absent from salinity comparisons, suggesting geographic structuring of disease pressure (see Discussion).

Table 2: Results from population pairwise differentiation (F_{ST}) and respective significance (*P* value) for Portuguese (CM_NS, CM_S, Gil_NS and Gil_S) and Tunisian (TN1, TN7, TN8 and TN9) genotypes grouped by site of origin.

	CM_NS	CM_S	Gil_NS	Gil_S	TN1	TN7	TN8	TN9
CM_NS	--	*P*=0.001	*P*=0.001	*P*=0.001	*P*=0.001	*P*=0.001	*P*=0.001	*P*=0.001
CM_S	0.184	--	*P*=0.001	*P*=0.001	*P*=0.001	*P*=0.001	*P*=0.001	*P*=0.001
Gil_NS	0.366	0.323	--	*P*=0.001	*P*=0.001	*P*=0.001	*P*=0.001	*P*=0.001
Gil_S	0.333	0.275	0.251	--	*P*=0.001	*P*=0.001	*P*=0.001	*P*=0.001
TN1	0.525	0.500	0.457	0.483	--	*P*=0.001	*P*=0.001	*P*=0.001
TN7	0.523	0.496	0.454	0.479	0.266	--	*P*=0.001	*P*=0.001
TN8	0.564	0.537	0.495	0.524	0.267	0.268	--	*P*=0.001
TN9	0.452	0.426	0.392	0.412	0.247	0.188	0.252	--

Table 3: Results from population pairwise differentiation (F_{ST}) and respective significance (*P* value) for Portuguese genotypes grouped by the most likely allele-frequency based genotype clustering using STRUCTURE (K=6).

	CM_NS	CM_S	Gil_NSA	Gil_NSB	Gil_SA	Gil_SB
CM_NS	--	*P*=0.001	*P*=0.001	*P*=0.001	*P*=0.001	*P*=0.003
CM_S	0.184	--	*P*=0.002	*P*=0.001	*P*=0.001	*P*=0.004
Gil_NSA	0.606	0.555	--	*P*=0.010	*P*=0.005	*P*=0.013
Gil_NSB	0.449	0.409	0.673	--	*P*=0.005	*P*=0.023
Gil_SA	0.389	0.320	0.621	0.394	--	*P*=0.081
Gil_SB	0.363	0.316	0.734	0.376	0.242	--

Planting site and country of origin impact plant performance

Reciprocal transplant experiments were used to test for adaptation to saline habitats. Fifty-six Mtr genotypes were grown at two locations in Portugal (CM_{loc} and Gil_{loc}) each of which consists of saline (CMS and GilS) and non-saline (CMNS and GilNS) sites. The 38 Portuguese genotypes constitute true reciprocal transplants,

while other 18 accessions originated from northern Tunisia (Lazrek et al. 2009; Friesen et al. 2014).

Table 4. Results from mixed-model generalized linear models on components of plant performance (germination, survival to reproduction, number of pods) and cumulative estimates of performance (fitness, plant biomass, genotype growth rate). Country of origin (Origin); Population nested within country of origin (Pop), Planting location (PltLoc), Planting site (PltSite), Transect nested within planting population and planting site, and genotype nested within country of origin and population (Genotype).

Source	df	Germin. X^2	Survival X^2	Pods F-value	Biomass F-value	Fitness F-value	Growth-λ F-value
Origin	1	8.76**	4.37*	0.02	3.54t	2.83t	0.65
Pop	6	7.79	15.48*	1.51	4.37**	0.49	1.03
PltLoc	1	1.94	10.78**	2.44	22.23****	2.53	15.71****
PltSite	2	15.83***	30.00****	342.63****	536.24****	170.10****	113.65****
Origin X PltLoc	1	0.14	0.00	4.11*	0.27	4.27*	7.89**
Origin X PltSite	2	1.12	7.23*	2.81t	0.20	6.06**	3.09*
Pop X PltLoc	6	5.71	4.93	1.81t	2.51*	2.73*	2.42*
Pop X PltSite	12	13.68	14.53	1.35	1.69t	1.29	1.32
Transect	4	NA	18.36**	2.36t	4.26**	2.13t	3.84**
Genotype	1	2.66***	2.44**	0.00	4.5*	12.7***	8.6**
Gen X PltSite	1	5.30*	0.00	0.00	1.60	1.40	0.00

t $P < 0.10$, * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, **** $P < 0.0001$

Components of fitness – germination, survival to reproduction and number of pods – were measured and analyzed individually. Plant biomass and composite measures (i.e., fitness and genotype growth rate, see Materials and Methods) were treated as performance traits. Planting sites were differentiated by all plant responses (Tables 4, S9, Supplemental Material), but only germination and survival to reproduction were negatively related with soil salinity (germination – NS: 86.4%, S: 79.0%, Fig. 2A; survival to reproduction – NS: 82.4%, S: 73.3%, Fig. 2B). By contrast, traits associated with later stages of the life cycle exhibited planting site- and/or location-specific responses (Tables 4, S9, Supplemental Material). In particular, among CM_{loc} transects, number of pods, biomass, fitness and genotype growth rate were at least 2.4 times greater at CMNS (6.5 pods/plant, 603.1 mg/plant, 4.17 pods/planted seed, 5.7 plants/seed) compared to CMS (1.3 pods/plant, 174.1 mg/plant; 0.85 pods/planted seed, 2.4 plants/seed; Table 4, Figs. 2C,D,E). In contrast, we observed the opposite pattern at Gil_{loc} in the same traits, with performance significantly enhanced at the saline compared to non-saline site. Plants at GilS had at least 8.5 times greater number of pods, biomass, and genotype growth rate (9.4 pods/plant, 734.4 mg/plant, 5.1pods/planted seed, 7.7

plants/seed) compared to plants at GilNS (1.1 pods/plant, 76.3 mg/plant, 0.91 pods/planted seed, 2.9 plants/seed Table 4, Figs. 2C,D,E), despite the facts that survival and germination were significantly lower at GilS sites.

Country of origin, or interactions with country of origin, influenced all measures of performance with the exception of plant biomass (Table 4). Genotypes from Tunisia had 10.3% greater germination (Tunisia: 88.5%, Portugal: 80.2%; χ^2 = 9.26, P = 0.0028; Fig. 2A) relative to genotypes from Portugal (Table S9, Supplemental Material). Greater survival to reproduction of Tunisian compared with Portuguese genotypes was due to the increased survival of Tunisian genotypes at non saline planting sites (χ^2 = 12.91, P = 0.0016; Fig. 2B). However reproduction per se (i.e., number of pods) differed between Portuguese and Tunisian genotypes depending on planting location (Tables 4, S9, Supplemental Material; Fig. 2D). Both composite measures of performance (i.e., fitness and genotype growth rate) were influenced by an interaction between country of origin and planting site (Table 4; Figs. 2C, E). Specifically, Tunisian genotypes had greater fitness relative to Portuguese genotypes at the most productive sites, i.e., CMNS (t = 1.93, P = 0.0533) and GilS (fitness: t = 3.47, P = 0.0005; genotype growth rate: t = 1.93, P = 0.0540). In contrast, Portuguese genotypes had greater genotype growth rate at the most stressful sites (i.e., CMS: t = 2.29, P = 0.0227; and GilNS: t = 1.69; P = 0.0913).

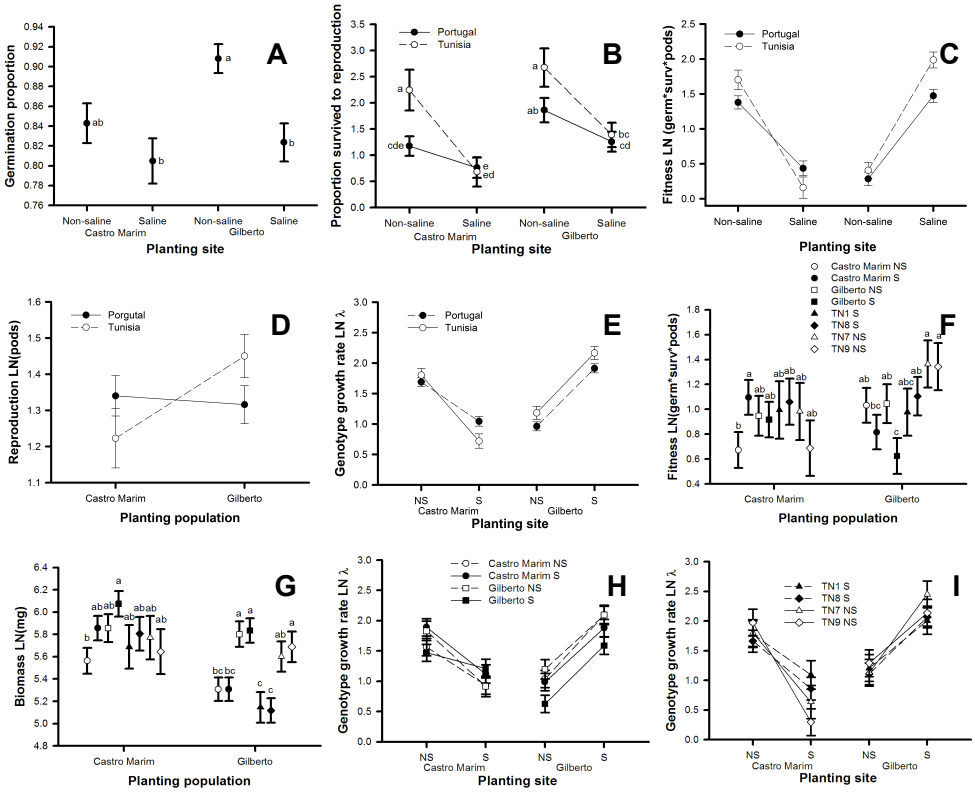


Figure 2: Least square means and standard error of main effects on different measures and components of plant performance. A. Proportion of seeds germinated by planting site. B. Proportion of Portuguese and Tunisian origin genotypes that survived to reproduction by planting site. C. Natural logarithm transformed number of pods of Portuguese and Tunisian origin genotypes at each planting location. D. Natural logarithm transformed fitness measured as the product of germination, survival, and total number of pods at the end of the experiment of Portuguese and Tunisian origin genotypes at each planting site. E. Natural logarithm transformed genotype growth rate (*l*) measured as the probability of a seed germinating, surviving to reproduction and remaining in the seed bank of Portuguese and Tunisian origin genotypes at each planting site. F. Natural logarithm transformed fitness of Portuguese and Tunisian populations based on site of origin at each planting location. G. Natural logarithm transformed total biomass of Portuguese and Tunisian populations based on site of origin at each planting location. H. Natural logarithm transformed genotype growth rate (*l*) of Portuguese populations based on site of origin at each planting site. I. Natural logarithm transformed genotype growth rate (*l*) of Tunisian populations at each planting site.

Patterns of performance suggest aspects of adaptation, mal-adaptation and the impact of demography

Under an hypothesis of local adaptation, significant population by planting site interactions are expected for at least one measure of performance. Biomass

had a marginally significant relationship (Table 4), and survival to reproduction showed significant signal when looking at the Portuguese origin populations only (Table S9, Supplemental Material), but no home site advantage was detected.

All cumulative measures of performance show overall signal for site of origin (i.e., population) by planting location interactions (i.e., marginally significant for number of pods, and significant for fitness, biomass and genotype growth rate; Figs. 2F, G; Table 4). Interestingly, the directionality of these responses at Gil_{loc} is opposite to the hypothesis of salinity adaptation (Figs 2F-I). Among Gil_{loc} genotypes, GilNS genotypes significantly outperformed GilS genotypes in both fitness (Fig. 2F) and genotype growth rate (Fig. 3H), irrespective of planting site (Table 4 and Fig. 2H). Similar and significant differences were observed for biomass among Tunisian genotypes at Gil_{loc}, with non saline-origin genotypes outperforming saline-origin genotypes (Fig. 2G). Taken together, these observations suggest that factors other than salinity per se are operating at GilS and that salinity adaptation is generally maladaptive at Gil_{loc}.

Nevertheless, signal for salinity adaptation was detected. Linear contrasts on genotype growth rate showed that saline-origin genotypes had better performance than non saline-origin genotypes at both saline planting sites (CMS: $F_{(1,48)} = 4.59$; P=0.0373; GilS: $F_{(1,48)} = 4.51$; P = 0.0389; Table 5). Considering only the planting sites at CM_{loc}, saline-origin consistently outperformed individuals of non saline-origin when transplanted into CMS (Fig. 2H, I), with 27.2% higher genotype growth rate. This outcome was independent of whether genotypes originated from saline sites in Portugal or Tunisia. In contrast, at CMNS, genotype growth rates of saline-origin and non saline-origin accessions were similar. Taken together, these observations are consistent with salinity adaptations in saline-origin genotypes, and that these adaptations confer increased genotype growth rate at CMS but are of non-measurable consequence at CMNS.

Patterns of selection and trait expression

Phenotypic and genotypic selection analyses were implemented to quantify the relationships between traits and performance. Towards this end, and given our focus on salinity, we quantified a set of non-correlated traits commonly associated with salinity avoidance (age and size at first flowering) and tolerance (tissue water content and tissue ion content, Table 6). Nodule number was also included in the analyses due to the importance of nitrogen fixation to legume performance and the fact that salinity can negatively impact symbiotic nitrogen fixation (de Lorenzo et al. 2007, Table 6).

Table 5. Results of linear contrasts on genotype growth rate comparing saline and non-saline populations from Portugal (PT) and Tunisia (TN) within each planting site.

Contrast	CM NS	CM S	Gil NS	Gil S
Original Groupings				
All Populations	0.49	4.59*	2.78	6.80*
PT Populations	0.01	1.42	6.43*	7.30**
TN Populations	0.58	3.17t	0.08	1.73
K6 Groupings				
All Populations	1.30	3.04t	8.51**	3.80t
PT Populations	0.77	0.72	11.76**	1.81
TN Populations	0.53	3.06t	0.12	2.09

Degrees of freedom for each contrast are 1,48 for the original groupings and 1, 42 for the K6 groupings.
t $P < 0.10$, * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, **** $P < 0.0001$

Flowering time and number of leaves at flowering

Consistent with Friesen et al. (2014), we observed that earlier flowering plants with greater number of leaves at first flower had greater performance across all planting sites and locations (Table 7). Selection for these phenotypes was the strongest at CMS (Table 7) where, despite this strong relationship, flowering occurred 11% later compared to all other planting sites (Fig. 3A). Genotypic selection analyses also revealed that early flowering was favored at CMS, and the magnitude of selection often did not differ from other sites (Table 7).

Site of origin effects on flowering time were also observed (Table 6). Genotypes from CM_{loc} flowered 5.3% earlier than those from Gil_{loc} overall (Fig. 3A; $F = 17.95$, $P = 0.0002$). Similarly, in agreement with greenhouse common garden experiments performed by Friesen et al. (2014), our field studies revealed that

saline-origin Tunisian genotypes (TN1, TN8) flowered 7.3% earlier than non saline-origin Tunisian genotypes (TN7, TN9) at all planting sites (Fig. 3A; F = 13.44, P = 0.0025). Moreover, all planting sites differed for the number of leaves at first flowering, with population differentiation dependent upon planting site, but independent of the salinity of origin environment (Table 6; Fig. 3B). In particular, plants flowered at larger sizes at CMNS (28.6 leaves ± 1.0 SE), followed by GilS, CMS and GilNS (14.9%, 54.2% and 64.1% fewer leaves than CMNS, respectively; P ≤ 0.0002 for all pairwise comparisons).

Table 6. Results from mixed-model generalized linear models on age at first flowering (Flw Age), number of leaves at first flowering (Lvs at Flw), root water content (RWC), shoot water content (SWC), and root nodule numbers (Nodules). Country of origin (Origin); Population nested within country of origin (Pop), Planting population (PltPop), Planting site (PltSite), Transect nested within planting population and planting site, and genotype nested within country of origin and population (Genotype).

Source	df	Flw Age F-value	Lvs at Flw F-value	RWC F-value	SWC F-value	Nodules F-value
Origin	1	1.13	0.19	19.99****	0.01	7.28**
Pop	6	6.33****	3.52**	1.66	3.52**	0.48
PltLoc	1	63.22****	32.92****	80.91****	6.14*	153.60****
PltSite	2	52.72****	321.84****	151.23****	38.46****	14.45****
Origin X PltLoc	1	2.44	1.71	0.72	6.60*	4.33*
Origin X PltSite	2	6.25**	2.32t	0.68	8.78***	1.26
Pop X PltLoc	6	1.00	2.40*	1.48	0.68	0.47
Pop X PltSite	12	1.82*	2.64**	0.69	2.43**	2.46**
Transect	4	3.73**	4.07**	12.83****	4.97***	2.05t
Genotype	1	54.3****	7.4**	3.0t	4.68*	10.9***
Genotype X PltSite	1	0.10	0.00	0.00	0.10	3.40t

t $P < 0.10$, * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, **** $P < 0.0001$

Root and shoot water content

Analysis of root and shoot water content revealed positive correlations with all three measures of performance, and that selection was significantly stronger in less productive environments (Table 7). Thus plants with greater root and shoot water content had higher relative performance at CMS, while selection favored either lower plant water content or was not significant at CMNS (Table 7). Similarly, selection favored plants with increased shoot water content at the less productive GilNS compared to GilS. Nevertheless, plants grown in stressful environments (i.e., CMS and GilNS) tended to have lower trait values (Figs. 3C, D), and the best performing individuals were better able to maintain water content homeostasis under stress. The symmetry of selection responses between the less (i.e., CMS

and GilNS) and more (i.e., CMNS and GilS) productive planting sites further reinforces the notion that GilNS, despite its relatively low salinity values, is indeed the most stressful of Gil_{loc} sites.

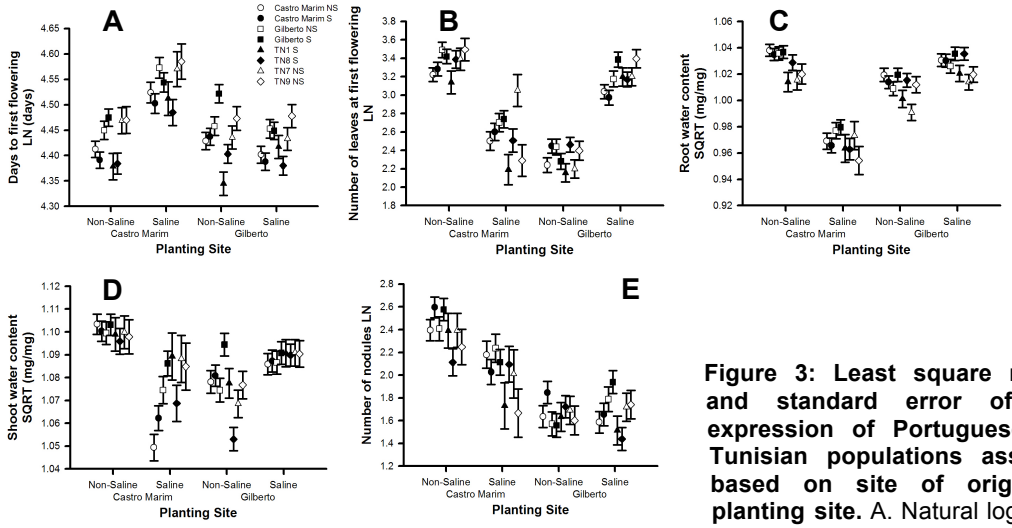


Figure 3: Least square means and standard error of trait expression of Portuguese and Tunisian populations assigned based on site of origin by planting site. A. Natural logarithm transformed number of days between germination and first flower, estimating age at flowering. B. Natural logarithm transformed number of leaves at first flower, estimating size at flowering. C. Squared root transformed root water content. D. Squared root transformed shoot water content. E. Natural logarithm transformed total number of nodules at the end of the experiment.

Nodule number.

Nodule number was positively correlated with performance only at CMS_{loc} and this relationship was significantly stronger at CMS than CMNS (Table 7). As we observed for flowering and water content traits, plants at the more stressful CMS had fewer nodules compared to CMNS, suggesting that performance at CMS is correlated with the ability to maintain nodulation rather than nodule number per se. More generally, we observed significant site of origin and planting site-specific differentiation for nodule number (Table 6; Fig. 3E), suggesting potentially complex factors governing symbiotic development.

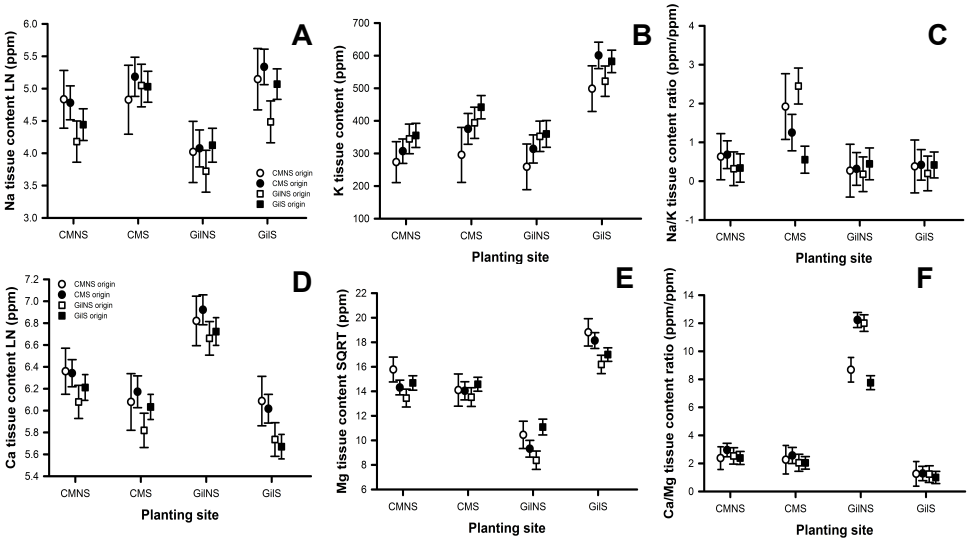


Figure 4: Least square means and standard error of Na, K, Na/K, Ca, Mg and Ca/Mg leaf ion content of a selection of Portuguese genotypes by planting site. A. Natural logarithm transformed Na leaf ion content. B. Mg leaf ion content. C. Na/K leaf ion content. D. Natural logarithm transformed Ca leaf ion content. E. Squared root transformed Mg leaf ion content. F. Ca/Mg leaf ion content.

Ion balances and absolute ion content in leaf tissue.

Significant planting site effects were detected for all ions measured in a subset of Portuguese genotypes (Table 8), suggesting that edaphic factors (e.g., soil nutrient composition) impact leaf ion content. Contrary to what is often observed in saline environments (e.g., Baxter et al. 2010; Arraouadi et al. 2012; Table S10, Supplemental Material), there was no correlation between leaf Na^+ content and performance at either saline planting site (Table 9, Fig. 4A). Nevertheless, saline-origin plants tended to maintain Na/K ratios more effectively than non saline-origin plants when grown in the high saline CMS environment (Figs. 4B, C), but this effect was not significant (Table 8). Moreover, plants with lower Na and higher K ion content had significantly greater fitness and biomass at GilNS.

We observed additional site-specific correlations for other ions. Thus at CMS, genotypes with higher leaf phosphorous showed significantly greater performance for all measures (e.g., fitness, biomass and genotype growth

Table 7. Phenotypic and genotypic, differential (or total, *s*) and gradient (direct, β), selection analyses are shown for both relative fitness and relative biomass and genotypic selection is shown for genotype growth rate for each planting site. Superscript symbols indicate significant differences between planting sites.

Trait	Castro Marim NS		Castro Marim S		Gilberto NS		Gilberto S	
	Total (*s*)	Direct (β)	Total (*s*)	Direct (β)	Total (*s*)	Direct (β)	Total (*s*)	Direct (β)
Relative Fitness								
Phenotypic								
Age at first flowering	-0.07^{**b}	-0.16****	-0.24^{****a}	-0.36***	-0.06^{tb}	-0.13*	-0.13^{****b}	-0.32****
Leaves at flowering	0.17^{****c}	0.27****	0.42^{****a}	1.04****	0.29^{****b}	0.56****	0.29^{****b}	0.50****
Root water content	-0.05^{tb}	-0.16***	0.32^{****a}	-0.28*	0.01^{b}	-0.13*	-0.02^{b}	-0.14*
Shoot water content	0.01^{c}	0.20****	0.30^{****a}	0.58****	0.17^{****b}	0.31****	-0.02^{c}	0.09**t**
Number of nodules	0.07^{**b}	0.11**	0.29^{****a}	0.11	0.12^{***b}	0.08****	0.10^{***b}	0.06
Genotypic								
Age at first flowering	-0.02^{ab}	-0.11*	-0.08^{**a}	-0.13	0.01^{b}	-0.10	-0.02^{b}	-0.10
Leaves at flowering	0.06^{****b}	0.14**	0.13^{****a}	0.24**	0.08^{****ab}	0.22**	0.06^{****b}	0.24**
Root water content	-0.02^{b}	-0.16**	0.05^{ta}	-0.13	0.01^{ab}	-0.08	-0.00^{b}	-0.11
Shoot water content	0.00^{a}	0.18**	0.02^{a}	0.15	0.05^{**a}	0.20*	-0.01^{a}	0.06
Number of nodules	0.04^{***b}	0.04	0.09^{***a}	0.13^{t}	0.01^{b}	-0.05	0.04^{**ab}	0.09
Relative Biomass								
Phenotypic								
Age at first flowering	-0.14^{****b}	-0.27****	-0.34^{****a}	-0.41****	-0.06^{tb}	-0.17***	-0.05^{b}	-0.23****
Leaves at flowering	0.23^{****c}	0.32****	0.44^{****a}	0.69****	0.29^{****b}	0.61****	0.09^{**ab}	0.59****
Root water content	-0.08^{**c}	-0.13**	0.33^{****a}	-0.13^{t}	0.01^{b}	-0.10^{t}	-0.05^{a}	-0.13*
Shoot water content	-0.06^{*b}	0.07^{t}	0.21^{****a}	0.30****	0.17^{****b}	0.29****	0.02^{a}	0.16*
Number of nodules	0.09^{***b}	0.09*	0.29^{****a}	0.19***	0.12^{***b}	-0.01	-0.01^{b}	0.13*
Genotypic								
Age at first flowering	-0.02^{*ab}	-2.30^{t}	-0.07^{****a}	-2.79***	0.01^{b}	-1.48	-0.03^{a}	-0.40
Leaves at flowering	0.05^{****b}	3.18**	0.11^{****a}	3.88****	0.08^{****ab}	3.19**	0.01^{a}	4.93****
Root water content	-0.02^{**c}	-2.53^{t}	0.06^{***a}	1.80*	0.01^{ab}	-0.01	-0.03^{*a}	-1.91*
Shoot water content	-0.01^{b}	1.87	0.03^{ab}	2.25*	0.05^{**a}	2.67*	-0.05*	1.10^{t}
Number of nodules	0.03^{**b}	1.61	0.07^{****a}	1.74*	0.01^{b}	-1.68^{t}	-0.07^{***a}	1.05
Genotype growth rate								
Genotypic Direct								
Age at first flowering	-0.17^{****ab}	-0.19****	-0.29^{****a}	-0.31***	-0.05^{b}	-0.05	-0.17^{****ab}	-0.16****
Leaves at flowering	0.02^{bc}	0.01	0.21^{**a}	0.22**	0.09^{**ab}	0.11***	-0.01^{c}	0.00
Root water content	-0.06^{a}	-0.11*	0.15^{ta}	-0.09	-0.05^{a}	-0.09**	0.05^{a}	-0.01
Shoot water content	0.01^{a}	0.13**	0.06^{a}	0.17^{t}	0.02^{a}	0.05	-0.02^{a}	0.02
Number of nodules	0.12^{**a}	0.12***	0.16^{*a}	0.01	-0.01^{b}	-0.01	0.07^{tab}	0.04
Genotypic Quadratic								
Age at first flowering	-0.04^{a}		0.03^{a}		-0.03^{a}		-0.07^{*a}	
Leaves at flowering	-0.08^{***a}		-0.05^{a}		0.01^{a}		0.02^{a}	
Root water content	0.01^{a}		-0.01^{a}		-0.07^{***a}		-0.00^{a}	
Shoot water content	-0.01^{a}		-0.13^{*a}		-0.03^{*a}		0.01^{a}	
Number of nodules	-0.02		-0.02		-0.05*		-0.01	

t $P < 0.10$, * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, **** $P < 0.0001$

rate), while genotypes with higher calcium content tended to only have greater biomass. Conversely, at CMNS, genotypes with greater fitness had lower Mg and Mn ion content, and genotypes with greater biomass had lower Ca, Mg and Zn ion content. Similar patterns were observed for Ca and Mg content (Figs. 4D,E) and for Ca/Mg ratios. Genotypes from CMNS and GilS showed better ability to maintain lower ratios when grown at GilNS (Fig. 4C), resulting in a significant site of origin by planting site interaction (Table 8). Zn content also revealed a significant site of origin by planting site interaction (Table 8). Unlike the phenological and

morphological traits (Table 6), variation in leaf ion concentration may reflect genotypic variation for plasticity (significant genotype by planting site interactions), with the strongest effects observed for Na^+, K^+ and Ca/Mg (Table 8). Leaf ion content (Na and K) of the Tunisian genotypes used in this experiment is further examined in Table S10 (Supplemental Material).

Table 8. Results from mixed-model generalized linear models on sodium (Na), potassium (K), calcium (Ca), magnesium (Mg), manganese (Mn), zinc (Zn), phosphorous (P), sodium-potassium ration (Na/K) and calcium-magnesium ratio (Ca/Mg) tissue ion content. Population (Pop), Planting location (PltLoc), Planting site (PltSite), Transect nested within planting population and planting site, and genotype nested within population (Gen).

Source	df	Na F-val	K F-val	Ca F-val	Mg F-val	Mn F-val	Zn F-val	P F-val	Na/K F-val	Ca/Mg F-val
Pop	3	0.95	1.40	1.40	2.60	0.78	0.18	2.99	0.50	3.15
PltLoc	1	7.22*	16.5***	5.38*	2.99	246.30****	72.00****	15.13**	8.14*	65.69****
PltSite	2	18.89****	24.34****	51.78****	79.15****	23.73****	13.64***	6.48**	4.34*	138.57****
Pop X PltLoc	3	0.51	0.41	0.64	0.15	2.50t	0.10	0.26	1.28	3.05^{t}
Pop X PltSite	6	0.51	0.41	0.29	0.97	1.79	4.18**	0.44	0.81	3.30*
Transect	4	1.44	1.39	0.89	1.76	2.81*	2.14^{t}	0.96	0.51	0.25
Gen	1	1.60	0.30	0.30	4.50*	0.70	0.30	4.10*	1.40	8.50**
Gen X PltSite	1	11.20***	3.10^{t}	16.50****	4.30*	0.10	0.80	4.50*	0.60	6.80**

t $P < 0.10$, * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, **** $P < 0.0001$

Discussion

In this study we sought to characterize adaptation to saline environments in populations of *Medicago truncatula* (Mtr) from Portugal, and to compare their responses *in natura* to those of previously characterized saline (S) and non-saline (NS) populations from Tunisia (e.g., Castro *et al.* 2013; Cordeiro *et al.* 2014; Friesen *et al*. 2014). Genomic data confirmed that Tunisian and Portuguese genotypes are reproductively isolated from one another and differentiated genetically, and that both the extent and structure of genetic variation differ between Tunisian and Portuguese populations. A field experiment was conducted using genotypes from four Portuguese and four Tunisian sites of origin that were planted into saline and non-saline sites at two locations in southern Portugal: Gilberto (Gil_{loc}: GilS and GilNS) and Castro Marim (CM_{loc}: CMS and CMNS). Plant performance and traits associated with mechanisms of salinity tolerance and/or avoidance were quantified to identify traits that contribute to fitness, to infer underlying mechanisms, and to test the hypothesis of local adaptation.

Table 9. Results from phenotypic and genotypic, differential (or total, s) and gradient (direct, β), selection analyses are shown for each planting site for both relative fitness and relative biomass, and genotypic selection for genotype growth rate. Superscript symbols indicate significant differences between planting sites.

Trait	Castro Marim NS		Castro Marim S		Gilberto NS		Gilberto S	
	Total (*s*)	Direct (β)	Total (*s*)	Direct (β)	Total (*s*)	Direct (β)	Total (*s*)	Direct (β)
Relative Fitness								
Phenotypic								
Na	-0.09a	-0.11	-0.04a	0.20	-0.36*a	-0.15	0.01a	-0.07
K	0.08a	-0.05	0.22a	-0.06	0.39*a	0.13	0.04a	0.03
Ca	-0.25***a	0.01	0.17b	0.56**	-0.04ab	0.15	-0.24ta	0.02
Mg	-0.25***a	-0.09	-0.01a	0.12	-0.01a	-0.21	-0.18a	0.02
Mn	-0.14ta	-0.05	0.09a	0.35*	-0.14a	-0.04	-0.11a	-0.13
Zn	-0.07a		0.01a		0.22a		-0.02a	
P	0.02a		0.41**b		0.19a		0.02a	
Na/K	-0.09a		-0.09a		-0.34*a		0.03a	
Ca/Mg	-0.09a		0.15a		-0.12a		-0.15a	
Genotypic								
Na	-0.12a	-0.30	-0.10a	-1.16	-0.33a	-0.19	-0.11a	-1.16
K	0.12a	0.06	0.20a	-0.33	0.19a	-0.42	0.02a	0.45
Ca	-0.15a	-0.05	0.06a	0.70	-0.03a	-1.22	-0.18a	1.41
Mg	-0.24*a	-0.16	-0.11a	0.49	-0.24a	-0.30	-0.16a	-0.71
Mn	-0.30*a	0.13	-0.19a	-0.91	-0.12a	-1.59	-0.21a	1.46
Zn	0.03a		-0.21a		-0.37a		0.00a	
P	0.09a		0.33b		-0.13a		0.10ab	
Na/K	-0.09a		-0.21a		-0.19a		-0.10a	
Ca/Mg	0.02a		0.11a		0.05a		-0.11a	
Relative Biomass								
Phenotypic								
Na	-0.03a		0.11a		-0.26ta		-0.00a	
K	0.04a		0.05a		0.36*a		0.05a	
Ca	-0.27*a		0.27tb		-0.10ab		-0.28a	
Mg	-0.23*a		0.03a		-0.10a		-0.24a	
Mn	-0.15a	-0.06	0.01a	0.17	-0.19a	-0.36	-0.11a	-0.05
Zn	-0.24*a	-0.18	-0.12ab	-0.10	0.32*b	0.29t	-0.01a	-0.09
P	-0.09a	-0.04	0.31ta	0.48**	0.19a	0.27	-0.02a	-0.04
Na/K	-0.02a	-0.07	0.07a	0.26	-0.27ta	-0.00	0.01a	0.01
Ca/Mg	-0.17ta	-0.13	0.23a	0.40*	-0.13a	-0.18	-0.18a	-0.19
Genotypic								
Na	-0.21ta		-0.02a		-0.15a		-0.19a	
K	0.04a		0.12a		0.05a		0.10a	
Ca	-0.16a		0.06a		0.02a		-0.22a	
Mg	-0.13a		-0.15a		-0.45*a		-0.19a	
Mn	-0.21a	-0.15	-0.20a	-1.55	-0.18a	-0.13	-0.21a	-1.49
Zn	-0.16a	-0.10	-0.28a	-0.82	0.05a	0.11	0.13a	0.72
P	0.01a	-0.12	0.22a	0.74	-0.06a	0.03	0.06a	1.67
Na/K	-0.16a	-0.31	-0.10a	0.81	-0.17a	-0.12	-0.19a	-0.99
Ca/Mg	-0.10a	0.15	0.13a	-1.53	0.14a	0.12	-0.15a	1.92
Gen growth rate								
Genotypic								
Na	0.10a		0.11a		-0.02a		0.21ta	
K	-0.09a		0.09a		-0.03a		-0.06a	
Ca	0.05a		0.16a		0.29*a		0.14a	
Mg	-0.03a		-0.00a		-0.04a		0.12a	
Mn	-0.05a	-0.07	-0.06a	-0.36*	0.06a	0.40t	-0.11a	-0.13
Zn	-0.04a	-0.06	-0.07a	0.16	-0.18a	-0.21	0.03a	0.03
P	-0.03a	0.09	0.27**b	0.45**	-0.22a	-0.27	-0.02ab	0.11
Na/K	0.11b	0.12	-0.15a	0.15	-0.01ab	-0.11	0.20ab	0.15
Ca/Mg	0.16ta	0.10	0.12a	0.07	0.28ta	0.19	0.07a	0.13

t $P < 0.10$, * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, **** $P < 0.0001$

Variable selection and adaptation at saline and non-saline sites

Soil analyses confirmed that all measures of soil salinity were greater in saline compared with non saline planting sites (Tables S2-4, Supplemental Material). As expected, viability traits (i.e., germination and survival to reproduction) were negatively affected at higher salinity (Figs. 2A and B). Thus, it is likely that viability selection is operative at the Portuguese saline sites, which is consistent with prior observations for Tunisian genotypes (Cordeiro *et al.* 2014) and with known importance of germination timing in the survival and early growth of annual plant seedlings, the most stress-sensitive phase of plant development (Donohue *et al.* 2005; Gremer & Venable 2014).

Fecundity-related measurements (i.e., fitness, biomass and genotype growth rate) indicate that salinity and environmental stress are concordant at CM_{loc} (Figs. 2D, E, H and I). The higher performance of saline-origin genotypes at CMS relative to CMNS, including crossing reaction norms in Figures 2H and I, support the conclusion that salinity is a main driver of performance in the CMS/CMNS pair and is consistent with local adaptation of CMS and CMNS genotypes. More generally, these results suggest that salinity adaptation has evolved in all saline-origin groups, including those from GilS (Fig. 2H, I). This latter observation is intriguing because all groups perform more poorly at GilNS compared to GilS. Thus, salinity and environmental stress are non-concordant at Gil_{loc}, suggesting that other factors (e.g., absolute and relative quantity of other soil nutrients; Tables S3, S5, Supplemental Material) impact plant performance and either obscure or obviate salinity stress at Gil_{loc} (Bernstein 1975; Marcelis & van Hooijdonk 1999; Singh *et al.* 2014). One interpretation of these data is that salinity adaptation is maladaptive at Gil_{loc}, because non-saline origin populations from both countries tended to outperform saline-origin populations at GilS. In particular, saline and non-saline Tunisian populations were strongly differentiated for total biomass (Fig. 2G) and to a lesser extent for genotype growth rate at GilS (Fig. 2I). Another non-mutually exclusive interpretation, consistent with the observation of salinity

adaptation in GilS origin genotypes, is that there is temporal variation of selection at this site.

Hereford's (2009) meta-analysis concluded that composite measures of performance and fecundity related traits tend to be more associated with local adaptation than viability traits. Indeed, our phenotypic selection analyses indicate that both flowering time and plant size at reproduction (number of leaves at first flower) are more strongly selected at CMS than at any other site (Table 7). Interestingly, the flowering time response was differentiated between saline and non-saline Tunisian populations, while for Portuguese genotypes, flowering time was differentiated by location of origin and not salinity (Fig. 3A). Such standing variation for flowering time responses to salinity at the Portuguese sites of origin correlates with a more complex geographic pattern of population structure in Portugal relative to Tunisia (see below).

Physiological traits (i.e., root and shoot water content, as well as nodulation by local rhizobia) were more strongly selected at CMS relative to other sites. Shoot water content was also under selection at the more stressful GilNS site, but less so than at CMS (Table 7). In the case of shoot water content, differentiation was strong and of similar directionality within the CMS/CMNS and GilS/GILNS pairs (Fig. 3D), suggesting that shoot water content responses may represent adaptation to soil salinity, despite the otherwise non-concordance between performance and salinity at Gilberto. More generally, across all origins, stressful planting sites (i.e., CMS and GilNS) were associated with reduced trait values, and the highest performing genotypes tended to be those that were best able to maintain water balances. Local populations had qualitatively greater performance at CMS and GilNS, which although not statistically significant, suggests the possibility of home site advantage at the most stressful planting sites (Fig. 2H).

Population diversity and structure differ markedly between Portugal and Tunisia

In their study of range-wide variation in Mtr, Yoder *et al.* (2014) determined that genetic differentiation is associated with geographic and climatic variables and that a portion of the variation is adaptive. Similarly, our data suggest that Portuguese and Tunisian genotypes comprise genetically distinct populations, that Portuguese populations harbor significantly more genetic variation than Tunisian populations, and that a portion of Portuguese variation is adaptive.

In contrast to the case in Tunisia, where genetic populations are synonymous with site of origin (Friesen *et al.* 2014), we observed a more complicated situation in Portugal. Primary differentiation was between locations and secondarily between and within sites, as reflected in STRUCTURE assignments (as well as neighbor joining and F_{ST} analyses; Tables 1-3, Fig. 1). However, genetic differentiation was not solely structured by geography, as within each location we observed both multiple populations at single sites and single populations (or populations with low differentiation) among sites. The substantially longer LD observed in Portuguese (~ 65 kb) compared to Tunisian (~10 kb; Friesen *et al.* 2014) genotypes likely reflects a more recent population expansion in Portugal through seed dispersal combined with population admixture, as may be the case for CMNS and CMS. In support of this conclusion, while Portuguese genotypes are on average more genetically diverse than Tunisian genotypes, there is less diversity among individuals within certain Portuguese genetic sub-populations. The maintenance of distinct lineages within and among individual Portuguese sites could also reflect the influence of fine-scale spatial or temporal variation and low gene flow among lineages (Kawecki & Ebert 2004; Postma & van Noordwijk 2005; Hereford 2009). In any case, such factors would likely confound detection of selection and could favor phenotypic plasticity over local adaptation for certain populations (Lenormand 2002; Hereford *et al.* 2009).

Do Tunisian and Portuguese genotypes share the same mechanism to deal with salinity?

Plants have evolved different mechanisms to deal with soil salinity including tolerance, avoidance and exclusion (Flowers *et al.* 1977; Munns & Tester, 2008). Salinity adaptation in Tunisian genotypes is consistent with the evolution of both salinity avoidance (e.g., earlier germination and flowering) and tolerance mechanisms (Castro *et al.* 2013; Cordeiro *et al.* 2014; Friesen *et al.* 2014), and candidate genes for both processes have been nominated among Tunisian genotypes (Friesen *et al.* 2014). Avoidance strategies allow saline adapted plants to reproduce before salt builds up in the soil or reaches toxic levels in the plant (Munns & Tester 2008; Friesen *et al.* 2014). Here we also observed that saline origin Tunisian genotypes tended to flower earlier and have more leaves at flowering than non saline-origin Tunisian genotypes (Fig. 3A). In particular, flowering time ranked among the most strongly selected traits for relative fitness and biomass at CM_{loc} (Table 7), where signal for salinity adaptation has been observed (Fig. 2I).

Selection also favored earlier flowering among Portuguese genotypes (Table 7) although, in contrast to Tunisian genotypes, flowering time did not assort with origin-site salinity but rather by location of origin. More generally, phenotypic selection analyses indicate that a suite of traits are under selection at saline sites (Table 7; Fig. 3A), including increased number of leaves at flowering, greater plant water content, and increased nodulation. Selection was strongest at the most stressful planting sites (i.e., CMS and GilNS, Table 7), and independent of country of origin (data not shown), suggesting that similar or related mechanisms have evolved in Tunisian and Portuguese genotypes. Examples of parallel phenotypic divergence are found in a range of species including humans (Tennessen & Akey 2011), beach mice (Hoekstra *et al.* 2006), marine snails (Westram *et al.* 2014), black cottonwood (Holliday *et al.* 2015) and *Arabidopsis lyrata* (Turner *et al.* 2010), and when trait expression is correlated with a common environmental variable it is usually interpreted as being adaptive. Interestingly, while this seems to be the case

for salinity tolerance related traits, the opposite is true for early flowering. Early flowering is under strong selection and appears to be a salinity avoidance mechanism in Tunisian Mtr populations, while it is decoupled from salinity in Portuguese populations.

Differentiation among Portuguese genotypes for flowering time, root water content and nodule number reflected location rather than site of origin, with Gil_{loc} plants tending to have higher trait values (Fig. 3A and B). Only in the case of shoot water content was there evidence of differentiation by site-of-origin salinity among locations (Fig. 3D), suggesting a possible basis of salinity adaptation among CMS genotypes (Fig. 2H). On average these saline-origin Portuguese genotypes were better able to maintain Na/K ion balances, with significant differences observed between GilS and GilNS origin (Fig. 4C), providing a mechanistic correlation with shoot water content. Interestingly, there was a greater range of trait expression for flowering time, number of leaves at first flower, and number of nodules (Fig. 3A, B and E) in Tunisian compared with Portuguese sites of origin. This observation is consistent with the colonization of new environments by Tunisian genotypes specialized for factors unique to Tunisian sites, while the relative uniformity of Portuguese genotypes suggests a regionally adapted generalist strategy (van Tienderen 1991; Kassen 2002).

Interestingly, Tunisian genotypes typically performed better or similarly to Portuguese genotypes (e.g., Figs. 2B-G), despite the fact that Portuguese genotypes would presumably have home site advantage. Various factors might explain why performance was greater for Tunisian compared with Portuguese genotypes (Fig. 2), including temporal variation in the strength of selection or the release from locally adapted biotic stresses. One intriguing possibility is shifting climatic conditions (Aitken *et al.* 2008; Wilczek *et al.* 2014; Hamilton *et al.* 2015). Indeed, the climate of southern Portugal is shifting towards that typical of historical northern Tunisia (New *et al.* 2002; Giorgi & Lionello 2008; Kuglitsch *et al.* 2009; Hoerling *et al*. 2012), which could theoretically shift species' biogeography (Etterson & Shaw 2001; Charmantier *et al.* 2008; Atkins *et al.* 2010). Alternatively,

differences between Portuguese and Tunisian accessions may derive from random demographic processes that limit genetic variation and thus the extent of possible adaptation, or from fitness trade-offs driven by adaptations to other environmental factors (Travisano *et al.* 1995; Hereford 2009; Yoder *et al.* 2014).

Which candidate genes may explain the patterns of trait expression in Portuguese genotypes?

Genotypes from CM and Gil are genetically differentiated (Table 2) and geographically separate. Between these locations we identified 84 genes within 36 haplotype blocks that contained location-specific, non-synonymous alleles (Table S8, Supplemental Material). Disease resistance genes in the NB-ARC family represented 7% of these total location-specific alleles, which is 6-fold more abundant than expected based on their genome-wide frequency (i.e., ~800 NB-ARC proteins among ~65,000 predictions in the reference Mtr genome; Young *et al.* 2011), suggesting that disease pressure and/or pathogen genotypes may vary between CM and Gil. Numerous other location-specific alleles encode regulators of transcription, translation and protein turnover, and components of signal transduction and hormone signaling, transport and biosynthesis pathways. The major developmental and physiological traits differentiated between CM and Gil origin genotypes, which include flowering time, nodule number and root water content (Figs. 3A, C, E), may directly or indirectly derive from functions of these differentiated protein isoforms. For example, Medtr8g010160 is orthologous to the FRIGIDA-like At5g48385 and could underlie the observed differences in flowering time between CM and Gil (Fig. 3A; Schläppi 2006; Schmalenbach *et al*. 2014). In particular, the Gil allele of Medtr8g010160 is altered in a highly conserved amino acid residue that is predicted by PROVEAN to result in altered function (Choi *et al.* 2012). Several genes located on chromosome 8 encode proteins involved in ethylene biosynthesis or signaling and given the pleiotropic nature of ethylene perception could impact site-specific disease resistance, abiotic stress adaptation,

or the observed differential nodulation phenotypes (Penmetsa & Cook 1997; Manavella *et al.* 2006; Penmetsa *et al.* 2008; Larrainzar *et al.* 2015; Fig. 3E). We also observed differentiation of genes implicated in ion homeostasis (Medtr1g011840 and Medtr8g016270), which might represent adaptations to soil differences among locations, including ionic balances and soil cation exchange capacity, and ultimately to differential root water content (Tables 8, S3 and S5, Supplemental Material, Fig. 3C).

We observed 38 genes, among 23 LD blocks, with non-synonymous allelic differentiation between saline and non saline origin genotypes across locations. The absence of NB-ARC disease resistance genes within this list of habitat-differentiated genes suggests that pathogen selection pressure is structured by location rather than habitat salinity. A substantial proportion of genes that assort by salinity are implicated in chromatin remodeling, regulation of transcription and translation, and signal transduction; these and other inferred functionalities are candidates for salinity adaptation. Medtr7g033135 is of particular interest because it is a close homolog of *A. thaliana* CRK42, At5g40380, a gene annotated as a salt stress-dependent, hormone-mediated protein kinase (Jones *et al.* 2014) and thus a candidate to confer salinity adaptation among Portuguese genotypes. Medtr2g461240 is homologous to *A. thaliana* KNAT3 (At5g25220), which regulates the expression of ABI3 and modulates ABA responses during germination and early seedling development (Kim *et al.* 2013). ABA is a key regulator of abiotic stress, especially salinity and water stress (Golldack *et al.* 2014), and germination and early seedling development are life cycle stages exhibiting differential tolerance to salinity among Tunisian *Mtr* genotypes (Cordeiro *et al.* 2014). Thus Medtr2g461240 is a strong candidate to confer for salinity adaptation during early development of Portuguese genotypes. Other *Mtr* genes that assort by habitat have functions that are consistent with adaptation to soil properties, including close homologs of *A. thaliana* genes involved in nitrate transport (Medtr5g038060; At1g69870, Fan *et al.* 2009) and phosphate starvation response (Medtr2g460821; At2g13100, Ramaiah *et al*. 2011), as well as the ABC transporter Medtr6g034265

that may regulate cellular homeostasis in response to unspecified factors (Kang *et al.* 2011). Including the many homologs of regulatory genes, these protein isoforms that assort by habitat salinity are candidates to maintain ion and water balances in saline environments.

Conclusion

Environmental factors other than salinity, such as soil ion balances and pathogen pressure, are likely to be driving performance at the Portuguese planting sites. Nevertheless, Portuguese saline-origin populations do display evidence of salinity adaptation, which we conclude based on i) their differential performance at the CMS site, ii) the parallel responses of saline-adapted genotypes from Tunisia, and iii) the nature of phenotypes under selection.

The strong population structure identified within sites of origin in Portugal suggests that seed dispersal occurs at higher rates in southern Portugal compared with northern Tunisia. Combined with the observation of increased genetic variation in Portuguese relative to Tunisian genotypes, especially between co-occurring genetic populations, there is potentially more adaptive capacity in the Portuguese populations. Nevertheless, the frequent equal or greater performance of the Tunisian genotypes in Portuguese environments is consistent with the possibility that current selective constraints in Portugal are more similar to those historically present in Tunisia, as might be expected if climate change is shifting environmental adaptations along the Mediterranean basin.

Due to the geographic and genetic differentiation between Tunisian and Portuguese populations, we suggest that evolution for salinity adaptation likely evolved independently in Tunisian and Portuguese genotypes. The absence of common candidate genes between this analysis and previous analysis of Tunisian genotypes (Friesen *et al.* 2014) is consistent with this interpretation. Nevertheless, the apparent parallel phenotypic divergence, with similar traits under selection, suggests that common developmental and physiological processes are targeted by

selection in both regions. For example, the early flowering of Tunisian saline-origin genotypes assorts with allelic variation in an ortholog of the *A. thaliana* CONSTANS (Friesen *et al.* 2014), while the same trait in Portuguese genotypes assorts by location of origin, not salinity, and a possible candidate mutation occurs in the homolog of an *A. thaliana* FRIGIDA-like protein. Thus, we suggest that despite our interpretation of parallel adaptive evolution, Portuguese and Tunisian populations have arrived at their specific adaptations by selection on distinct genetic components.

Acknowledgements

Matilde Cordeiro (MAC) was involved in all the conceptualization, materialization, analyses and interpretation of the work described in this chapter. As explicit in the submitted manuscript: MAC, KSM, EVW, MLF, SYS, SVN and DRC designed research; MAC, KSM, EVW, MLF, SYS and SP collected the Portuguese populations; MAC and KMS ran the field experiment with help from EVW, MLF, WV, PSF, SYS, SVN and DRC; MAC generated PT ions data; EVW with help from KS and JM generated TN ion data; HKC provided material for sequencing; JHM and MJO with help from CL and JHK sequenced the Portuguese genotypes; PLC and NK (with help from WJL) and MAC analyzed genomic data; MAC, KSM and EVW analyzed data; MAC, KSM and DRC wrote the paper with input from all authors.

The authors thank Srs. Gilberto Ferro, António Firmino and Carlos Lopes for providing the field sites for collection and experiments; Srs. Drs. João Alves e Marco Rebelo from Instituto da Conservação da Natureza e das Florestas, Sr. Eng. Nuno Grade and others from Parque Natural da Ria Formosa and from Reserva Natural do Sapal de Castro Marim e Vila Real de Stº António for supporting the field experiments and making the park's infrastructures available; Lisa Vance, Emily Bergmann and Luciano for hands-on help; David Salt for advice in ion content protocol and analysis; Joel Commisso from the ICP-MS lab in

University of California, Davis; and the anonymous reviewers for useful comments. This work was financed by the National Science Foundation (grant PGRP 08-20846), the Next-Generation BioGreen21 Program (PJ011182) by the Rural Development Administration Republic of Korea, and a fellowship from the Fundação para a Ciência e a Tecnologia (grant SFRH/ BD/39905/2007 to MAC) and the USDA NNF program (USDA-NIFA-NNF 2011-38420-20053, to EvW support KS).

References

Aitken SN, Yeaman S, Holliday JA, Wang T, Curtis-McLane S (2008) Adaptation, migration or extirpation: climate change outcomes for tree populations. *Evolutionary applications,* **1**, 95-111.

Arraouadi S, Badri M, Abdelly C, Huguet T, Aouani ME (2012) QTL mapping of physiological traits associated with salt tolerance in *Medicago truncatula* Recombinant Inbred Lines. *Genomics,* **99**, 118-125.

Atkins KE, Travis JMJ (2010) Local adaptation and the evolution of species' ranges under climate change. *Journal of Theoretical Biology*, **266**, 449-457.

Badri M, Ilahi H, Huguet T, Aouani ME (2007) Quantitative and molecular genetic variation in sympatric populations of *Medicago laciniata* and *M. truncatula* (Fabaceae): relationships with eco-geographical factors. *Genetical Research,* **89**, 107-122.

Baxter I, Brazelton JN, Yu D, Huang YS, Lahner B, Yakubova E, Li Y, Bergelson J, Borevitz JO, Nordborg M, et al (2010) A coastal cline in sodium accumulation in *Arabidopsis thaliana* is driven by natural variation of the sodium transporter AtHKT1;1. *PLoS genetics,* **6**: e1001193.

Bernstein L (1975) Effects of salinity and sodicity on plant growth. *Annual Review of Phytopathology,* **13**, 295-312.

Branca A, Paape TD, Zhou P, Briskine R, Farmer AD, Mudge J, Bharti AK, Woodward JE, May GD, Gentzbittel L, et al (2011) Whole-genome nucleotide diversity, recombination, and linkage disequilibrium in the model legume *Medicago truncatula*. *Proceedings of the National Academy of Sciences of the United States of America,* **108**, E864-E8.

Busoms S, Teres J, Huang XY, Bomblies K, Danku J, Douglas A, Weigel D, Poschenrieder C, Salt DE (2015) Salinity is an agent of divergent selection driving local adaptation of Arabidopsis to coastal habitats. *Plant physiology,* **168**, 915-929.

Castro BM, Moriuchi KS, Friesen ML, Badri M, Nuzhdin SV, Strauss SY, Cook DR, Wettberg E (2013) Parental environments and interactions with conspecifics alter salinity tolerance of offspring in the annual *Medicago truncatula*. *Journal of Ecology,* **101**, 1281-1287.

Charmantier A, McCleery RH, Cole LR, Perrins C, Kruuk LE, Sheldon BC (2008) Adaptive phenotypic plasticity in response to climate change in a wild bird population. *Science,* **320**, 800-803.

Choi Y, Sims GE, Murphy S, Miller JR, Chan AP (2012) Predicting the functional effect of amino acid substitutions and indels. *PloS one,* **7**: e46688.

Clausen J, Keck DD, Hiesey WM (1940) Experimental studies on the nature of species. I. The effects of varied environments on western American plants. Carnegie Institution of Washington Publications **520**.

Conner JK (1996) Understanding natural selection: an approach integrating selection gradients, multiplicative fitness components, and path analysis. *Ethology Ecology & Evolution*, **8**, 387-397.

Cook DR (1999) *Medicago truncatula*: a model in the making! *Current Opinion in Plant Biology,* **2**, 301-304.

Cordeiro MA, Moriuchi KS, Fotinos TD, Miller KE, Nuzhdin SV, von Wettberg EJ, Cook DR (2014) Population differentiation for germination and early seedling root growth traits under saline conditions in the annual legume *Medicago truncatula* (Fabaceae). *American journal of botany,* **101**, 488-498.

Courbot M, Willems G, Motte P, Arvidsson S, Roosens N, Saumitou-Laprade P, Verbruggen N (2007) A major quantitative trait locus for cadmium tolerance in *Arabidopsis halleri* colocalizes with HMA4, a gene encoding a heavy metal ATPase. *Plant physiology,* **144**, 1052-1065.

De Lorenzo L, Merchan F, Blanchet S, Megías M, Frugier F, Crespi M, Sousa C (2007) Differential expression of the TFIIIA regulatory pathway in response to

salt stress between *Medicago truncatula* genotypes. *Plant physiology,* **145**, 1521-1532.

DePristo MA, Banks E, Poplin R, Garimella KV, Maguire JR, Hartl C, Philippakis AA, del Angel G, Rivas MA, Hanna M, et al (2011) A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nature genetics,* **43**, 491-498.

Donohue K, Dorn L, Griffith C, Kim E, Aguilera A, Polisetty CR, Schmitt J (2005) The evolutionary ecology of seed germination of *Arabidopsis thaliana*: variable natural selection on germination timing. *Evolution; international journal of organic evolution,* **59**, 758-770.

Donohue K, Messiqua D, Pyle EH, Heschel MS, Schmitt J (2000) Evidence of adaptive divergence in plasticity: density-and site-dependent selection on shade-avoidance responses in *Impatiens capensis*. *Evolution,* **54**, 1956-1968.

Etterson JR, Shaw RG (2001) Constraint to adaptive evolution in response to global warming. *Science*, **294**, 151-154.

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology*, **14**, 2611-2620.

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

Fan SC, Lin CS, Hsu PK, Lin SH, Tsay YF (2009) The Arabidopsis nitrate transporter NRT1.7, expressed in phloem, is responsible for source-to-sink remobilization of nitrate. *The Plant cell,* **21**, 2750-2761.

Flowers TJ, Troke PF, Yeo AR (1977) The mechanism of salt tolerance in halophytes. *Annual Review of Plant Physiology*, **28**, 89-121.

Fournier-Level A, Korte A, Cooper MD, Nordborg M, Schmitt J, Wilczek AM (2011)

A map of local adaptation in *Arabidopsis thaliana*. *Science (New York, N.Y.),* **334**, 86-89.

Friesen ML, Cordeiro MA, Penmetsa RV, Badri M, Huguet T, Aouani ME, Cook DR, Nuzhdin SV (2010) Population genomic analysis of Tunisian *Medicago truncatula* reveals candidates for local adaptation. *The Plant journal,* **63**, 623-635.

Friesen ML, von Wettberg EJ, Badri M, Moriuchi KS, Barhoumi F, Chang PL, Cuellar-Ortiz S, Cordeiro MA, Vu WT, Arraouadi S, et al (2014) The ecological genomic basis of salinity adaptation in Tunisian *Medicago truncatula*. *BMC genomics,* **15**: 1160.

Fry JD (1996) The evolution of host specialization: are trade-offs overrated? *American Naturalist*, **148**, S84-S107.

Galloway LF, Fenster CB (2000) Population differentiation in an annual legume: local adaptation. *Evolution*, **54**, 1173-1181.

Giorgi F, Lionello P (2008) Climate change projections for the Mediterranean region. *Global and Planetary Change*, **63**, 90-104.

Golldack D, Li C, Mohan H, Probst N (2014) Tolerance to drought and salt stress in plants: Unraveling the signaling networks. *Frontiers in Plant Science*, **5**, 151.

Gould B, McCouch S, Geber M (2014) Variation in soil aluminium tolerance genes is associated with local adaptation to soils at the Park Grass Experiment. *Molecular Ecology*, **23**, 6058-6072.

Gremer JR, Venable DL (2014) Bet hedging in desert winter annual plants: optimal germination strategies in a variable environment. *Ecology Letters*, **17**, 380-387.

Hall BG, Ehrlich GD, Hu FZ (2010) Pan-genome analysis provides much higher strain typing resolution than multi-locus sequence typing. *Microbiology*, **156**, 1060-1068.

Hamilton JA, Okada M, Korves T, Schmitt J (2015) The role of climate adaptation in colonization success in *Arabidopsis thaliana*. *Molecular Ecology*, **24**, 2253-2263.

Hansen TF, Carter AJR, Pélabon C (2006) On adaptive accuracy and precision in natural populations. *American Naturalist*, **168**,168–181.

Hereford J (2009) A quantitative survey of local adaptation and fitness trade-offs. *American Naturalist*, **5**, 579-588.

Hereford J, Winn AA (2008) Limits to local adaptation in six populations of the annual plant *Diodia teres*. *New Phytologist,* **178**, 888–896.

Hoekstra HE, Hirschmann RJ, Bundey RA, Insel PA, Crossland JP (2006) A single amino acid mutation contributes to adaptive beach mouse color pattern, *Science*, **313**, 101-104.

Hoerling M, Eischeid J, Perlwitz J, Quan X, Zhang T, Pegion P (2012) On the increased frequency of Mediterranean drought. *Journal of Climate,* **25**, 2146-2161.

Holliday JA, Zhou L, Bawa R, Zhang M, Oubida RW (2015) Evidence for extensive parallelism but divergent genomic architecture of adaptation along altitudinal and latitudinal gradients in *Populus trichocarpa*. *New Phytologist*, **doi**, 10.1111/nph.13643.

Huang X, Schmitt J, Dorn L, Griffith C, Effgen S, Takao S, Koornneef M, Donohue K (2010) The earliest stages of adaptation in an experimental plant population: strong selection on QTLS for seed dormancy. *Molecular ecology,* **19**, 1335-1351.

Ingvarsson PK, Street NR (2011) Association genetics of complex traits in plants. *New Phytologist,* **189**, 909-922.

Jones AM, Xuan Y, Xu M, Wang RS, Ho CH, Lalonde S, You CH, Sardi MI, Parsa SA, Smith-Valle E, et al (2014) Border control--a membrane-linked interactome

of Arabidopsis. *Science (New York, N.Y.),* **344**, 711-716.

Kang J, Park J, Choi H, Burla B, Kretzschmar T, Lee Y, Martinoia E (2011) Plant ABC Transporters. *The Arabidopsis book / American Society of Plant Biologists,* **9**: e0153.

Kassen R (2002) The experimental evolution of specialists, generalists, and the maintenance of diversity. *Journal of Evolutionary Biology*, **15**, 173-190.

Kawecki TJ, Ebert D (2004) Conceptual issues in local adaptation. *Ecology Letters*, **7**, 1225-1241.

Kim D, Cho YH, Ryu H, Kim Y, Kim TH, Hwang I (2013) BLH1 and KNAT3 modulate ABA responses during germination and early seedling development in Arabidopsis. *The Plant journal,* **75**, 755-766.

Kuglitsch FG, Toreti A, Xoplaki E, Della-Marta PM, Luterbacher J, Wanner H (2009) Homogenization of daily maximum temperature series in the Mediterranean. *Journal of Geophysical Research: Atmospheres,* **114**.

Lahner B, Gong J, Mahmoudian M, Smith EL, Abid KB, Rogers EE, Guerinot ML, Harper JF, Ward JM, McIntyre L (2003) Genomic scale profiling of nutrient and trace elements in *Arabidopsis thaliana*. *Nature biotechnology,* **21**, 1215-1221.

Lande R, Arnold SJ (1983) The measurement of selection on correlated characters. *Evolution*, **37**, 1210-1226.

Larrainzar E, Riely BK, Kim SC, Carrasquilla-Garcia N, Yu HJ, Hwang HJ, Oh M, Kim GB, Surendrarao AK, Chasman D, et al (2015) Deep Sequencing of the Medicago truncatula Root Transcriptome Reveals a Massive and Early Interaction between Nodulation Factor and Ethylene Signals. *Plant physiology,* **169**, 233-265.

Lazrek F, Roussel V, Ronfort J, Cardinet G, Chardon F, Aouani ME, Huguet T (2009) The use of neutral and non-neutral SSRs to analyse the genetic structure of a Tunisian collection of Medicago truncatula lines and to reveal

associations with eco-environmental variables. *Genetica,* **135**, 391-402.

Le Corre V, Kremer A (2012) The genetic differentiation at quantitative trait loci under local adaptation. *Molecular Ecology*, 21, 1548-1566.

Lenormand T (2002) Gene flow and the limits to natural selection. *Trends in Ecology & Evolution*, **17**, 183-189.

Li H, Durbin R (2010) Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics*, **26**, 589-595.

Lowry DB, Hall MC, Salt DE, Willis JH (2009) Genetic and physiological basis of adaptive salt tolerance divergence between coastal and inland *Mimulus guttatus*. *New Phytologist*, **183**, 776-788.

Lowry DB, Rockwood RC, Willis JH (2008) Ecological reproductive isolation of coast and inland races of *Mimulus guttatus*. *Evolution*, **62**, 2196-22.

Macnair MR (1983) The genetic control of copper tolerance in the yellow monkey flower*, Mimulus guttatus*. *Heredity*, **50**, 283-293.

Manavella PA, Arce AL, Dezar CA, Bitton F, Renou JP, Crespi M, Chan RL (2006) Cross-talk between ethylene and drought signalling pathways is mediated by the sunflower Hahb-4 transcription factor. *The Plant journal,* **48**, 125-137.

Marcelis LFM, Van Hooijdonk J (1999) Effect of salinity on growth, water use and nutrient use in radish (*Raphanus sativus* L.). *Plant and Soil*, **215**, 57-64.

McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, Garimella K, Altshuler D, Gabriel S, Daly M, et al (2010) The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome research,* **20**, 1297-1303.

Meirmans PG, Van Tienderen PH (2004) GENOTYPE and GENODIVE: two programs for the analysis of genetic diversity of asexual organisms. *Molecular Ecology Notes*, **4**, 792-794.

Mitchell-Olds T, Shaw RG (1987) Regression analysis of natural selection: statistical inference and biological interpretation. *Evolution*, **41**, 1149-1161.

Molles MC, Cahill JF (1999) Ecology: concepts and applications. WCB/McGraw-Hill Dubuque, IA.

Munns R, Tester M (2008) Mechanisms of salinity tolerance. *Annual Reviews on Plant Biology*, **59**, 651-681.

New M, Lister D, Hulme M, Makin I (2002) A high-resolution data set of surface climate over global land areas. *Climate research*, **21**, 1-25.

Penmetsa RV, Cook DR (1997) A legume ethylene-insensitive mutant hyperinfected by its rhizobial symbiont. *Science*, **275**, 527-530.

Penmetsa RV, Uribe P, Anderson J, Lichtenzveig J, Gish JC, Nam YW, Engstrom E, Xu K, Sckisel G, Pereira M, et al (2008) The Medicago truncatula ortholog of Arabidopsis EIN2, sickle, is a negative regulator of symbiotic and pathogenic microbial associations. *The Plant journal,* **55**, 580-595.

Perrier X, Jacquemoud-Collet JP (2006). **DARwin** software http://darwin.cirad.fr/

Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, Handsaker RE, Lunter G, Marth GT, Sherry ST, et al (2011) The variant call format and VCFtools. *Bioinformatics,* **27**, 2156-2158.

Peverill KI, Sparrow LA, Reuter DJ (1999) Soil analysis: an interpretation manual. CSIRO publishing.

Postma E, van Noordwijk AJ (2005) Gene flow maintains a large genetic difference in clutch size at a small spatial scale. *Nature*, **433**, 65-68.

Pritchard JK, Di Rienzo A (2010) Adaptation--not by sweeps alone. *Nature Reviews Genetics*, **11**, 665-667.

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945-959.

R Development Core Team (2008). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Ramaiah M, Jain A, Baldwin JC, Karthikeyan AS, Raghothama KG (2011) Characterization of the phosphate starvation-induced glycerol-3-phosphate permease gene family in Arabidopsis. *Plant physiology,* **157**, 279-291.

Ronfort J, Bataillon T, Santoni S, Delalande M, David JL, Prosperi J-M (2006) Microsatellite diversity and broad scale geographic structure in a model legume: building a set of nested core collection for studying naturally occurring variation in *Medicago truncatula*. *BMC plant biology,* **6**: 28.

Sanford E, Kelly MW (2011) Local adaptation in marine invertebrates. *Annual Review of Marine Science*, **3**, 509-535.

SAS Institute Inc. (2011) SAS/STAT user's guide, version 9.33. SAS Institute, Cary, North Carolina, USA.

Schläppi MR (2006) FRIGIDA LIKE 2 is a functional allele in *Landsberg erecta* and compensates for a nonsense allele of FRIGIDA LIKE 1. *Plant Physiology*, **142**, 1728-1738.

Schmalenbach I, Zhang L, Ryngajllo M, Jiménez-Gómez JM (2014) Functional analysis of the *Landsberg erecta* allele of FRIGIDA. *BMC Plant Biology*, **14**, 218.

Singh M, Kumar J, Singh VP, Prasad SM (2014) Plant tolerance mechanism against salt stress: the nutrient management approach. *Biocheminstry & Pharmacology*, **3**, e165.

Slavich PG, Petterson GH (1993) Estimating the electrical conductivity of saturated paste extracts from 1: 5 soil, water suspensions and texture. *Australian Journal of Soil Research*, **31**, 73-81.

Tang H, Krishnakumar V, Bidwell S, Rosen B, Chan A, Zhou S, Gentzbittel L, Childs KL, Yandell M, Gundlach H, et al (2014) An improved genome release

(version Mt4.0) for the model legume Medicago truncatula. *BMC genomics,* **15**: 312.

Tennessen JA, Akey JM (2011) Parallel adaptive divergence among geographically diverse human populations. *PLoS Genetics*, **7**, e1002127.

Travisano M, Mongold JA, Bennett AF, Lenski RE (1995) Experimental tests of the roles of adaptation, chance, and history in evolution. *Science*, **267**, 87-90.

Turesson G (1922) The species and the variety as ecological units. *Hereditas*, **3**, 100-113.

Turner TL, Bourne EC, Von Wettberg EJ, Hu TT, Nuzhdin SV (2010) Population resequencing reveals local adaptation of *Arabidopsis lyrata* to serpentine soils. *Nature genetics,* **42**, 260-263.

Van der Auwera GA, Carneiro MO, Hartl C, Poplin R, Del Angel G, Levy-Moonshine A, Jordan T, Shakir K, Roazen D, Thibault J, et al (2013) From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. *Current Protocols in Bioinformatics,* **11**, 11.10.1-11.10.33.

Westram AM, Galindo J, Alm Rosenblad M, Grahame JW, Panova M, Butlin RK (2014) Do the same genes underlie parallel phenotypic divergence in different *Littorina saxatilis* populations? *Molecular ecology,* **23**, 4603-4616.

Wilczek AM, Cooper MD, Korves TM, Schmitt J (2014) Lagging adaptation to warming climate in *Arabidopsis thaliana*. *Proceedings of the National Academy of Science United States of America,* **111**, 7906-7913.

Yoder JB, Stanton-Geddes J, Zhou P, Briskine R, Young ND, Tiffin P (2014) Genomic signature of adaptation to climate in *Medicago truncatula*. *Genetics,* **196**, 1263-1275.

Young ND, Debellé F, Oldroyd GE, Geurts R, Cannon SB, Udvardi MK, Benedito VA, Mayer KF, Gouzy J, Schoof H, et al (2011) The Medicago genome provides insight into the evolution of rhizobial symbioses. *Nature,* **480**, 520.

Supplemental material

Genotype information

Table S1. List of genotypes (lines) used in the field experiment and/or used for sequencing, where they were originally collected (country of origin, location, soil salinity, and spatial coordinates for the Portuguese lines), and the nomenclature used when the analysis was done based on site of origin (Field) or STRUCTURE based division of the genotypes into populations (K6).

Origin	Location	Soil Salinity	Line	Spatial Coordinates	Field	K6
Portugal	Castro Marim	Non-Saline	651	N37 14.035 W7 25.969	CMNS	CMNS
Portugal	Castro Marim	Non-Saline	654	N37 14.041 W7 25.959	CMNS	CMNS
Portugal	Castro Marim	Non-Saline	655	N37 14.043 W7 25.958	CMNS	CMNS
Portugal	Castro Marim	Non-Saline	656	N37 14.051 W7 25.953	CMNS	CMNS
Portugal	Castro Marim	Non-Saline	657	N37 14.052 W7 25.954	CMNS	CMNS
Portugal	Castro Marim	Non-Saline	658	N37 14.046 W7 25.959	CMNS	CMNS
Portugal	Castro Marim	Non-Saline	659	N37 14.044 W7 25.961	CMNS	CMNS
Portugal	Castro Marim	Non-Saline	660	N37 14.042 W7 25.964	CMNS	CMNS
Portugal	Castro Marim	Non-Saline	662	N37 14.039 W7 25.971	CMNS	CMNS
Portugal	Castro Marim	Non-Saline	663	N37 14.041 W7 25.971	CMNS	CMNS
Portugal	Castro Marim	Saline	551	N37 14.066 W7 26.494	CMS	CMS
Portugal	Castro Marim	Saline	552	N37 14.065 W7 26.494	CMS	CMS
Portugal	Castro Marim	Saline	553	N37 14.063 W7 26.492	CMS	CMS
Portugal	Castro Marim	Saline	554	N37 14.063 W7 26.488	CMS	CMS
Portugal	Castro Marim	Saline	556	N37 14.059 W7 26.485	CMS	CMNS
Portugal	Castro Marim	Saline	557	N37 14.059 W7 26.486	CMS	CMS
Portugal	Castro Marim	Saline	558	N37 14.061 W7 26.491	CMS	
Portugal	Castro Marim	Saline	559	N37 14.073 W7 26.505	CMS	CMNS
Portugal	Castro Marim	Saline	560	N37 14.073 W7 26.509	CMS	CMNS
Portugal	Castro Marim	Saline	567	N37 14.070 W7 26.528	CMS	CMNS
Portugal	Gilberto	Non-Saline	601	N37 05.979 W7 40.315	GilNS	
Portugal	Gilberto	Non-Saline	602	N37 05.978 W7 40.319	GilNS	GilNSA
Portugal	Gilberto	Non-Saline	603	N37 05.971 W7 40.327	GilNS	GilNSA
Portugal	Gilberto	Non-Saline	604	N37 05.964 W7 40.332	GilNS	GilNSA
Portugal	Gilberto	Non-Saline	605	N37 05.966 W7 40.336	GilNS	GilNSB
Portugal	Gilberto	Non-Saline	607	N37 05.977 W7 40.320	GilNS	GilNSB
Portugal	Gilberto	Non-Saline	609	N37 05.978 W7 40.315	GilNS	GilNSB
Portugal	Gilberto	Non-Saline	611	N37 05.978 W7 40.316	GilNS	GilNSA
Portugal	Gilberto	Non-Saline	612	N37 05.976 W7 40.312	GilNS	GilNSA
Portugal	Gilberto	Non-Saline	614	N37 05.970 W7 40.301	GilNS	GilNSB
Portugal	Gilberto	Non-Saline	615	N37 05.967 W7 40.299	GilNS	GilNSB
Portugal	Gilberto	Saline	501	N37 06.629 W7 39.025	GilS	GilSA
Portugal	Gilberto	Saline	502	N37 06.629 W7 39.025	GilS	GilSA
Portugal	Gilberto	Saline	504	N37 06.629 W7 39.028	GilS	GilSA
Portugal	Gilberto	Saline	506	N37 06.632 W7 39.027	GilS	GilSA

Origin	Location	Soil Salinity	Line	Spatial Coordinates	Field	K6
Portugal	Gilberto	Saline	507	N37 06.633 W7 39.028	GilS	GilSA
Portugal	Gilberto	Saline	508	N37 06.634 W7 39.029	GilS	
Portugal	Gilberto	Saline	510	N37 06.636 W7 39.019	GilS	GilSB
Portugal	Gilberto	Saline	512	N37 06.635 W7 39.021	GilS	GilSB
Portugal	Gilberto	Saline	513	N37 06.634 W7 39.022	GilS	GilSB
Portugal	Gilberto	Saline	515	N37 06.637 W7 39.021	GilS	
Tunisia	El Kef	Non-Saline	TN7.17	Unknown	TN7	TN7
Tunisia	El Kef	Non-Saline	TN7.19	Unknown	TN7	TN7
Tunisia	El Kef	Non-Saline	TN7.22	Unknown	TN7	TN7
Tunisia	El Kef	Non-Saline	TN7.23	Unknown	TN7	TN7
Tunisia	Bulla Regia	Non-Saline	TN9.12	Unknown	TN9	TN9
Tunisia	Bulla Regia	Non-Saline	TN9.15	Unknown	TN9	TN9
Tunisia	Bulla Regia	Non-Saline	TN9.17	Unknown	TN9	TN9
Tunisia	Bulla Regia	Non-Saline	TN9.21	Unknown	TN9	TN9
Tunisia	Enfidha	Saline	TN1.11	Unknown	TN1	TN1
Tunisia	Enfidha	Saline	TN1.15	Unknown	TN1	TN1
Tunisia	Enfidha	Saline	TN1.16	Unknown	TN1	TN1
Tunisia	Enfidha	Saline	TN1.21	Unknown	TN1	TN1
Tunisia	Soliman	Saline	TN8.03	Unknown	TN8	TN8
Tunisia	Soliman	Saline	TN8.04	Unknown	TN8	TN8
Tunisia	Soliman	Saline	TN8.15	Unknown	TN8	TN8
Tunisia	Soliman	Saline	TN8.22	Unknown	TN8	TN8
Tunisia	Soliman	Saline	TN8.25	Unknown	TN8	TN8
Tunisia	Soliman	Saline	TN8.28	Unknown	TN8	TN8
Tunisia	El Kef	Non-Saline	TN7.18	Unknown	Ref	
Unknown	Unknown	Unknown	A17	Unknown	Ref	

Table S2. List of PT genotypes used to measure leaf ion content. Portuguese origin genotypes were ranked based on their relative standardized performance (based both on biomass and reproduction) in saline compared with non-saline sites at each Castro Marim and Gilberto locations (ranked from 1 to 38, with one being the highest performance in saline compared with non-saline sites). The average rank of these four measurements was used to select the five best and five worst performers in salt compared with no salt.

Location	Soil Salinity	Line	Rank	Relative performance
Gilberto	Saline	510	7.00	Best
Gilberto	Saline	504	7.00	Best
Gilberto	Non-Saline	611	7.50	Best
Gilberto	Saline	502	9.25	Best
Gilberto	Saline	512	9.50	Best
Gilberto	Non-Saline	601	29.75	Worst
Castro Marim	Non-Saline	654	29.75	Worst
Castro Marim	Saline	560	30.75	Worst
Castro Marim	Saline	554	32.50	Worst
Castro Marim	Saline	557	33.25	Worst

Soil nutrient, soil water content and electro-conductivity analyses for the field planting sites

Table S3. Two-way fixed effect ANOVA results for soil samples collected from the Portuguese planting sites. F-values, amount of variance explained by the model, and significance for location (CM vs Gil; Loc), salinity (S vs NS; Salt), and location by salinity (e.g., site: CMNS vs CMS vs GilNS vs GilS), together with percent variance explained by location, soil salintiy, location by salinity and within location.

Variable	Location (Loc)	Salinity (Salt)	Loc ×Salt	R2	% Var Loc	% Var Salt	% Var Loc ×Salt	% Var Within Loc
SAR (cmol kg^{-1})$^{0.5}$	0.03	26.39****	0.00	64.1	0.1	62.2	0.0	37.7
ESP (%)	0.00	24.59****	0.02	62.0	0.0	60.6	0.0	39.4
Na (meq/L)	2.16	33.84****	1.28	71.2	4.1	63.5	2.4	30.0
Na (ppm)	8.70**	17.58***	9.25**	72.8	16.9	34.1	18.0	31.0
E.C. (dS/m)	11.83**	14.09**	3.91t	67.6	25.8	30.7	8.5	34.9
Ca (meq/L)	2.56	0.31	1.90	25.5	12.3	1.5	9.1	77.0
Ca (ppm)	99.46****	18.70****	0.00	89.4	74.1	13.9	0.0	11.9
Mg (meq/L)	3.70t	5.44*	0.20	40.3	14.6	21.5	0.8	63.1
Mg (ppm)	33.06****	40.51****	66.99****	91.7	21.1	25.9	42.8	10.2
K (ppm)	109.38****	56.34****	88.75****	95.0	40.4	20.8	32.8	5.9
NO3-N (ppm)	0.76	2.42	0.59	20.0	3.8	12.2	3.0	80.9
Zn (ppm)	1.16	13.23**	0.02	50.4	3.8	43.5	0.1	52.6
Mn (ppm)	63.48****	0.32	5.18*	81.2	74.7	0.4	6.1	18.8
Fe (ppm)	3.04	2.22	17.08***	63.5	7.9	5.8	44.6	41.7
Cu (ppm)	0.10	11.97**	27.84****	76.1	0.2	2.8	61.6	35.4
B (ppm)	35.58****	29.55****	26.93****	87.3	32.9	27.3	24.9	14.8
pH	47.63****	10.11**	2.17	79.4	21.1	21.1	2.9	21.1
Organic Matter (%)	2.14	0.66	2.64	27.6	10.0	3.1	12.3	74.6
Organic Matter ENR (lbs/A)	2.18	0.68	2.78	28.3	10.1	3.1	12.9	73.9
Phosphorus (Weak Bray- (ppm))	3.36t	8.07*	1.82	50.8	11.5	27.6	6.2	54.7
Phosphorus (Olsen Method- (ppm))	0.07	4.52*	3.48t	39.5	0.3	18.8	14.4	66.5
Cation Exchange Capacity (meq/100g)	152.71****	0.03	18.77***	92.3	81.4	0.0	10.0	8.5

t 0.10P>0.05; * P 0.05, ** P < 0.01, *** P < 0.001, **** P < 0.0001

Table S4. Means (±1standard error) for soil samples collected from Portuguese and Tunisian saline and non-saline environments where the original populations were collected.

Variable	PT Non-saline	PT Saline	TN Non-saline	TN Saline
SAR (cmol kg^{-1})$^{0.5}$	0.72 (0.39)	3.36 (0.33)	0.78 (1.38)	10.02 (1.54)
ESP (%)	0.13 (0.52)	3.56 (0.45)	0.26 (1.43)	15.02 (1.61)
Na (meq/L)	1.13 (0.59)	5.68 (0.51)	1.62 (8.76)	46.17 (9.79)
Na (ppm)	21.92 (34.33)	211.42 (29.40)	115.80 (394.91)	1747.00 (441.52)
E.C. (dS/m)	0.45 (0.12)	1.05 (0.10)	0.82 (1.06)	4.97 (1.19)
Ca (meq/L)	3.73 (0.91)	4.39 (0.78)	8.24 (4.76)	14.55 (5.32)
Ca (ppm)	2516.9 (145.4)	1689.2 (124.5)	3186.2 (443.79)	3168.25 (496.17)
Mg (meq/L)	1.32 (0.59)	3.14 (0.50)	1.30 (4.32)	13.60 (4.83)
Mg (ppm)	161.42 (38.62)	485.07 (33.07)	141.40 (143.74)	1009.00 (160.70)
K (ppm)	106.58 (18.91)	293.42 (16.19)	329.40 (99.82)	424.50 (111.60)
NO3-N (ppm)	2.42 (8.50)	19.83 (7.28)	62.00 (14.60)	26.00 (16.32)
Zn (ppm)	0.27 (0.09)	0.72 (0.08)	1.62 (0.79)	2.47 (0.89)
Mn (ppm)	4.00 (0.44)	4.33 (0.38)	3.00 (0.70)	3.50 (0.78)
Fe (ppm)	5.92 (0.90)	7.68 (0.77)	4.80 (1.45)	8.75 (1.62)
Cu (ppm)	0.47 (0.07)	0.80 (0.06)	1.28 (0.41)	1.02 (0.45)
B (ppm)	0.48 (0.55)	4.46 (0.47)	0.82 (0.31)	2.75 (0.35)
pH	7.45 (0.10)	7.86 (0.08)	7.90 (0.10)	8.05 (0.11)
Organic Matter (%)	3.12 (0.47)	2.61 (0.40)	3.02 (0.59)	2.87 (0.66)
Organic Matter ENR (lbs/A)	92.17 (9.30)	82.07 (7.96)	90.20 (11.73)	87.00 (13.11)
Phosphorus (Weak Bray- (ppm))	5.58 (1.63)	11.68 (1.40)	13.80 (5.35)	4.50 (5.99)
Phosphorus (Olsen Method- (ppm))	7.42 (2.12)	13.36 (1.82)	22.40 (4.31)	10.50 (4.82)
Cation Exchange Capacity (meq/100g)	14.27 (0.81)	14.09 (0.69)	18.42 (4.21)	32.77 (4.70)

Table S5. Soil sample number and soil analysis means (±1 standard error) by site where original populations were collected. Portuguese locations Castro Marim (CM) and Gilberto (Gil) saline and non-saline sites, and for Tunisian sites El Kef (TN7), Bulla Regia (TN9), Enfidha (TN1), and Soliman (TN8).

	PT CM NS	PT CM S	PT Gil NS	PT Gil S	TN7 NS	TN9 NS	TN1 S	TN8 S
Soil Sample	3	4	6	6	1	1	1	1
SAR (cmol kg^{-1})$^{0.5}$	0.77 (0.63)	3.40 (0.49)	0.67 (0.45)	3.32 (0.45)	0.3	1	8.3	16.65
ESP (%)	0.10 (0.86)	3.62 (0.66)	0.17 (0.61)	3.50 (0.61)	0.1	0.2	9.9	18.85
Na (meq/L)	1.27 (0.97)	6.70 (0.75)	1.00 (0.69)	4.67 (0.68)	1.0	2.5	18.5	68.2
Na (ppm)	24.0 (56.1)	76.0 (43.4)	19.8 (39.6)	346.8 (39.6)	30	91	625	2536
E.C. (dS/m)	0.57 (0.20)	1.48 (0.15)	0.33 (0.13)	0.62 (0.14)	1.0	1	1.4	7.65
Ca (meq/L)	3.87 (1.48)	6.18 (1.15)	3.60 (1.05)	2.62 (1.05)	14.9	11	6.4	21.55
Ca (ppm)	1562.7 (237.4)	734.8 (183.9)	3471.2 (167.8)	2643.7 (167.8)	4257	4550	3302	2976
Mg (meq/L)	1.90 (0.96)	4.06 (0.75)	0.75 (0.68)	2.22 (0.68)	1.60	1.4	3.4	21.85
Mg (ppm)	223.3 (63.1)	130.8 (48.8)	99.5 (44.6)	839.3 (44.6)	157	158	624	1088
K (ppm)	93.7 (30.9)	46.0 (23.9)	119.5 (21.8)	540.8 (21.8)	329	699	308	407.5
NO3-N (ppm)	3.00 (13.89)	29.00 (10.75)	1.83 (9.82)	10.67 (9.82)	111	48	46	20
Zn (ppm)	0.33 (0.15)	0.80 (0.12)	0.22 (0.11)	0.65 (0.11)	1.0	3.4	0.8	3.8
Mn (ppm)	7.00 (0.73)	6.00 (0.56)	1.00 (0.51)	2.67 (0.51)	3.0	6	3	3.5
Fe (ppm)	7.33 (1.47)	4.20 (1.14)	4.50 (1.04)	11.17 (1.04)	4.0	5	6	11
Cu (ppm)	0.73 (0.12)	0.56 (0.09)	0.20 (0.08)	1.03 (0.08)	0.7	2.9	0.5	0.95
B (ppm)	0.20 (0.91)	0.38 (0.70)	0.77 (0.64)	8.53 (0.64)	0.8	1.2	1.5	2.95
pH	7.10 (0.16)	7.32 (0.12)	7.80 (0.11)	8.40 (0.11)	7.8	7.9	8.3	7.95
Organic M (%)	3.17 (0.76)	1.66 (0.59)	3.07 (0.54)	3.57 (0.54)	3.2	5.3	2.7	2.65
Organic M ENR	93.3 (15.2)	62.8 (11.8)	91.0 (10.7)	101.3 (10.7)	95.0	135	84	82.5
Phosphorus (WB)	9.00 (2.66)	12.20 (2.06)	2.17 (1.88)	11.17 (1.88)	4.0	41	3	5
Phosphorus(O)	9.67 (3.47)	10.40 (2.69)	5.17 (2.45)	16.33 (2.45)	15.0	44	9	12.5
CEC (meq/100g)	10.00 (1.32)	5.22 (1.02)	18.53 (0.93)	22.97 (0.93)	23.5	26.2	25.1	35.85

Table S6. ANOVA results for soil water content (SWC, %) and electro-conductivity (EC, μS) location, soil salinity classification at each site nested within location, and month (SWC: March, April, May; EC: January, March, May).

Source	d.f.	SS	MS	F	P
Soil water content					
Location	1	4.96	4.96	29.73	<0.0001
Salinity (Location)	2	1.17	0.59	3.53	0.0335
Month	2	6.54	3.26	19.61	<0.0001
Location X Month	2	1.90	0.95	5.70	0.0047
Salinity (Location) X Month	4	1.18	0.29	1.77	0.1420
Error	90	15.00	0.17		
Soil EC					
Location	1	0.35	0.35	1.34	0.2505
Salinity (Location)	2	42.67	21.33	81.66	<0.0001
Month	2	0.13	0.06	0.24	0.7856
Location X Month	2	2.41	1.21	4.26	0.0131
Salinity (Location) X Month	4	0.39	0.10	0.37	0.8264
Error	90	18.03	0.26		
Salinity (Population)	**CM NS**	**CM S**	**Gil NS**	**Gil S**	
Mean soil moisture LN (%)	1.56 (0.09)^{c}	1.58 (0.08)^{c}	1.88 (0.08)^{b}	2.17 (0.07)^{a}	
Mean soil EC LN (μ S)	4.40 (0.14)^{d}	6.50 (0.11)^{a}	5.18 (0.12)^{c}	6.00 (0.11)^{b}	

Superscripts that have different letters indicate significant differences in mean trait values ($P < 0.05$).

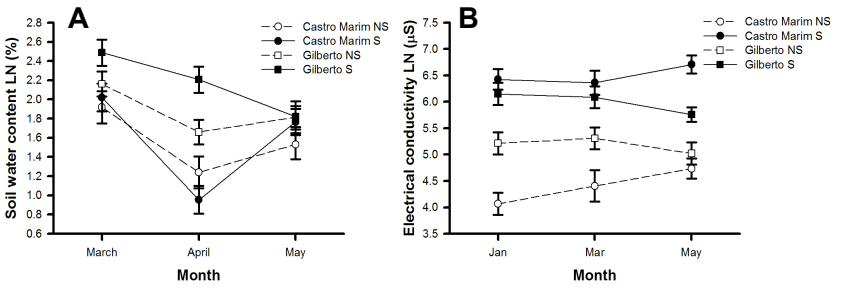


Figure S1. Soil measurements taken for three months on the course of the experiment from each of the four planting sites: location (CM vs Gil), salinity classifications at each site (S vs NS). A: Electro-conductivity (EC) measurements (µS). B: Soil water content.

Genomics and genetics analysis for PT genotypes and comparisons with TN genotype

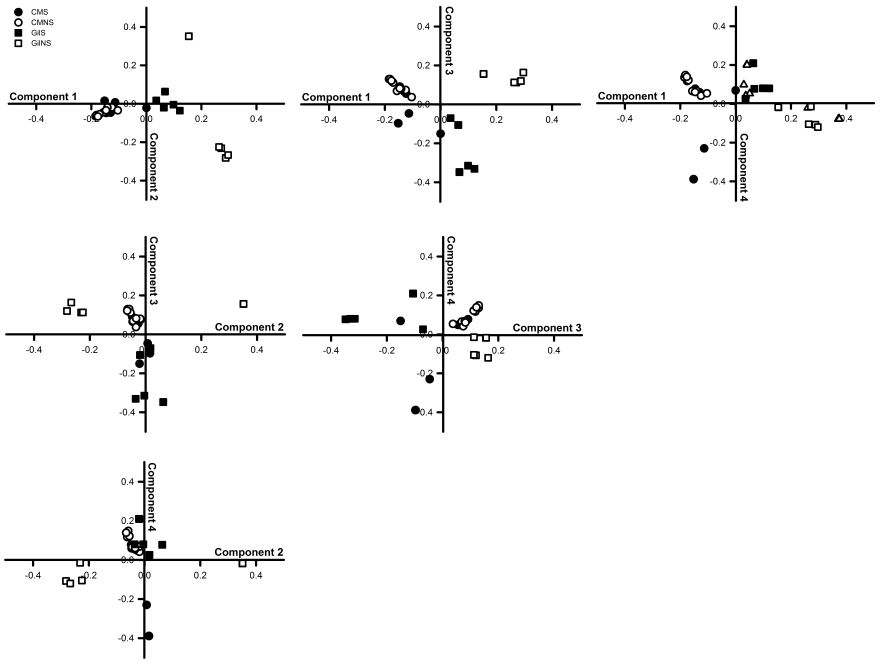


Figure S2: Multidimensional scaling, principal component analysis (PCA), to reveal similarities of genetic variation among PT genotypes with populations assigned based site of origin. Coordinate 1 explains ~5.8% of the variation and separates Gilberto (in the positive axis) from Castro Marim (in the negative axis) origin genotypes. Coordinate 2 explains ~4.9% of the variation and separates Gilberto NS origin genotypes in the two groups distinguished by STRUCTURE: GilNSA genotypes in the positive axis and GilNSB genotypes in the negative axis. Component 3 explains ~ 3.5% of the variation and mostly separates non-saline (positive axis) from saline (negative axis) origin genotypes. Component 4 explains ~3.3% of the variation and separates in the negative axis genotypes from imbalanced soils: the five CMS origin genotypes assigned by STRUCTURE, and closer to zero, all genotypes from GilNS.

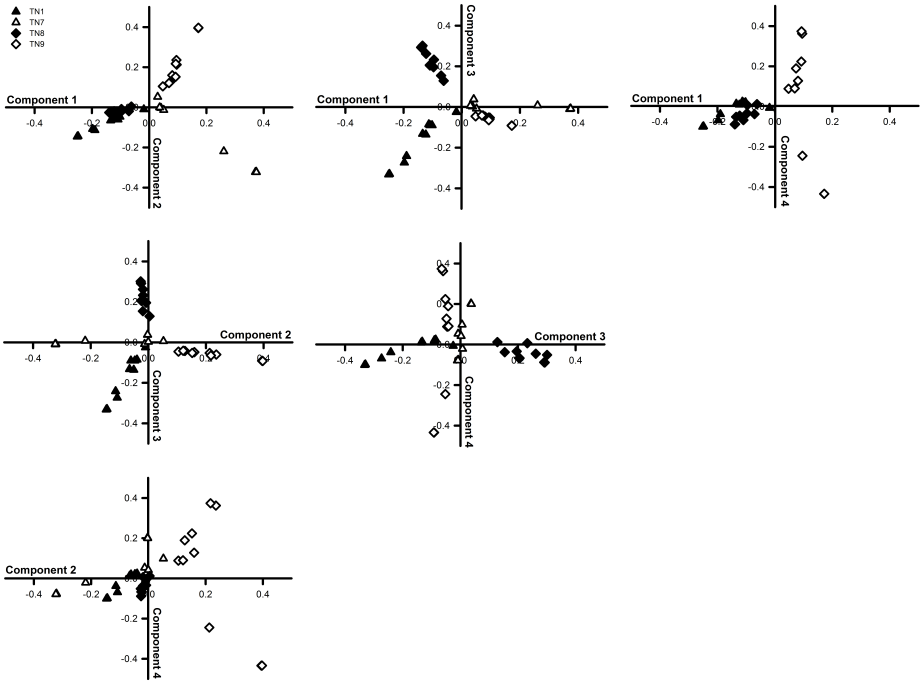


Figure S3: Multidimensional scaling - principal component analysis (PCA) - to reveal similarities of genetic variation among TN genotypes with populations assigned based site of origin. Coordinate 1 explains ~4.5% of the variation and separates non saline (in the positive axis) and saline (in the negative axis) origin genotypes. Coordinate 2 explains ~3.9% of the variation and separates separates TN9 (in the positive axis) from the other TN populations (in the negative axis). Component 3 explains ~ 3.5% of the variation and separates TN8 (in the positive axis) from TN1 (in the negative axis). Component 4 explains ~2.3% of the variation and separates a individuals from TN9 (TN9.05, TN9.12 and TN9.20; in the negative axis) from all other genotypes.

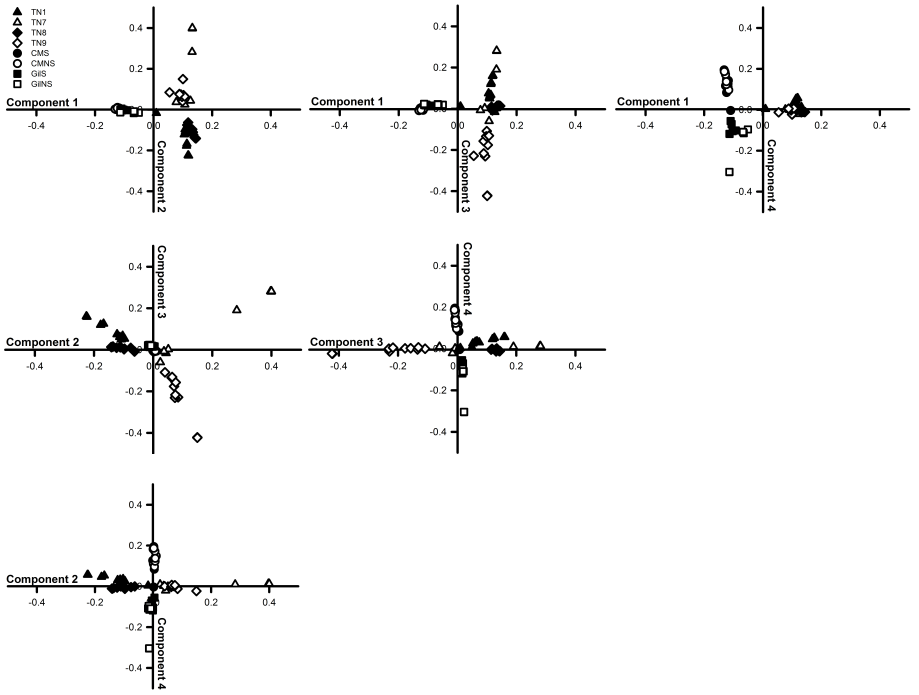


Figure S4: Multidimensional scaling - principal component analysis (PCA) - to reveal similarities of genetic variation among PT and TN genotypes with populations assigned based site of origin. Coordinate 1 explains ~15.4% of the variation and separates Tunisian (positive axis) and Portuguese (negative axis) origin genotypes. Coordinate 2 explains ~4.9% of the variation and separates Tunisian non-saline origin genotypes (positive axis) from Tunisian saline origin genotypes (negative axis). Component 3 explains ~ 4.2% of the variation and mostly separates TN9 (negative axis) from other Tunisian populations (positive axis). Component 4 explains ~3.8% of the variation and separates CM origin genotypes (positive axis) from Gil origin genotypes (negative axis)

Candidate non-synonymous substitutions differentiated between environments

Table S7. List of genes with non-synonymous substitutions differentiated between at least 80% of the genotypes from S and NS environments assuming genetic based population assignment. Gene ID and position are based on Mtr v4.0. RefA refers to the group that contains the reference allele (Ref) rather than the alternative allele (Var). Consecutive genes with the same highlight are in the same haplotype block (in full LD and less than 1 Mbp away).

Gene ID	Annotation	At ortholog	Chr	Position	Ref	Var	Ref	AA
Medtr1g046680	GATA type zinc finger transcription factor	AT3G24050	chr1	17601889	T	C	NS	V/F
Medtr1g040610	Hypothetical protein		chr1	18563194	G	A	NS	N/K
Medtr2g445210	Actin-related protein ARP4	AT3G33520	chr2	19682644	G	A	NS	T/M
Medtr2g449800	Transmembrane protein 2C putative		chr2	21945488	T	C	NS	T/A
			chr2	21945491	T	C	NS	I/V
			chr2	21946768	A	G	NS	M/L
Medtr2g460780	26S proteasome non-ATPase regulatory	AT5G57950	chr2	25071963	C	T	NS	M/L
Medtr2g460790	Fatty-acid desaturase		chr2	25079015	C	T	NS	E/K
Medtr2g460810	Glycerol-3-phosphate transporter	AT2G13100	chr2	25099291	C	T	NS	V/I
Medtr2g460920	Hypothetical protein		chr2	25135197	C	T	NS	E/K

Gene ID	Annotation	At ortholog	Chr	Position	Ref	Var	Ref	AA
			chr2	25135320	A	G	NS	Y/H
			chr2	25135326	C	T	NS	A/T
			chr2	25136530	G	A	NS	A/V
Medtr2g461020	Zinc finger 2C C3HC4 type (RING finger)	AT2G15580	chr2	25185710	A	C	NS	M/L
Medtr2g461120	Hypothetical protein		chr2	25208986	C	T	NS	R/H
Medtr2g461240	Class II knotted-like homeobox protein	AT5G25220	chr2	25277330	C	T	NS	A/T
Medtr2g461440	Aspartyl protease		chr2	25374109	G	A	NS	R/C
Medtr2g082010	Terpene synthase family 2C metal-binding domain protein		chr2	34459566	G	T	NS	C/*
			chr2	34461852	T	G	NS	E/D
Medtr3g062890	NADH-quinone oxidoreductase cyanobacterial	AT5G58260	chr3	28437819	T	G	S	W/G
Medtr3g073400	Hypothetical protein		chr3	33095241	C	A	NS	P/H
			chr3	33095288	G	A	NS	V/I
Medtr4g053785	Gland-specific fatty acyl-CoA reductase	AT4G33790	chr4	19478941	A	C	NS	K/Q
			chr4	19483081	A	T	NS	K/I
			chr4	19483163	C	A	NS	S/R
Medtr4g079650	Brdt subfamily bromodomain	AT1G06230	chr4	30811252	C	G	NS	P/R
Medtr4g080580	Zinc ion-binding protein	AT4G13970	chr4	31222735	C	T	NS	E/K
Medtr4g115930	RING-H2 zinc finger protein		chr4	47906118	G	A	NS	R/C
Medtr4g119428	Late embryogenesis abundant protein	AT3G11650	chr4	49494511	C	A	NS	S/Y
Medtr5g016370	Transducin family protein/WD-40 repeat protein	AT5G67320	chr5	5874020	A	G	S	M/L
			chr5	5874065	A	G	S	M/L
			chr5	5874067	G	GT	S	
Medtr5g017020	Histone acetyltransferase of the CBP family	AT3G12980	chr5	6163801	A	G	S	M/L
Medtr5g017025	Hypothetical protein		chr5	6164414	T	C	S	T/A
Medtr5g037180	RNA recognition motif, RRM 2C RBD protein		chr5	16225017	G	A	NS	V/I
Medtr5g037400	TPR-like protein	AT1G10430	chr5	16381798	A	AT	NS	
			chr5	16381063	G	T	NS	*/L
Medtr5g037560	Hypothetical protein		chr5	16439546	G	A	NS	V/I
Medtr5g037680	F-box protein interaction domain protein	AT3G06240	chr5	16487094	T	A	NS	N/Y
Medtr5g037880	Potyviral capsid protein interacting protein		chr5	16557747	C	T	S	M/L
Medtr5g038060	Peptide/nitrate transporter	AT1G69870	chr5	16629566	T	A	NS	M/L
			chr5	16626724	A	C	S	V/G
Medtr5g048910	Hypothetical protein		chr5	21379446	T	A	NS	Q/H
			chr5	21379531	T	A	NS	K/M
Medtr5g089220	Agenet domain protein		chr5	38782632	C	T	NS	T/I
Medtr6g034265	ABC transporter family protein		chr6	11664527	T	C	NS	S/P
Medtr6g080160	Pentatricopeptide (PPR) repeat protein		chr6	30135307	A	G	S	K/R
Medtr7g007080	F-box/RNI/FBD-like domain protein	AT1G51370	chr7	1333525	C	T	S	R/K
Medtr7g033135	Cysteine-rich receptor-kinase-like protein	AT5G40380	chr7	11723314	T	C	S	E/G
Medtr7g053500	Glucan endo-1 2C3-beta-glucosidase-like protein	AT3G07320	chr7	18857630	C	T	S	M/L
			chr7	18857634	T	C	S	M/L
Medtr7g072070	PPR containing plant-like protein	AT2G38420	chr7	26781059	A	T	NS	M/K
Medtr8g466990	Hypothetical protein		chr8	23951298	C	T	NS	R/Q

Table S8. List of genes with non-synonymous substitutions fully differentiated between genotypes from CM and Gil locations. Gene ID and position based on Mtr v4.0. RefA refers to the group that contains the reference allele (Ref) rather than the alternative allele (Var). Consecutive genes with the same highlight are in the same haplotype block (in full LD and less than 1 Mbp away).

Gene ID	Annotation	At homolog	Chr	Pos	Ref	Var	Ref	AA
Medtr1g011840	Boron transporter-like protein	AT3G06450	chr1	2236082	C	T	CM	A/P
Medtr1g027160	Stress-induced receptor-like kinase		chr1	8964143	C	T	CM	A/V
			chr1	8964147	T	C	CM	T/A
Medtr1g044020	Rpp4C3	AT4G16860	chr1	16469961	G	T	Gil	C/W
Medtr1g051085	Hypothetical protein		chr1	19917255	T	G	Gil	A/V
Medtr2g040230	Disease resistance protein (TIR-NBS-LRR class)_2C putative		chr2	17653999	A	G	Gil	D/G
			chr2	17654350	A	T	Gil	N/I
			chr2	17654699	A	C	Gil	L/F

Gene ID	Annotation	At homolog	Chr	Pos	Ref	Var	Ref	AA
Medtr2g040240	Hypothetical protein		chr2	17657040	G	A	CM	R/C
Medtr2g040850	Glycoside hydrolase family 18 protein		chr2	17895153	G	A	CM	E/K
Medtr2g048360	DUF630 family protein	AT2G34670	chr2	21301447	C	T	CM	
Medtr2g048870	Eukaryotic translation initiation factor 3c	AT3G56150	chr2	21627711	T	C	CM	I/T
			chr2	21627713	C	A	CM	Q/K
Medtr2g049310	PPR containing plant-like protein		chr2	21729897	G	A	CM	P/L
			chr2	21730930	C	T	CM	A/T
Medtr2g072150	Disease resistance protein (CC-NBS-LRR class)		chr2	30279017	A	T	CM	K/N
Medtr2g101950	F-box/RNI superfamily protein_2C putative	AT1G16930	chr2	43866887	T	C	Gil	N/D
Medtr2g461780	GRF zinc finger protein		chr2	25559421	C	T	CM	S/L
Medtr2g461900	Hypothetical protein		chr2	25599849	T	C	CM	W/R
Medtr3g023310	MORN domain protein		chr3	7045830	C	T	Gil	M/I
Medtr3g023630	F-box/RNI/FBD-like domain protein		chr3	7175713	T	G	CM	K/T
Medtr3g029930	60S ribosomal L7-like protein		chr3	9438472	A	G	CM	*/W
Medtr3g048690	NB-ARC domain disease resistance protein		chr3	18079710	C	T	CM	R/H
			chr3	18082214	A	C	CM	I/S
			chr3	18082235	G	T	CM	T/K
			chr3	18082422	A	T	CM	Y/N
Medtr3g048710	PIF1-like helicase		chr3	18095681	T	G	Gil	F/L
Medtr3g451900	Hypothetical protein		chr3	18877461	G	A	CM	R/*
Medtr3g451910	Hypothetical protein		chr3	18880531	T	A	CM	H/L
Medtr3g088160	Thylakoid-bound ascorbate peroxidase	AT1G77490	chr3	39980072	G	C	CM	G/A
Medtr3g088380	Replication factor C subunit 2	AT1G21690	chr3	40162237	T	G	CM	M/L
			chr3	40162279	A	C	CM	M/V
			chr3	40157503	T	A	CM	*/Y
Medtr3g088495	DnaJ domain protein		chr3	40223781	G	A	CM	A/V
Medtr3g088560	BTB/POZ domain plant protein		chr3	40077580	G	A	CM	
Medtr3g088575	NAD/NADH kinase family protein	AT1G21640	chr3	40303708	T	C	CM	H/R
			chr3	40304981	C	T	CM	G/R
Medtr3g088580	F-box only protein	AT1G21760	chr3	40068320	T	G	CM	M/V
Medtr3g088625	Histone-lysine N-methyltransferase ASHH2_2C putative	AT1G77300	chr3	40336035	G	A	CM	E/K
			chr3	40347242	C	G	CM	R/G
			chr3	40347522	G	A	CM	S/N
			chr3	40333484	A	C	CM	
Medtr4g417260	Verticillium wilt resistance-like protein		chr4	5348840	T	G	Gil	K/Q
Medtr4g417270	Verticillium wilt disease resistance protein		chr4	5354517	C	T	Gil	T/I
Medtr4g087520	O-acetylserine (thiol) lyase	AT2G43750	chr4	34368822	A	G	Gil	M/L
Medtr4g087620	MAP kinase-like Ntf4 protein	AT2G43790	chr4	34418802	C	G	CM	T/S
Medtr4g087635	Hypothetical protein		chr4	34427028	T	C	Gil	I/T
Medtr4g087690	Ferrochelatase	AT2G30390	chr4	34454031	A	T	CM	M/L
Medtr4g087830	Phospholipase A1		chr4	34558372	C	G	CM	G/A
Medtr4g087920	Helix loop helix DNA-binding domain protein		chr4	34478057	G	A	Gil	G/D
Medtr4g087960	Hypothetical protein		chr4	34493901	T	G	Gil	V/G
Medtr4g088030	Trichome birefringence-like protein	AT2G34070	chr4	34505564	G	C	CM	L/F
Medtr5g018570	Wall associated kinase-like protein	AT2G23450	chr5	6931470	T	C	CM	M/L
			chr5	6931499	T	A	CM	
Medtr5g029940	Hypothetical protein		chr5	12598119	T	C	CM	N/D
Medtr5g029950	Disease resistance protein (TIR-NBS-LRR class)		chr5	12599075	T	A	CM	N/I
			chr5	12599376	A	G	CM	Y/H
Medtr5g430520	Hypothetical protein		chr5	12856433	A	G	Gil	L/S
			chr5	12856637	T	G	Gil	K/Q
Medtr5g062110	Hypothetical protein		chr5	25798892	T	C	CM	L/S
Medtr5g073460	Exocyst subunit exo70 family protein		chr5	31277395	GT	G	CM	
Medtr6g014890	UDP-glucose 6-dehydrogenase		chr6	4819737	C	CA	Gil	
			chr6	4819601	A	G	Gil	S/P
			chr6	4819752	A	C	Gil	N/K
Medtr6g025730	Alpha-L-arabinofuranosidase-like protein		chr6	8946712	A	G	CM	V/A
			chr6	8947268	G	A	Gil	P/S
Medtr6g025790	Alpha-L-arabinofuranosidase-like protein		chr6	8973029	G	T	Gil	S/I

Gene ID	Annotation	At homolog	Chr	Pos	Ref	Var	Ref	AA
Medtr6g028050	Calcium-dependent lipid-binding (CaLB domain)	AT5G17980	chr6	9937730	T	G	CM	H/Q
Medtr6g046450	NB-ARC domain disease resistance protein		chr6	16787550	C	A	Gil	D/Y
Medtr6g047650	Hypothetical protein		chr6	17167236	T	C	Gil	C/R
Medtr6g053260	F-box and associated interaction domain protein		chr6	19126091	C	T	CM	T/I
Medtr6g060630	Hypothetical protein		chr6	21006192	T	C	Gil	V/A
			chr6	23901952	C	T	Gil	E/K
			chr6	23901956	C	A	Gil	R/S
Medtr6g465880	Hypothetical protein		chr6	23521401	C	G	CM	A/P
Medtr7g110730	Hypothetical protein		chr7	45377196	C	T	CM	P/S
Medtr8g006780	Auxin efflux carrier family protein		chr8	957044	G	C	CM	M/V
Medtr8g007285	Equilibrative nucleoside transporter 6	AT4G05140	chr8	1427847	A	G	CM	I/V
Medtr8g007340	Elongator complex protein	AT3G11220	chr8	1468158	T	C	CM	Q/R
Medtr8g007435	Transmembrane protein_2C putative		chr8	1538196	A	T	CM	F/I
Medtr8g008540	Actin-related protein ARP4		chr8	1757626	T	C	CM	E/G
Medtr8g008920	Hypothetical protein		chr8	1945953	C	T	Gil	P/S
Medtr8g008970	UDP-glucosyltransferase family protein		chr8	1963324	G	A	Gil	E/K
Medtr8g009020	Helix loop helix DNA-binding domain protein		chr8	1994937	T	C	CM	M/V
Medtr8g009063	UDP-glucosyltransferase family protein		chr8	2011982	A	G	CM	Y/H
Medtr8g009170	1-aminocyclopropane-1-carboxylate oxidase-like		chr8	2055892	C	G	CM	P/A
Medtr8g010160	ABI3-interacting protein	AT5G48385	chr8	2585964	C	T	CM	D/N
Medtr8g011410	Cysteine-rich RLK (receptor-like kinase) protein		chr8	3057515	G	A	Gil	G/D
Medtr8g012655	Ethylene response factor	AT4G13040	chr8	3726275	T	C	CM	*/R
Medtr8g012880	Hypothetical protein		chr8	3896683	C	A	CM	E/*
Medtr8g015170	Receptor-like kinase		chr8	4898610	G	A	CM	T/I
			chr8	4898703	A	G	CM	L/S
Medtr8g015190	LRR receptor-like kinase plant		chr8	4914979	C	A	CM	
Medtr8g015610	Occludin-like domain protein		chr8	5130706	T	C	CM	R/G
			chr8	5131489	C	G	CM	E/Q
			chr8	5132023	G	A	CM	P/S
			chr8	5132229	A	G	CM	I/T
			chr8	5132463	A	T	CM	V/E
			chr8	5132807	T	A	CM	K/N
			chr8	5133028	T	C	CM	M/V
			chr8	5134128	G	A	CM	A/V
			chr8	5134142	A	T	CM	H/Q
			chr8	5134159	T	G	CM	T/P
			chr8	5136900	T	C	CM	Q/R
Medtr8g015620	Dentin sialophosphoprotein-like protein	AT3G21290	chr8	5142530	T	C	CM	Q/R
			chr8	5142869	A	G	CM	L/S
			chr8	5143229	C	T	CM	R/Q
			chr8	5143545	T	G	CM	N/H
Medtr8g015670	Hypothetical protein		chr8	5170584	A	G	CM	D/G
Medtr8g015680	F-box/FBD-like domain protein		chr8	5174349	G	C	CM	A/P
			chr8	5174418	C	G	CM	Q/E
			chr8	5175433	G	C	CM	Q/H
			chr8	5175459	T	A	CM	F/Y
Medtr8g015920	Hypothetical protein		chr8	5252784	G	A	CM	W/*
Medtr8g015970	ABC transporter-like family-protein		chr8	5281138	T	C	CM	I/V
			chr8	5281685	T	C	CM	T/A
			chr8	5283985	C	T	CM	A/T
Medtr8g016020	ABC transporter-like family-protein		chr8	5312691	C	T	CM	T/I
			chr8	5316869	A	C	CM	L/F
Medtr8g016150	PRA1 family protein		chr8	5396340	A	G	CM	
Medtr8g016270	Endomembrane protein 70 family protein	AT3G13772	chr8	5471362	T	G	CM	K/Q
Medtr8g040860	2-hydroxyisoflavanone dehydratase		chr8	15234059	A	C	CM	L/V
Medtr8g041670	Receptor-like kinase		chr8	15704672	G	C	CM	H/Q
Medtr8g042440	Disease resistance protein (TIR-NBS-LRR class)		chr8	16353806	G	A	Gil	H/Y
			chr8	16353996	T	G	Gil	E/D
Medtr8g064230	E3 ubiquitin-protein ligase BRE1-like protein		chr8	26928882	C	A	CM	A/S

Gene ID	Annotation	At homolog	Chr	Pos	Ref	Var	Ref	AA
			chr8	26928886	C	A	CM	E/D
Medtr8g069750	DUF239 domain protein		chr8	29383419	TG	T	CM	
			chr8	29383262	A	C	CM	Y/S
			chr8	29384025	G	A	CM	

Performance analysis divided by country of origin

Table S9. Results from generalized linear mixed-models on components of plant performance (germination: germ, survival to reproduction: surv, number of pods: pods) and cumulative estimates of performance (fitness, plant biomass) divided by country of origin. For PT, population is represented by type of soil origin (SoilOrg) nested within location of origin (Loc). For TN, population (Pop) is nested within origin type of soil (SoilOrg). Planting site is represented by planting location (PltLoc) nested within planting soil type (PltSoil). Genotype nested within population was treated as a random effect.

Source	**df**	Germin. X^2	Survival X^2	Pods F-value	Biomass F-value	Fitness F-value
PT only						
Loc	1	5.87*	1.38	22.64****	1.70	4.26*
SoilOrg(Loc)	2	6.56*	16.22***	2.02	1.86	3.32
Pltsite	1	3.24t	0.13	6.08*	7.24**	2.90t
PltSoil(PltLoc)	2	10.18**	53.73****	526.66****	121.69****	267.84****
Loc X PltLoc	1	0.00	0.34	1.69	0.38	0.01
Loc X PltSoil(PltLoc)	2	0.50	1.10	10.75**	1.23	8.57*
SoilOrg(Loc) X PltLoc	2	3.10	1.65	6.34*	1.89	14.52***
SoilOrg(Loc) X PltSoil(PltLoc)	4	7.32	10.03*	6.46	0.92	7.15
Genotype(Loc SoilOrg)	34			44.22	0.0	53.86*
TN only						
SoilOrg	1	0.89	2.78t	2.17	0.75	0.24
Pop (SoilOrg)	2	2.54	6.81*	2.33	0.82	0.34
PltLoc	1	0.98	14.87***	7.29**	36.24****	0.92
PltSoil(PltLoc)	2	10.54**	50.16****	423.47****	138.82****	276.32****
SoilOrg X PltLoc	1	0.75	0.08	3.56t	7.99**	5.69*
SoilOrg X PltSoil(PltLoc)	2	6.09*	2.56	1.51	0.44	3.72
Pop(SoilOrg)xPltLoc	2	0.44	3.54	0.12	0.15	0.31
Pop(SoilOrg)xPltSoil(PltLoc)	4	1.07	5.96	5.24	0.25	9.31t
Genotype(Pop)	15			40.50***	2.9t	45.69****

t $P < 0.10$, * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, **** $P < 0.0001$

Sodium and potassium content of Tunisian genotypes growing in the greenhouse under 0 and 100 mM NaCl.

Table S10. Results from mixed-model ANOVA on sodium (Na), potassium (K), and sodium-potassium ratio (Na/K) tissue ion content from TN genotypes grown under 0mM and 100mM NaCl. Population (Pop) and salinity treatment (Trt) were treated as fixed effects with F-values reported, and genotype nested within population (Gen) was treated as a random effect with Chi-values reported.

	df	Na F-value	K F-value	Na/K F-value
Population	3, 33	2.21	0.24	1.78
Salinity treatment	1, 79	27.68****	0.19	20.34****
Population X Salinity treatment	3, 79	1.32	0.20	1.09
Genotype(Population)	1	41.6****	10.3**	26.8****

t 0.10>P>0.05; * P ≤ 0.05, ** P < 0.01, *** P < 0.001, **** P < 0.0001

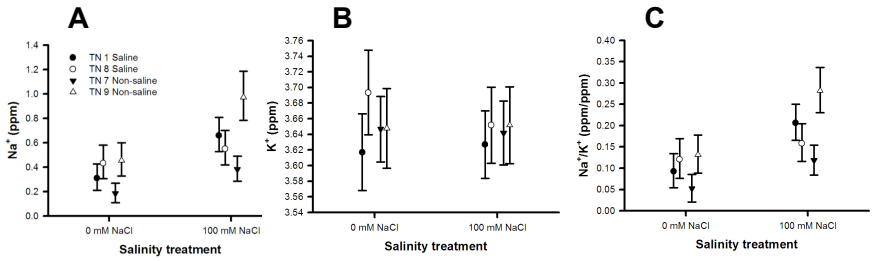


Figure S5: Least square means and standard error of Na, K and Na/K leaf ion content of a selection of TN genotypes used in the reciprocal transplant field experiment in Portugal grown in the greenhouse under 0mM and 100mM NaCl. A. Natural logarithm transformed Na leaf ion content. B. Squared root transformed K leaf ion content. C. Na/K leaf ion content.

Chapter 5.

Conclusions and future perspectives

Soil salinity is a major factor reducing crop productivity worldwide, a situation that is expected to aggravate with the continuation of intensive agricultural practices and the consequences of climate change (Flowers and Yeo 1995; Provin and Pitt 2001; IPCC 2007; FAO 2008, 2013; Nelson et al. 2014). Cultivated species have undergone narrow bottlenecks with domestication and subsequent artificial selection and continuous inbreeding. The resulting reduced genetic diversity is reflected in the decrease of adaptive potential, which is expected to have negative consequences on the size and distribution patterns of natural populations and crop yield within agricultural systems. For this dissertation, ecological, evolutionary, and molecular methods were used to disentangle the mechanisms and genetic basis of adaptation to soil salinity within natural populations, an approach that is increasingly favored to develop new strategies for crop improvement (Galloway and Fenster 2000; Hall et al. 2010; Turner et al. 2010; Friesen and von Wettberg 2010; Fournier-Level et al. 2011; Ingvarssom and Street 2011; Warschefsky et al. 2014; Milla et al. 2015). Such studies come at a critical time in human history as climate change is rapidly shifting the biogeographical range of natural populations and may help identify mechanisms that allow populations to cope with changes in the environment (Etterson and Shaw 2001; Charmantier et al. 2008; Atkins and Travis 2010).

Legumes are among the most important resources for food and feed worldwide, combining about one third of human total nitrogen intake (Graham and Vance 2003; Gepts et al. 2005). *Medicago truncatula* is used as model species and allows the study of legume specific processes (Cook 1999). *M. truncatula* is native from the Mediterranean basin where it occurs in a wide range of environments and a broad range of genetic, developmental and physiological material has been made available over the years.

For the work presented in this dissertation, Tunisian and Portuguese *M. truncatula* populations that have evolved in saline and non-saline habitats were used to: i) evaluate local adaptation and potential maladaptation to soil salinity (Chapters 2 and 4); ii) study the mechanisms of salinity resistance evolved in these

populations (Chapter 2 to 4); iii) test whether populations from different countries evolved the same or different mechanisms to deal with soil salinity (Chapter 4); iv) identify potential genes and traits underlying adaptive responses (Chapter 2 to 4); and v) test evolutionary hypotheses regarding salinity adaptation within these populations (Chapter 2 to 4).

Natural populations from saline and non-saline soils in northern Tunisia were genotyped and searched for genome-wide signatures of salinity-dependent differentiation, which revealed candidate regions and genes for salinity adaptation (Chapter 2; Friesen et al. 2010). To investigate the natural variation in Tunisian saline and non-saline *M. truncatula*, an existing Affymetrix microarray was used, and a genome-wide polymorphism scan, linkage disequilibrium (LD) estimates, and patterns of population differentiation were obtained. Probe hybridization was analyzed using a new algorithm to call single feature polymorphisms (SFPs) that were validated using traditional Sanger sequencing. A total of 52 SFPs were perfectly differentiated between saline and non-saline origin populations, clustering into 18 genome regions in Mt3.0 that contain candidate genes for local adaptation to high-salinity habitats, including several genes with putative roles in abiotic stress responses. The initial genotyping of Tunisian *M. truncatula* revealed salinity associated genetic differentiation that includes a large number of candidate genes for salinity responses (Chapter 2). The findings in Chapter 2 support that these populations make a good study system to test evolutionary question of local adaptation and pursue the genes and traits underlying adaptive responses to salinity.

Following up on the population genomics characterization of Tunisian *M. truncatula* (Friesen et al. 2010) more detailed evolutionary ecology studies were developed on these populations, including whole genome sequencing and extensive phenotyping (Friesen et al. 2014). This work involved reciprocal soil field and greenhouse experiments in Tunisia, and a greenhouse experiment in the USA. Phenotyping revealed the importance of both salt tolerance and salt avoidance for salinity adaptation in Tunisian genotypes (Friesen et al. 2014). Additional candidate

genes with possible roles in early flowering time and abiotic stress tolerance were discovered (Friesen et al. 2014). For example, the Medtr3g098090.1 codes for a CPK that has three homologs in *Arabidopsis thaliana* related with ABA mediated germination and early seedling growth, drought, and salinity stress responses (Zhao et al. 2011).

Germination and initial root expansion represent the first developmental stages exposed to the soil environment and are important for salinity adaptation, particularly because they are among the most sensitive to water deficit and high salt (Chapter 3; Cordeiro et al. 2014). Population differentiation on germination and early root growth between saline and non-saline adapted Tunisian *M. truncatula* depending on salinity and ABA was studied using a combination of plate and aeroponical experiments (Cordeiro et al. 2014). Saline origin genotypes displayed greater salinity tolerance for germination and seedling traits relative to non-saline origin genotypes. Maintaining greater primary root elongation rates and lateral branching under salt may allow saline adapted plants to prospect for relatively lower salinity micro-patches and effectively maintaining biomass allocation to roots improving water relations (Munns and Tester 2008; Galvan-Ampudia and Testerink 2011; Rahnama et al. 2011). Although the general responses of root growth due to ABA and NaCl were similar and seemed to be additive, saline population are less sensitive to NaCl and more sensitive to ABA compared to non-saline populations. Altogether, this suggests that saline origin populations may have evolved an ABA-independent pathway to shape root architecture under high salt, which could potentially allow responding to water deficit separately from ion toxicity and adjusting root growth differentially when experiencing each stress. Ion balance and salinity tolerance mechanisms seem to be critical during early seedling establishment, supporting that different mechanisms of salinity adaptation are important at different developmental stages. Interestingly, the CPK candidate gene mentioned earlier is involved in ABA mediated signaling, potentially during seedling establishment (Friesen et al. 2010, 2014). Identifying allelic variation for this gene

and screening for early seedling traits may help to find other resistant genotypes and have positive implications on research for saline resistant crops.

Parental environmental effects of salinity on germination were tested, and parental exposure to salt had a general negative effect on germination (Cordeiro et al. 2014). Moreover, population specific differences were observed for the effects of salinity on time to germination and for the impact of parental environment on germination rates (Cordeiro et al. 2014). Besides germination, parental environmental effects due to salinity were tested in adult plants of the Tunisian *M. truncatula* populations (Moriuchi et al. 2016). Of particular note, parental environmental exposure to salt accentuated fitness differences between saline and non-saline origin genotypes and was associated with salinity avoidance traits, while tolerance traits were associated with offspring exposure to salinity (Moriuchi et al. 2016). On an evolutionary perspective, parental exposure to salt forecasts and queues improved performance of the offspring under this stress. Transgenerational plasticity is transmitted to the offspring by mechanisms such as epigenetic modifications or differential seed provisioning and is commonly interpreted as noise, but in predictable environments parental experience is more likely to anticipate the offspring's environment and function as a favorable strategy. Therefore, for crop species that share such transgenerational effects, farmers would benefit from raising the stash of seed under similar salinity to the one they will experience.

Generally, Tunisian populations tended to grow better under soil salinity more similar to their home soil salinity levels. Tunisian *M. truncatula* plants from saline soils evolved avoidance mechanisms to deal with high salinity, including the constitutive earlier flowering that allows plants to terminate their life cycle before the seasonal salinity built up in the soil as the rain cease, the soil dries, and salt percolates towards the root zone (Friesen et al. 2014). Additionally, the capacity to maintain development, growth and physiology of saline origin genotypes under high salinity - i.e, germination rates, root elongation and carbon assimilation - is associated with salinity tolerance mechanisms (Cordeiro et al. 2014, Moriuchi et al.

2016). Salinity associated genome-wide differentiation enabled naming several candidate genes to underlie the expression and regulation of these avoidance and tolerance related responses, such as a CPK for ABA mediated salinity stress regulation and CONSTANS for the constitutive flowering time differences (Friesen et al. 2010, 2014). But salinity adaptation of Tunisian *M. truncatula* has fitness tradeoffs in non-saline environments, where earlier flowering and lower growth potential translate into a cost on performance (Friesen et al. 2014; Moriuchi et al. 2016). Such fitness tradeoffs suggest that using salinity resistant germplasm may only be beneficial when growing them in saline environments. In other words, adaptation is environment specific, which was expected. But this also raises the question of whether genes and traits that confer salinity adaptation to Tunisian *M. truncatula* populations also provide fitness advantages under different saline soil environments, or if it is specific to soils from these populations.

To test whether salinity responses involve the same or different mechanisms in distinct populations, saline and non-saline genotypes were collected in southern Portugal and used in a reciprocal transplant field experiment in Portugal together with an informative subset of Tunisian genotypes to test; i) whether Tunisian populations are adapted to Portuguese saline environments; ii) if Portuguese saline and non-saline populations are adapted to local soil salinity levels; and iii) whether isolated populations that evolved under saline conditions share the same or different mechanism to deal with salinity (Chapter 4). The whole genome of the Portuguese genotypes was sequenced and compared with Tunisian genotypes. Salinity is not the only environmental factor driving performance at the Portuguese planting sites. But the parallel responses and traits under selection at Castro Marim (CM), with signal for salinity adaptation in Tunisian and Portuguese saline genotypes, is evidence that Portuguese saline-origin populations are adapted to salinity. Also, viability selection seems to be operational in both Tunisian and Portuguese saline environments. Population's geographic and genetic differentiation between countries suggests the independent evolution for salinity adaptation in Tunisian and Portuguese genotypes. This is additionally

supported by the lack of common candidate genes between this analysis and previous analysis of Tunisian genotypes (Friesen et al. 2014). But the same traits seem to be under selection. Therefore, Portuguese and Tunisian populations evolved independently, with selection acting on different genetic components in spite of the signal for parallel adaptive evolution (Cordeiro et al. submitted).

By studying natural populations of *M. truncatula*, novel mechanisms of salinity adaptation and potential genes that confer salinity adaptation (e.g., concordant role of tolerance and avoidance mechanisms) were identified, supporting the independent evolution of distant populations. Commonly, studies on environmental stress focus on a single mechanism, but incorporating a range of mechanisms has shown to be important to understand adaptive evolution. Furthermore, the mechanisms of salinity adaptation revealed to be dependent upon the stage of plant development. Altogether, the results from this dissertation confirm that salinity adaptation is a result of complex trait interactions, which can be disentangled via the integration of ecological, evolutionary, and molecular methods. The independent evolution of salinity adaptation suggests that studying natural populations may lead to the identification of novel mechanisms and their genetic basis in populations that have evolved in saline habitats. Such discoveries should enable the identification of candidate genes in natural populations for integration into crops towards obtaining improved yields in marginal agricultural fields. Moreover, studying salinity adaptation in multiple populations may lead to the identification of multiple candidate gene networks best-suited particular associations of salinity with other environmental factors. Such evolutionary studies constitute a more integrated approach to guide crop improvement and allow for more rigorous ecological and evolutionary tests.

Future perspectives

Additional experimental work is necessary to confirm the genes and traits that underlie salinity adaptation in these populations. The large amount of genomic and phenotypic data made available from the studies presented in this dissertation

make a great basis for more directed molecular and evolutionary studies. Several examples of potential follow up studies come to mind given a preferred target gene, trait, or even mechanism. In this case, phenotyping in controlled conditions is key to disentangle the response to the specific environmental variable. Classical genetic studies can be conducted with the available germplasm by performing controlled crosses, to quantitatively test questions of heredity of the traits and taking advantage of genetic markers from the whole genome sequencing data to link to the genetic basis of the target. Moreover, “proof of concept” can be achieved using the available mutant collections coupled with transgenic approaches where the target alleles can be directly tested or be used for complementation studies. Taking a more evolutionary ecology approach, further collections could be made from the same and from additional isolated populations. Genetic characterization of the novel genotypes would inform about which individuals contributed the most to the gene pool at each site already collected, and whether other distant populations are genetically differentiated. Multiple reciprocal transplant field experiments could be conducted at the different origins using a subset of genotypes that span the genetic variation in the collection to depict the traits favored by selection at each site, i.e., how distinct sets of environmental variables other than salinity affect plant performance in saline habitats. Additionally, by doing it for consecutive years, temporal environmental variation and ecological predictability could be assessed and the genotypes that are overall more successful at each site identified. For example, salinity is not the major factor driving performance at the Gilberto (Gil) location, but populations from the saline site (GilS) showed salinity adaptation at Castro Marim (CM), meaning that very likely salinity has been a major selective factor during past (and maybe future) growing seasons (Chapter 4). Expanding the temporal and spatial dimensions in evolutionary ecology studies could give a more integrated view of salinity adaptation, and aid selecting the best candidate mechanism to introduce in cultivated germplasm depending on the target region and their combine environmental conditions.

Given the importance of population level adaptation and the role of past and current environmental effects on the study of salinity adaptation in *M. truncatula*, another source of genetic variation may also play a role in salinity adaptation. While it has been known for some time that *M. truncatula* and *M. littoralis* hybridize (Lesins and Lesins 1979), the contribution of gene flow between species to population level adaptation is still unknown. During the course of the Portuguese field experiment, two natural hybrid populations between *M. truncatula* and *M. littoralis* were identified, collected and studied. Preliminary data were generated in parallel with the work developed for this dissertation and is presented here as the foundation for future studies.

Hybridization consists on the reproduction between members of genetically distinct groups such as distinct but closely related species, and is among the most common sources of variation in sexually reproducing organisms, particularly in plants (Ellstrand and Schierenbeck, 2000; Mallet 2005; Soltis and Soltis 2009; Abbott et al. 2013; Schumer et al. 2014). Hybridization has the potential to generate novel genetic and phenotypic variation from the recombination of alleles from different species, and it is still unclear what drives hybrid taxa to persist rather than form transiently. If hybrids are just temporary, they serve as a mere genomic sink; but if hybrids are more persistent, then hybridization may lead to novel adaptations capable of altering species breadth, and ultimately result in speciation (Arnold 1992; Abbott 1992; Rieseberg 1997; Ellstrand and Schierenbeck 2000; Barton et al. 2001; Mallet 2005; Buerkle and Rieseberg 2008; Soltis and Soltis 2008; Abbot et al. 2013; Seehausen 2013; Schumer et al. 2014; Warschefsky et al. 2014).

M. truncatula and *M. littoralis* are locally adapted to different habitats but natural hybrid populations can be found in highly disturbed areas in steep clines of soil texture and salinity (Supplementary material; Lesins and Lesins, 1979; Small 2011). Because of the autogamic nature of these species, reduced gene flow can allow the maintenance of favorable gene combinations in hybrids, potentiating the adaptive evolution of hybrids that express novel genetic (Figure S1) and

phenotypic variation (Figure S2) relative to the parents and may therefore colonize other environmental ranges. Systems that allow hybridization may translate into faster evolution rates while allowing for further exploration of the mechanisms that permit hybridization, as well as its barriers or potential benefits.

Most commonly, new traits arise with hybridization from the reorganization of genetic interactions (Rieseberg et al. 1999). Transgressive segregation and newly generated traits likely allow some hybrids to colonize and adapt to new habitats (Abbot 1992, Barton 2001, Rieseberg 2003, Rieseberg et al. 2007, Turner et al. 2010, Gagnaire et al. 2012, Fishman et al. 2015). This system is ideal to study the genetic patterns of hybridization success and to test hypotheses on the contribution of hybridization to adaptive evolution.

Combining whole genome sequencing with RAD sequencing genotyping allows determining the direction, frequency, and stability of hybridization events, and associate them with disturbance and other environmental variables (Coop et al. 2010). LD breakdown is expected if hybrids are stable, while low recombination and the evolution towards one of the parent's genome is expected if hybrids are transient. Reciprocal transplant experiments can be used to test for local adaptation and foreign advantage as both allopatric *versus* sympatric, and home *versus* away performance and identify the main environmental factors driving performance, following existing approaches (Kawecki and Ebert 2004; Hereford 2009; Blanquart et al 2013). Artificial hybrids can be developed using controlled crosses and RAD genotyping of viable and unviable crosses allows studying genome evolution by looking at favorable and unfavorable gene combinations, respectively. Hybridization may be disadvantageous, i.e., hybrids may not be able to grow and reproduce effectively, and be maintained by recurrent events, which could potentially make them a sink for biotic factors. Another possibility is that hybrids are intermediate, meaning trait expression would be the average between the two parent species, which doesn't necessarily mean that this is ideal in intermediate environments. Also, transgressive segregation may lead to the expression of traits outside the range of the parent species, and is therefore more

likely to enable the evolution of novel adaptations. Preliminary data on Portuguese hybrid populations suggests that the observed patterns are dependent on the trait, that hybridization tends to be relatively stable (aberrant phenotypes occur regularly), and that new useful traits may arise from such crosses. Overall, such experiments can help reveal the genetic, phenotypic, and ecological consequences of hybridization, and the set of genes underlying complex adaptive traits using a model system for hybridization.

Plant domestication is usually accompanied by reproductive isolation and, consequently, speciation. *Medicago* can be used as a model system to study the use of hybridization in crop species. This is particularly important because findings on hybridization are facilitating the introduction of genes and traits identified in the wild relatives in cultivated germplasm (Friesen and von Wettberg 2010; Warschefsky et al. 2014). Thus, novel adaptations could be identified and introduced into crops potentiating the development of crop species with a broader ecological and/or agricultural range (Etterson and Shaw 2001; Friesen and von Wettberg 2010; Nelson et al. 2014; Rippke et al. 2016).

References

Abbott R, Albach D, Ansell S, Arntzen JW, Baird SJE, Bierne N, Boughman J, Brelsford A, Buerkle CA, Buggs R (2013) Hybridization and speciation. *Journal of evolutionary biology,* **26**, 229-246.

Abbott R, Albach D, Ansell S, Arntzen JW, Baird SJE, Bierne N, Boughman J, Brelsford A, Buerkle CA, Buggs R (2013) Hybridization and speciation. *Journal of evolutionary biology,* **26**, 229-246.

Arnold ML (1992) Natural hybridization as an evolutionary process. *Annual review of Ecology and Systematics,* , 237-261.

Atkins KE, Travis JMJ (2010) Local adaptation and the evolution of species' ranges under climate change. *Journal of Theoretical Biology*, **266**, 449-457.

Barton NH (2001) The role of hybridization in evolution. *Molecular ecology,* **10**, 551-568.

Blanquart F, Kaltz O, Nuismer SL, Gandon S (2013) A practical guide to measuring local adaptation. *Ecology letters,* **16**, 1195-1205.

Bonnin I, Ronfort J, Wozniak F, Olivieri I (2001) Spatial effects and rare outcrossing events in Medicago truncatula (Fabaceae). *Molecular ecology,* **10**, 1371-1383.

Buerkle CA, Rieseberg LH (2008) The rate of genome stabilization in homoploid hybrid species. *Evolution; international journal of organic evolution,* **62**, 266-275.

Charmantier A, McCleery RH, Cole LR, Perrins C, Kruuk LE, Sheldon BC (2008) Adaptive phenotypic plasticity in response to climate change in a wild bird population. *Science,* **320**, 800-803.

Cook DR (1999) Medicago truncatul: a model in the making! *Current opinion in plant biology,* **2**, 301-304.

Coop G, Witonsky D, Di Rienzo A, Pritchard JK (2010) Using environmental

correlations to identify loci underlying local adaptation. *Genetics,* **185**, 1411-1423.

Cordeiro MA, Moriuchi KS, Fotinos TD, Miller KE, Nuzhdin SV, von Wettberg EJ, Cook DR (2014) Population differentiation for germination and early seedling root growth traits under saline conditions in the annual legume Medicago truncatula (Fabaceae). *American journal of botany,* **101**, 488-498.

Ellstrand NC, Schierenbeck KA (2000) Hybridization as a stimulus for the evolution of invasiveness in plants? *Proceedings of the National Academy of Sciences of the United States of America,* **97**, 7043-7050.

Elshire RJ, Glaubitz JC, Sun Q, Poland JA, Kawamoto K, Buckler ES, Mitchell SE (2011) A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PloS one,* **6**: e19379.

Etterson JR, Shaw RG (2001) Constraint to adaptive evolution in response to global warming. *Science*, **294**, 151-154.

FAO (2008) FAO Land and Plant Nutrition Management Service. Available at http://www.fao.org/ag/agl/agll/spush.

FAO *FAOSTAT* (2013), Available at http://faostat.fao.org.

Fishman L, Beardsley PM, Stathos A, Williams CF, Hill JP (2015) The genetic architecture of traits associated with the evolution of self-pollination in Mimulus. *The New phytologist,* **205**, 907-917.

Flowers, T.J., and A. R. Yeo. 1995. Breeding for salinity resistance in crop plants: where next? Australian Journal of Plant Physiology 22: 875–884.

Fournier-Level A, Korte A, Cooper MD, Nordborg M, Schmitt J, Wilczek AM (2011) A map of local adaptation in *Arabidopsis thaliana*. *Science*, **334**, 86-89.

Friesen ML, von Wettberg EJ (2010) Adapting genomics to study the evolution and ecology of agricultural systems. *Current opinion in plant biology,* **13**, 119-125.

Friesen ML, Cordeiro MA, Penmetsa RV, Badri M, Huguet T, Aouani ME, Cook DR, Nuzhdin SV (2010) Population genomic analysis of Tunisian *Medicago*

truncatula reveals candidates for local adaptation. *The Plant journal,* **63**, 623-635.

Friesen ML, von Wettberg EJ, Badri M, Moriuchi KS, Barhoumi F, Chang PL, Cuellar-Ortiz S, Cordeiro MA, Vu WT, Arraouadi S, et al (2014) The ecological genomic basis of salinity adaptation in Tunisian *Medicago truncatula*. *BMC genomics,* **15**: 1160.

Galloway LF, Fenster CB (2000) Population differentiation in an annual legume: local adaptation. *Evolution*, **54**, 1173-1181.

Galvan-Ampudia CS, Testerink C (2011) Salt stress signals shape the plant root. *Current Opinion in Plant Biology*, **14**, 296–302.

Gepts P, Beavis WD, Brummer EC, Shoemaker RC, Stalker HT, Weeden NF, Young ND (2005) Legumes as a model plant family. Genomics for food and feed report of the cross-legume advances through genomics conference. *Plant physiology,* **137**, 1228-1235.

Graham PH, Vance CP (2003) Legumes: importance and constraints to greater use. *Plant Physiology*, **131**, 872-877.

Hall BG, Ehrlich GD, Hu FZ (2010) Pan-genome analysis provides much higher strain typing resolution than multi-locus sequence typing. *Microbiology*, **156**, 1060-1068.

Hereford J (2009) A quantitative survey of local adaptation and fitness trade-offs. *American Naturalist*, **173**, 579-588.

Ingvarsson PK, Street NR (2011) Association genetics of complex traits in plants. *New Phytologist*, **189**, 909-922.

IPCC (2007) Climate change 2007: the physical science basis. Agenda, **6**, 333.

Kawecki TJ, Ebert D (2004) Conceptual issues in local adaptation. *Ecology Letters*, **7**, 1225-1241.

Lesins KA, Lesins I (1979) *Genus Medicago (Leguminosae): A Taxogenetic Study.* W. Junk.

Mallet J (2005) Hybridization as an invasion of the genome. *Trends in ecology &*

evolution, **20**, 229-237.

Milla R, Osborne CP, Turcotte MM, Violle C (2015) Plant domestication through an ecological lens. *Trends in Ecology and Evolution*, **30**, 463-469.

Miller MR, Dunham JP, Amores A, Cresko WA, Johnson EA (2007) Rapid and cost-effective polymorphism identification and genotyping using restriction site associated DNA (RAD) markers. *Genome research,* **17**, 240-248.

Moriuchi KS, Friesen ML, Cordeiro MA, Badri M, Vu WT, Main BJ, Aouani ME, Nuzhdin SV, Strauss SY, von Wettberg EJ (2016) Salinity Adaptation and the Contribution of Parental Environmental Effects in Medicago truncatula. *PloS one,* **11**: e0150350.

Munns R, Tester M (2008) Mechanisms of salinity tolerance. *Annual Review of Plant Biology*, **59**, 651–681.

Nelson GC, Valin H, Sands RD, Havlík P, Ahammad H, et al (2014) Climate change effects on agriculture: economic responses to biophysical shocks. *Proceedings of the National Academy of Sciences of the U S A,* **111**, 3274-3279.

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics,* **155**, 945-959.

Provin T, Pitt JL (2001) Managing soil salinity. Texas AgriLife Extension Service publication E-60. Texas A&M Univ. Publication. College Station, TX.

Rahnama A, Munns R, Poustini K, Watt M (2011) A screening method to identify genetic variation in root growth response to a salinity gradient. *Journal of experimental botany,* **62**, 69-77.

Rieseberg LH, Raymond O, Rosenthal DM, Lai Z, Livingstone K, Nakazato T, Durphy JL, Schwarzbach AE, Donovan LA, Lexer C (2003) Major ecological transitions in wild sunflowers facilitated by hybridization. *Science (New York, N.Y.),* **301**, 1211-1216.

Rieseberg LH, Kim SC, Randell RA, Whitney KD, Gross BL, Lexer C, Clay K

(2007) Hybridization and the colonization of novel habitats by annual sunflowers. *Genetica,* **129**, 149-165.

Rieseberg LH (1997) Hybrid origins of plant species. *Annual Review of Ecology and Systematics,* **28**, 359-389.

Rieseberg LH, Archer MA, Wayne RK (1999) Transgressive segregation, adaptation and speciation. *Heredity,* **83**, 363-372.

Rippke U, Ramirez-Villegas J, Jarvis A, Vermeulen SJ, Parker L, Mer F, Diekkrüger B, Challinor AJ, Howden M (2016) Timescales of transformational climate change adaptation in sub-Saharan African agriculture. *Nature Climate Change,* .

Schumer M, Cui R, Powell DL, Dresner R, Rosenthal GG, Andolfatto P (2014) High-resolution mapping reveals hundreds of genetic incompatibilities in hybridizing fish species. *eLife,* **3**.

Seehausen O (2013) Conditions when hybridization might predispose populations for adaptive radiation. *Journal of evolutionary biology,* **26**, 279-281.

Small E (2011) Alfalfa and relatives. *Evolution and classification of Medicago*.

Soltis PS, Soltis DE (2009) The role of hybridization in plant speciation. *Annual review of plant biology,* **60**, 561-588.

Turner TL, Bourne EC, Von Wettberg EJ, Hu TT, Nuzhdin SV (2010) Population resequencing reveals local adaptation of *Arabidopsis lyrata* to serpentine soils. *Nature Genetics*, **42**, 260-263.

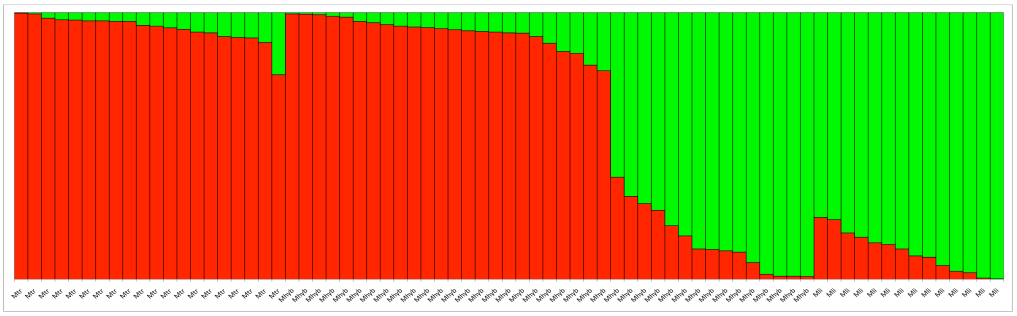
Warschefsky E, Penmetsa RV, Cook DR, von Wettberg EJ (2014) Back to the wilds: tapping evolutionary adaptations for resilient crops through systematic hybridization with crop wild relatives. *American journal of botany*, **101**, 1791-1800.

Supplemental material

Preliminary results from analysis of hybrid zones

M. truncatula and *M. littoralis* are closely related species that produce natural homoploid hybrids. Putative hybrids occur in steep clines of soil texture and salinity between co-occurring pure species populations: “pure” *M. truncatula* occurs in heavy non-saline soils and *M. littoralis* in sandy more saline soils. Pod morphology is a good character for delimiting these species because *M. truncatula* has sutures in the coils, the spines are thicker in the bottom third, and the pods are hairy, while *M. littoralis* has smooth coils, it has no or homogenous spines, and is glabrous (Lesins and Lesins, 1979; Small 2011). Pods that contain intermediate combinations of these morphologies are considered hybrids.

Figure S1. STRUCTURE (Pritchard et al. 2000) based population assignment for K=2, using 1114 SNPs from RAD-tag sequencing in a subset of lines from the hybrid zones (10 Mtr from each site; 21 Mhyb from FU, 18 Mhyb from MR; and 7 Mli from each site), sorted by assigned species (based on pod morphology) and assigned population.



Several populations of pure *M. truncatula* and *M. littoralis* connected by apparent hybrid zones were identified in southern Portugal. Pure and hybrid pods were identified based on pod morphology in two areas: Fuzeta (FU) and Manta Rota (MR). Genotyping of the hybrid zones was done using restriction-site associated DNA sequencing (RAD; Miller et al. 2007, Elshire et al. 2011) on 72 genotypes, including ‘pure’ species and putative hybrids, which resulted in more than 20k polymorphic loci. These diagnostic SNPs were used to distinguish *M. truncatula* from *M. littoralis*. Allele-frequency based analysis using STRUCTURE

(Pritchard et al. 2000) shows genetic differentiation between *M. truncatula* and *M. littoralis* as well as different levels of admixture in the putative hybrids (Figure 1).

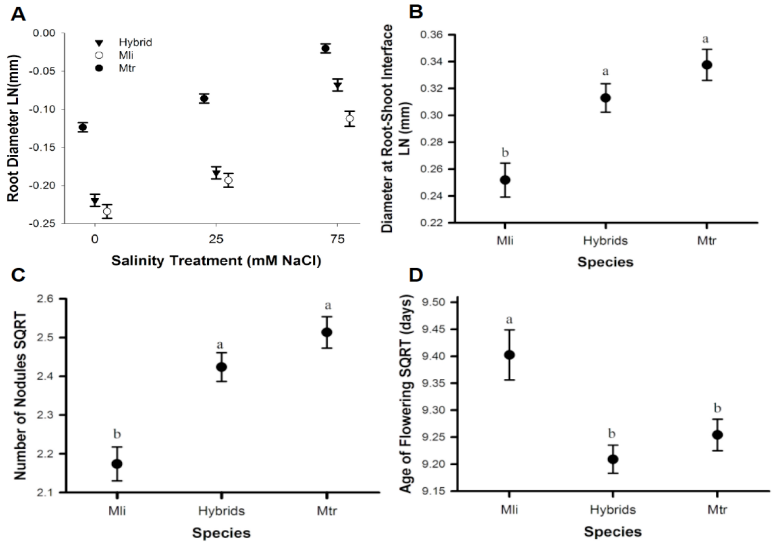


Figure S2. Least square means and standard error of the Mli and Mtr and putative hybrid phenotyping in a growth chamber (aeroponics chamber with 3 levels of NaCl – 0, 25 and 75 mM – during 14 days using 5 replicates of 10 Mtr, 4 Mli and 5 putative hybrid accessions) and in the field (reciprocal transplant field experiment in two hybrid zones divided in Mli, Mtr and Mhyb zone – FU and MR – using 6 replicates of 20 Mtr, 21 Mli and 29 putative hybrids). A. Average root diameter of growth chamber grown seedlings under salinity treatments; B: Average root-shoot interface diameter of field grown plants; C. Nodule number of field grown plants; D. Time to flowering in field grown plants.

Moreover, root architecture (Figs. 2A-B), nodulation (Fig 2C) and flowering time (Fig. 2D) are differentiated between morpho-species assignment based on pod morphology. Putative hybrid roots tend to be thicker (Fig 2A) and have lower root growth rates at increasing salt concentrations compared with the 'pure' parent species. As further explored in Chapter 3, such patterns of root growth might be advantageous under drought and saline conditions and therefore give hybrid vigor under other unfavorable soil water balances. In a reciprocal transplant field experiment, putative hybrids and *M. truncatula* showed similar trends regarding root growth and architecture (Fig 2B-D). *M. littoralis* had thinner roots (Fig 2B), fewer nodules (Fig 2C) and later flowering (Fig 2D) than *M. truncatula* and putative hybrids.

Genetic data suggests that back-crosses occur between hybrids and parental species, most likely with *M. truncatula*. Additionally, these observed patterns suggest that hybrid phenotypes can be environment dependent (Fig. 2A), intermediate between parental species (Fig. 2B,C), or similar to one parental species (Fig. 2D).