RESEARCH ARTICLE

Transcriptomic response to nitrogen availability reveals signatures of adaptive plasticity during tetraploid wheat domestication

Alice Pieri[1], Romina Beleggia[2]*, Tania Gioia[3], Hao Tong[4,5], Valerio Di Vittori[1], Giulia Frascarelli[1], Elena Bitocchi[1], Laura Nanni[1], Elisa Bellucci[1], Fabio Fiorani[6], Nicola Pecchioni[2], Stefania Marzario[3], Concetta De Quattro[7], Antonina Rita Limongi[7], Pasquale De Vita[2], Marzia Rossato[7], Ulrich Schurr[6], Jacques L. David[8], Zoran Nikoloski[4,5], Roberto Papa[1]*

[1] Department of Agricultural, Food and Environmental Sciences, Marche Polytechnic University, via Brecce Bianche, 60131, Ancona, Italy
[2] Council for Agricultural Research and Economics (CREA), Research Centre for Cereal and Industrial Crops (CREA-CI), 71122 Foggia, Italy
[3] School of Agricultural, Forestry, Food and Environmental Sciences, University of Basilicata, 85100 Potenza, Italy
[4] Bioinformatics Department, Institute of Biochemistry and Biology, University of Potsdam, 14476 Potsdam, Germany
[5] Systems Biology and Mathematical Modeling Group, Max Planck Institute of Molecular Plant Physiology, 14476 Potsdam, Germany
[6] Institute of Biosciences and Geosciences (IBG-2): Plant Sciences, Forschungszentrum Julich GmbH, 52428 Julich, Germany
[7] Department of Biotechnology, University of Verona, Strada Le Grazie 15, 37134 Verona, Italy
[8] AGAP, Univ. Montpellier, CIRAD, INRAE, Institut Agro, 34060 Montpellier, France

* Corresponding authors: Roberto Papa: r.papa@univpm.it; Romina Beleggia: romina.beleggia@crea.gov.it

Short title: Role of nitrogen in wheat domestication

The author responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (https://academic.oup.com/plcell/pages/General-Instructions) is: Roberto Papa (r.papa@univpm.it).

Abstract

The domestication of crops, coupled with agroecosystem development, is associated with major environmental changes and provides an ideal model of phenotypic plasticity. Here, we examined 32 genotypes of three tetraploid wheat (*Triticum turgidum* L.) subspecies, wild emmer, emmer and durum wheat, which are representative of the key stages in the domestication of tetraploid wheat. We developed a pipeline that integrates RNA-Seq data and population genomics to assess gene expression plasticity and identify selection signatures under diverse nitrogen availability conditions. Our analysis revealed differing gene expression responses to nitrogen availability across primary (wild emmer to emmer) and secondary (emmer to durum wheat) domestication. Notably, nitrogen triggered the expression of twice as many genes in durum wheat compared to that in emmer and wild emmer. Unique selection signatures were identified at each stage: primary domestication mainly influenced genes related to biotic interactions, whereas secondary domestication affected genes related to amino acid metabolism, in particular lysine. Selection signatures were found in differentially expressed genes, notably those associated with nitrogen metabolism, such as the gene encoding glutamate dehydrogenase. Overall, our study highlights the pivotal role of nitrogen availability in the domestication and adaptive responses of a major food crop, with varying effects across different traits and growth conditions.

Introduction

Domestication influences the genetic diversity of animals and plants as they adapt to agroecosystems and undergo selection to meet human preferences and needs. This process is typically associated with the genome-wide loss of nucleotide diversity due to the combined consequences of selection and genetic drift, which is known as the domestication bottleneck. The loss of genetic diversity has been documented in many domesticated species by comparing them with wild relatives (Bitocchi et al., 2017). A parallel effect is the reprogramming of gene expression and the loss of expression diversity, which was first reported in the common bean (*Phaseolus vulgaris*) (Bellucci et al., 2014) and subsequently in other domesticated plants and animals (Sauvage et al., 2017; Liu et al., 2019; Burgarella et al., 2021). Similar observations have been reported at the level of metabolic diversity (Beleggia et al., 2016).

Changes in nucleotide and gene expression diversity during the domestication of tetraploid wheat (*Triticum turgidum* L., 2*n* = 4*x* = 28; AABB genome) are not fully understood. Tetraploid wheat was domesticated in two well-defined phases. Primary domestication from wild emmer (*Triticum turgidum* ssp. *dicoccoides*) to emmer (*Triticum turgidum* ssp. *dicoccum*) started ~12,000 years ago in the Fertile Crescent. This was followed by secondary domestication from emmer to durum wheat (*Triticum turgidum* ssp. *durum*), which started 8,000~10,000 years ago in the Near East and gave rise to durum wheat, the most important form of tetraploid wheat and currently the most widespread Mediterranean crop (Gioia et al., 2015; Taranto et al., 2020; Levy and Feldman, 2022).

The transition from wild environments to early farming and eventually to modern high-input agroecosystems had profound ecological consequences. Throughout history, humans have employed various methods to enhance soil fertility, such as: soil preparation to facilitate organic matter mineralization and the use of livestock manure, with evidence dating back to Neolithic early farming sites ~7900 years ago (Bogaard et al., 2013). However, the scale and intensity of fertilizer use have escalated over time, especially with the advent of the Haber-Bosch industrial process, which heavily relies on non-renewable fossil fuels. Notably, the widespread overreliance on nitrogen (N) fertilizers in modern industrial agriculture can be traced back to the Donald model (Donald, 1968). This model aims to optimize crop yields by

minimizing intraspecific competition and providing substantial agronomic inputs, including fertilizers (Fréville et al., 2022). Today, the global application of N fertilizers, particularly to cereal crops, exceeds 80 million tons annually (Ludemann et al., 2022). N is an essential macronutrient whose availability is directly linked to crop yield and grain quality (protein content) (Barneix, 2007; Howarth et al., 2008; Laidò et al., 2013), but it is also directly harmful to humans and the environment. Indeed, excess N from agricultural sources is one of the major fresh water pollutants, causing the eutrophication of aquatic ecosystems (Rockström et al., 2009). The production of industrial fertilizers contributes ~3% of global CO_2 and is a primary source of N_2O (Wood and Cowie, 2004). Understanding genetic variations in N acquisition, assimilation and metabolism can therefore provide strategies for crop improvement to meet the United Nations Sustainable Development Goals (SDGs) (Plett et al., 2018; Hawkesford and Griffiths, 2019).

Environmental changes that accompanied the domestication of crops over thousands of years can be tolerated by organisms that exhibit phenotypic plasticity, defined as the ability of a genotype to exhibit changes in a specific trait across different environments, and through the modulation of gene expression (Bradshaw, 1965; Laitinen and Nikoloski, 2019). Understanding the molecular basis of phenotypic plasticity in crops and their wild relatives can help to address the challenges faced by modern agriculture. In tetraploid wheat, phenotypic differences in below-ground and above-ground growth traits related to N availability primarily arose during secondary domestication (Gioia et al., 2015), but the relationship between N metabolism and changes in gene expression plasticity during domestication is unclear.

Here, we analyzed 32 wild emmer, emmer and durum wheat genotypes by RNA-Seq in contrasting N availability scenarios, to investigate the potential role of N during the domestication of tetraploid wheat. Our study elucidates the subspecies-specific responses to nitrogen, as well as nucleotide and gene expression diversity during both primary and secondary domestication phases. Our results provide insight into the pivotal role of N during the domestication and adaptive plasticity of one of our major food crops.

Results and Discussion

Choice of the reference genome: mapping accuracy and comparative analysis across available tetraploid wheat references

The inclusion of genotypes from three distinct tetraploid wheat subspecies (i.e. wild emmer, emmer, and durum wheat) within our panel presented challenges when comparing results across groups. In this context, the choice of the reference genome was of crucial importance. Given that reference genomes were only available for two of the three subspecies examined in our study, namely wild emmer and durum wheat, we considered using the A and B subgenomes of bread wheat (*Triticum aestivum*) as an outgroup reference closely related to all subspecies in the panel. This option aimed to mitigate potential biases that might arise from favoring any subspecies, ensuring a balanced comparison across the genotypes. Indeed, the A and B subgenomes of bread wheat serve as a good intermediate between emmer (wild and domesticated) and durum wheat. Bread wheat shares ancestral A and B subgenomes with the wild and cultivated emmer, originating from the same founding population. On the other hand, the A and B subgenomes of durum wheat underwent differentiation coinciding with the origin of bread wheat (Haudry et al., 2007; Levy and Feldman, 2022).

However, we rigorously tested these assumptions and validated our choice by also mapping reads to currently available tetraploid wheat reference genomes: wild emmer Zavitan (Zhu et al., 2019) and durum wheat Svevo (Maccaferri et al., 2019).

We prepared 128 RNA-Seq libraries from the 4-week-old leaves of 32 tetraploid wheat genotypes representing wild emmer, emmer and durum wheat, grown in two contrasting N conditions (N starvation and optimal N availability) (Supplementary Data Set S1). The mapping frequency across the entire genome was consistent among the three references (86–87%) with an average of 6.8 million mapped reads per genotype (Supplementary Data Set S1 A-C). However, while the proportion of reads mapping to genic regions was similar when comparing bread wheat and wild emmer wheat (average 73%), it was lower when using the durum wheat reference genome (average 52%) primarily due to the absence of untranslated regions (UTRs) in the Svevo reference annotation (Supplementary Data Set S1 A-C).

Pairwise genetic distances between the three reference genomes were computed using an alignment-free method based on the MinHash technique (Ondov et al., 2016). This method compresses large genomic sequences (the three entire genomes, in our case) into sketch representations, allowing for rapid similarity estimations with bounded error. We found that the Mash distance (*D*), which is an approximation of the mutation rate (Ondov et al., 2016), between the bread wheat Chinese Spring A and B subgenomes was ~ 0.014 (*P* value < 10^{-10}) when compared with wild emmer Zavitan, and ~ 0.008 (*P* value < 10^{-10}) when compared with durum wheat Svevo. The distance between durum wheat Svevo and wild emmer Zavitan was the same as that between bread wheat and wild emmer Zavitan (*D* ~ 0.014, *P* value < 10^{-10}). From the resulting distances, the average nucleotide identity (ANI) can be extracted as *D* ~ 1 – ANI (Ondov et al., 2016), obtaining a value for ANI of ~ 99% between each genome, thus confirming their close relationship.

Computing raw read counts and filtering out genes with weak expression (see Materials and Methods) resulted in 32,358 genes from the bread wheat Chinese Spring reference genome. We performed the same process using the wild emmer Zavitan and durum wheat Svevo reference genomes, resulting in 33,586 and 29,784 genes, respectively. A comprehensive comparison of the sequences of the three gene sets also showed an ANI exceeding 98%.

Overall, despite minor differences in the outcomes across different references, the overall patterns remain consistent and the use of an outgroup species for reference might help mitigate biases and ensure fair representation of all subspecies, maintaining mapping accuracy and coverage of gene regions, making it a suitable choice for our analysis.

A greater loss of nucleotide diversity occurred during the secondary domestication of tetraploid wheat

Variant calling on the RNA-Seq of the whole panel of 32 genotypes produced 800,996 high-quality single-nucleotide polymorphisms (SNPs) including “population-specific” SNPs found only in one of the three subspecies. The number of polymorphic sites was similar in wild emmer (617,128) and emmer (613,509), but was much lower in durum wheat (425,513). Site frequency spectra for each

subpopulation are provided in Supplementary Figure S1. SNPs were categorized as "private" if found exclusively in a single subspecies or "shared" if they were distributed across two or three subspecies. We identified 190,377 common SNPs shared by all three *taxa*. As expected, wild emmer and emmer shared the highest percentage of SNPs (33%, 206,578). In contrast, durum wheat shared only 11% (46,352) of its SNPs with wild emmer and 17% (71,147) with emmer (Supplementary Data Set S2).

SNP principal component analysis (PCA) revealed the broad genetic structure of the three wheat *taxa* (Figure 1) and confirmed that secondary domestication had a greater impact than primary domestication in differentiating the durum wheat subspecies. The 12 durum wheat genotypes are genetically very similar, forming a dense cluster that is clearly distinguishable from the wild emmer and emmer genotypes. In contrast, the wild emmer and emmer genotypes are loosely clustered, indicating a greater genetic admixture. These results are consistent with previous genetic studies on the origins of domesticated tetraploid wheat and reflect the multiple stages of domestication (Haudry et al., 2007; Luo et al., 2007; Civáň et al., 2013; Oliveira et al., 2020).

Nucleotide diversity estimates (π and θ) revealed the expected substantial loss of diversity during domestication, highlighting the greater impact of secondary domestication (Table 1). Considering π, the average nucleotide diversity of durum wheat was ~17% lower than that of domesticated emmer, which in turn was ~11% lower than that of wild emmer. The cumulative effect of primary and secondary domestication was a ~26% reduction in the nucleotide diversity of durum wheat compared to its wild ancestor (Table 1).

To ensure that our results were not biased towards the chosen reference genome, we also repeated the variant calling, PCA based on SNPs, and nucleotide diversity estimates for each subspecies using the wild emmer and durum wheat reference genomes to allow comparison with the bread wheat reference. The numbers of SNPs, nucleotide diversity estimates, and diversity loss estimates are summarized in Supplementary Data Set S2. Both the wild emmer and durum wheat references yielded fewer SNPs (604,479 and 544,406, respectively) compared to using the bread wheat reference (800,996). However, the similarity in the number of polymorphic sites between wild emmer and emmer along with the lower number of

durum wheat SNPs, as well as the ratio of private and shared SNPs were reaffirmed with both of these alternative references (Supplementary Data Set S2). The equivalence in the utilization of the three references was reinforced by identical PCA results obtained from the two sets of SNPs derived from the calls using wild emmer and durum wheat references (Supplementary Figure S2).

In addition, the estimates of π and θ for wild emmer and emmer exhibited lower subspecies-specific values compared to the bread wheat reference (Supplementary Data Set S2). Considering π, the values obtained using the wild emmer reference were 0.0027 for wild emmer, 0.0027 for emmer and 0.0024 for durum, whereas similar but lower values were obtained when using the durum wheat reference, especially for durum wheat (π = 0.0011). Even so, the overall trend of diversity loss confirmed previous findings using the bread wheat reference, indicating a more pronounced impact of secondary domestication. The percentage losses of π nucleotide diversity observed using the bread wheat reference (11.4 for primary domestication and 16.8 for secondary domestication) were in the range obtained using wild emmer (3.2 for primary domestication and 10.1 for secondary domestication) and durum wheat (12.5 for primary domestication and 48.1 for secondary domestication).

In summary, although we observed slight variations in SNP detection and nucleotide diversity estimates across various reference genomes, the general trends of genetic variation and diversity loss remained consistent. By detailing our methodology for testing available reference genomes, we aim to provide guidance for other researchers encountering similar challenges. In the case of tetraploid wheat, we propose that the use of an outgroup species as a reference genome does not introduce bias in gene identification and subsequent analyses. On the contrary, it may mitigate biases and ensure fair representation for all subspecies in the study.

The variability of gene expression during domestication was influenced by N availability

To quantify the diversity of gene expression in each subspecies, we calculated evolvability scores under high and low N-availability conditions. Evolvability was estimated using the additive coefficient of variation (CV_A) in read counts

(Supplementary Data Set S3). In contrast to heritability, CV_A is a standardized measure of additive genetic variation that is not influenced by other sources of variance (Houle, 1992; Hansen et al., 2011), and is therefore well suited for comparative analysis (Garcia-Gonzalez et al., 2012; Gioia et al., 2015). We found that the CV_A was the highest in wild emmer, followed by emmer and then durum wheat, indicating a decline during domestication under both N-availability conditions (Figure 2A,B; Table 2). The loss of diversity in gene expression has been observed across the domestication process in other crops, such as: common bean (Bellucci et al., 2014), tomato (*Solanum lycopersicum*) (Sauvage et al., 2017) and sorghum (*Sorghum bicolor*) (Burgarella et al., 2021) as well as domesticated animal species (Liu et al., 2019). However, we observed a higher mean CV_A across all three subspecies under low-N compared to high-N conditions (Figure 2A,B; Table 2). This suggests that higher N availability promotes a more uniform gene expression pattern, whereas higher variability (plasticity) is observed during N starvation.

We used the contrasting N conditions of our samples to examine whether the loss of expression diversity is associated with the specific aspects of the cultivation environment, causing primary and secondary domestication to have a substantially different impact. Under high-N conditions, we observed a ~9% loss in expression diversity in emmer compared to wild emmer (effect of primary domestication) and a ~15% loss in durum wheat compared to emmer (effect of secondary domestication). In contrast, these losses were ~18% and 11% under N starvation conditions, revealing twice the loss of expression diversity during primary domestication, but a lower value during secondary domestication (Table 2). All four values differed significantly from each other (Mann–Whitney U-test, $p < 0.001$). The opposing expression diversity profiles during domestication under high-N and low-N conditions were observed not only for overall gene expression, but also for the subgroup comprising all differentially expressed genes (DEGs) and the subgroup comprising all unmodulated genes (Supplementary Table S1). The loss of expression diversity among the DEGs due to primary domestication was ~9% and ~15% under high-N and low-N conditions, respectively, whereas the corresponding losses due to secondary domestication were ~18% and ~14% (Supplementary Table S1). The loss of expression diversity among the unmodulated genes was similar to the values for overall gene expression (Supplementary Table S1).

A phenotypic study of the same accessions used in the present work has already shown that secondary domestication reduced the phenotypic diversity under high-N conditions, but the reduction was smaller and not significant under N starvation (Gioia et al., 2015). In the case of durum wheat, selection has apparently enhanced the growth response to N availability, indicating a putative focus on improving N uptake and utilization efficiency. Our expression diversity results indicate that selection may have favored specific traits and thus led to a more uniform set of cultivars, as also suggested in an earlier study based on morphological traits (Gioia et al., 2015).

Domestication and nitrogen availability played a role in the divergence of tetraploid wheat

Genetic differentiation among the three subspecies was estimated by calculating the pairwise fixation index (F_{ST}) for every gene locus in our dataset. As shown in Figure 3A, the lowest genetic differentiation was observed between wild emmer and emmer (median F_{ST} ~ 0.08), whereas much higher genetic differentiation was found between emmer and durum wheat (median F_{ST} ~ 0.24) and, similarly, between wild emmer and durum wheat (median F_{ST} ~ 0.26). These values align with earlier findings that examined broad collections of tetraploid wheat accessions (Luo et al., 2007), suggesting an indication of the representativeness of the utilized genotypes.

Divergence at the transcriptomic level was estimated by calculating Q_{ST}, the quantitative analog of F_{ST}, taking N availability into account as an environmental variable. Under both N conditions, we observed the same trend shown for F_{ST} (Figure 3B). Specifically, secondary domestication had a stronger impact on differentiation (emmer *vs* durum wheat, median $Q_{ST\,LN}$ ~ 0.04, median $Q_{ST\,HN}$ ~ 0.15) than primary domestication (wild emmer *vs* emmer, median $Q_{ST\ LN}$ ~ 4.7×10^{-9}, median $Q_{ST\ HN}$ ~ 0.018). Interestingly, the Q_{ST} distributions of every pairwise comparison showed higher values under high-N conditions compared to N starvation (Figure 3B), suggesting that the response to N availability has been under selection during domestication and breeding substantially contributed to the differentiation of gene expression in tetraploid wheat in response to different agroecosystems.

We therefore dissected the transcriptome as a multidimensional plastic phenotype, in which the abundance and expression patterns of thousands of genes in different environmental conditions can be processed as phenotypic traits, reflecting variability in levels of gene expression. Studying the transcriptome as a plastic phenotype is a powerful approach to understand the molecular basis of plasticity and its evolutionary potential (Leinonen et al., 2013; Oostra et al., 2018). By assessing how gene expression patterns over an environmental gradient vary within populations (reflecting phenotypic plasticity) and among populations (reflecting genetic differentiation), it is possible to identify candidate genes and regulatory pathways that may be under selection for adaptation (Oleksiak et al., 2002; Whitehead and Crawford, 2006a, 2006b). In tetraploid wheat, the Q_{ST}–F_{ST} comparison method has been used to detect selection signatures for metabolites, treated as molecular phenotypic traits (Beleggia et al., 2016). Until now, despite its suitability for the analysis of gene expression data, this method has been rarely adopted in transcriptomics studies (Roberge et al., 2007; Kohn et al., 2008; Aykanat et al., 2011).

We implemented a methodology based on Q_{ST} distributions and Q_{ST}–F_{ST} comparisons to perform a “selection scan”, seeking genes whose expression was potentially under selection. To ensure the reliability of gene selection and minimize the risk of false positives, we established two thresholds based on the distributions of gene expression heritability (Supplementary Figure S3), while acknowledging that we developed this methodology as a proof of concept with a limited number of genotypes. Consequently, we aimed to establish a high confidence level to ensure the robustness of our approach. Genes with $H^2 < 0.7$ were removed in order to retain only candidates in the top 15% of the distribution for which we can ascertain that their variation is predominantly genetic. However, to consider those genes whose expression was strongly influenced by N availability, we also evaluated the percentage of the species × environment ($S \times N$) variance component (i.e., every species subgroup × N condition), retaining those genes meeting at least the 20% threshold (Supplementary Figure S3). Altogether, 5,868 genes (~18% of the total number) met these criteria and only those with Q_{ST} values in the 5% right-hand tail of the distributions were considered as candidates for selection. The Q_{ST} and F_{ST} values of the filtered genes were then compared (Supplementary Figure S4) to

confirm that their divergent expression (high Q_{ST} values) was caused by directional selection ($Q_{ST} > F_{ST}$) and not by genetic drift ($Q_{ST} \approx F_{ST}$) or stabilizing selection ($Q_{ST} < F_{ST}$) (Bonnin et al., 1996; Whitlock and Guillaume, 2009). After removing F_{ST} values < 0.01, we observed that all the resulting 967 genes satisfied the criterion $Q_{ST} > F_{ST}$, indicating that their expression was probably subjected to directional selection in at least one of the evolutionary contexts examined herein (i.e., primary and/or secondary domestication under high-N and/or low-N availability conditions). Notably, ~280 genes were consistently detected in each of the six comparisons (Supplementary Data Set S4). These genes exhibited high Q_{ST} values, indicating substantial divergence in expression levels, coupled with lower F_{ST} values, suggesting limited nucleotide-level divergence. This suggests that the observed signals likely arise from upstream regulatory mechanisms influencing gene expression, rather than mutations within the gene's coding regions, potentially leading to altered gene products. Consequently, we conducted SNP annotation on the 967 genes under selection, accounting for both synonymous and non-synonymous mutations. We then compared these results by annotating variants present in another set of 967 genes randomly selected from the total pool of 32,358 genes. As detailed in Supplementary Table S2, for each *taxon*, we detected an average lower number of both synonymous and non-synonymous mutations in the genes under selection compared to the randomly selected group. Specifically, the ratios of non-synonymous/synonymous mutations were ~0.78 for selected genes and ~0.81 for random ones in wild emmer and emmer wheat, showing comparable values. However, in durum wheat, this ratio was significantly lower: ~0.69 for selected genes and ~0.76 for random ones (Kolmogorov-Smirnov two-sided test $p < 2.2e\text{-}16$). These findings bolster our identified signals of selection, suggesting that our 967 candidate genes, precisely because they are putatively under directional selection, may have also undergone a process of purifying selection, thereby preventing the accumulation of deleterious mutations.

Gene Ontology (GO) enrichment analysis was performed on the six groups of genes. Not all six comparisons yielded significant results, but significant and distinct GO categories emerged specifically during primary and secondary domestication under high-N (Supplementary Figure S5). However, examining the entire evolutionary process (from wild emmer to durum wheat), categories associated with the amino

acid biosynthesis were enriched also in genes showing selection signatures when comparing wild emmer and durum wheat under low-N conditions. (Supplementary Figure S5). Our results suggest that primary and secondary domestication involved different natural and artificial selection pressures affecting distinct sets of genes and phenotypes influenced by N availability, indicating that certain genetic pathways may have been particularly important for the adaptation of wheat to different environmental conditions during its domestication history.

Among the genes under selection during primary domestication, categories linked to "defense-related programmed cell death, modulated by biotic interactions" were enriched, suggesting an enhanced hypersensitive response to pathogens. As wild genotypes transition to agroecosystems characterized by dense crop monocultures, they encounter increased disease pressure from crop-specific pathogens (Savary et al., 2019). This prompts a hypersensitive response, potentially leading to programmed cell death and necrosis as a defense mechanism. Pathogen defense mechanisms in plants often intersect with the regulation of beneficial symbiotic interactions, suggesting a trade-off between symbiosis-associated traits and innate immunity (Porter and Sachs, 2020). Moreover, domesticated crops are less able to leverage microbial interactions compared to wild counterparts, as shown by comparative studies involving bread wheat landraces and old *vs* modern varieties (Valente et al., 2023). This reduced capacity may in part reflect the widespread adoption of high-input agricultural practices, wherein the availability of fertilizers diminishes the need for plants to invest in symbiotic relationships (Martín-Robles et al., 2018).

Among the genes under selection during secondary domestication, we observed the enrichment of categories associated with amino acid metabolism, particularly those related to the "lysine catabolic process" (Supplementary Figure S5). This included genes encoding the bifunctional enzyme lysine ketoglutarate reductase/saccharopine dehydrogenase (LKR/SDH), which breaks down lysine via the saccharopine pathway (SACPATH). The structure and transcription of the LKR/SDH gene have been investigated in durum wheat, revealing species-dependent differences in expression and lineage-specific variations between monocots and dicots (Anderson et al., 2010). Lysine is a limiting essential amino acid in cereal grains, and efforts have been made to enhance its content in crops like

maize (*Zea mays*) and rice (*Oryza sativa*) by targeting the catabolic pathway (Houmard et al., 2007; Frizzi et al., 2008; Long et al., 2013). However, lysine-rich proteins generally do not accumulate to high levels in cereal seeds, which instead stockpile prolamins (such as gliadin in wheat). The SACPATH appears to direct lysine toward the production of glutamic acid, a precursor of proline, which is abundant in gluten (Arruda and Barreto, 2020).

Evolutionary metabolomics has revealed signatures of selection affecting amino acid metabolism during secondary domestication (Beleggia et al., 2016). Changes in amino acid metabolism during domestication have been observed in crops such as sunflower, maize, and common bean based on nucleotide data (Chapman et al., 2008; Swanson-Wagner et al., 2012; Bellucci et al., 2014). In durum wheat, domestication has been associated with the selection of specific protein compositions, reducing the diversity of gliadin and glutenin subunits, thus affecting grain yield and gluten properties (Laidò et al., 2013, 2014). Moreover, the SACPATH is upregulated in response to drought stress in spring wheat genotypes, particularly in drought-tolerant varieties, suggesting a role in stress adaptation (Michaletti et al., 2018). Proline, derived from this pathway, may serve as a major constituent of storage proteins and also a key osmoprotectant produced in response to stress (Kavi Kishor et al., 2022). Overall, selection during wheat domestication may have influenced the expression of SACPATH genes, favoring not only protein composition but also abiotic stress tolerance.

Changes in nitrogen availability trigger gene expression, resulting in a twofold increase in the number of differentially expressed genes in durum wheat compared to emmer and wild emmer wheat.

We identified DEGs in each subspecies that discriminated between high-N conditions and N starvation using a stringent pipeline and strict thresholds ($p_{adj} < 0.001$) to reduce the number of false positives. We found 3,326 DEGs in wild emmer, 3,305 in emmer and 5,901 in durum wheat, with more upregulated than downregulated genes in all three subspecies. Durum wheat had the highest percentage of private DEGs (~42%, 2,479), whereas similar numbers were found in wild emmer (~14%, 458) and emmer (~15%, 486). Wild emmer and emmer shared

~23% (749) and ~21% (700), respectively, of their DEGs with durum wheat. The percentage of DEGs shared only between wild emmer and emmer was 4% (146), but almost 60% of wild emmer and emmer DEGs and ~33% of durum wheat DEGs were shared by all three *taxa* (Figure 4A). The proportions of private and shared DEGs were preserved when we separated them into upregulated and downregulated subsets (Figure 4B,C). In all three *taxa*, most DEGs were located on chromosomes 2A, 2B, 3A, 3B, 5A and 5B, each carrying > 7.5% of the DEGs, whereas chromosomes 6A and 6B each contained only ~5% of the DEGs (Supplementary Figure S6).

GO enrichment analysis of the DEGs meeting the threshold FDR < 0.05 revealed 23 macro-categories in wild emmer, 21 in emmer and 25 in durum wheat (Supplementary Figure S7). The main differences between the three subspecies were observed for categories related to “signaling”, “regulation of biological process”, “developmental process”, and “metabolic process” (Supplementary Figure S7). We observed the uniform enrichment of GO categories associated with upregulated genes in all three subspecies, including terms linked to N and amino acid metabolism as well as carbon (C) metabolism and photosynthesis (Supplementary Figure S7). In contrast, the enrichment of GO categories associated with downregulated genes was more selective, with some GO categories related to N metabolism enriched only in durum wheat, including GO:0006807 and GO:0034641 (N compound and cellular N compound metabolic process, respectively) and GO:0006536 “glutamate metabolic process” (Supplementary Data Set S5). Functional annotations of the most strongly modulated genes (top 5% $|\log_2 FC|$ values) are reported in Supplementary Data Set S6.

Our data confirm, on a larger set of samples, earlier observations on the response of wheat to N starvation based on transcriptomics and metabolomics data. These earlier studies included one emmer and one durum wheat genotype also present in our sample set (Beleggia et al., 2021), but also considered the durum wheat cultivar Svevo (Curci et al., 2017) and various bread wheat cultivars (Sultana et al., 2020). As expected, genes involved in N metabolism were modulated during N starvation. Among the key genes for N assimilation, those encoding asparagine synthetase and nitrite reductase were upregulated in every *taxon*, whereas those encoding glutamate carboxypeptidase and glutamate decarboxylase were downregulated. We

observed contrasting profiles for genes encoding ureide permease (a ureide transporter), which were strongly upregulated in all three subspecies in response to N stress, whereas genes encoding nitrate transporters were strongly downregulated. The modulated genes also included transporters of amino acids and other nutrients.

N starvation also influenced other metabolic pathways, revealing many further DEGs involved in C metabolism, especially fatty acid metabolism, glycolysis, photosynthesis, and the tricarboxylic acid (TCA) cycle. About 10% of the highest-ranking DEGs represented transcription factors and protein kinases. The most common functional category (accounting for 17% of annotated DEGs) reflected the general stress response to N starvation, including the mitigation of oxidative stress and detoxification. Examples included genes encoding cytochrome P450s, glutaredoxin family proteins, glutathione S-transferases and peroxidases (Supplementary Data Set S6).

To compare gene expression between the three *taxa* while taking the environmental effects into account, we also identified DEGs between each pair of subspecies under all N conditions. Accordingly, we compared emmer *vs* wild emmer (primary domestication, high- and low-N), durum wheat *vs* emmer (secondary domestication, high and low N) and durum wheat *vs* wild emmer (cumulative effect, high- and low-N) (Supplementary Figure S8). The wild emmer *vs* emmer comparison revealed few DEGs regardless of N availability (12 and 11 DEGs under high-N and low-N conditions, respectively), whereas the emmer *vs* durum wheat comparison revealed 41 DEGs associated with high N and 29 associated with N starvation, and the wild emmer *vs* durum wheat comparison revealed 46 DEGs associated with high-N and only 10 associated with N starvation. These data indicate that the number of DEGs increases during domestication but only when there is a sufficient N supply (Supplementary Figure S8). Interestingly, there were more upregulated than downregulated genes in all pairwise comparisons under high-N conditions (~65%) but the proportion increased under N starvation, particularly for the comparison of wild emmer *vs* durum wheat (90%). The preponderance of upregulated genes during domestication has also been observed in maize (Lemmon et al., 2014), whereas domestication was shown to increase the proportion of downregulated genes in common bean (Bellucci et al., 2014), eggplant (*Solanum melongena*) (Page et al., 2019) and sorghum (Burgarella et al., 2021) landraces compared to wild relatives.

The absence of consistent patterns suggests that the evolution of domesticated phenotypes is driven by specific processes that are unique to each crop.

Among the 102 DEGs found in at least one of the six pairwise comparisons between subspecies (Supplementary Data Set S7), 35 were also found among DEGs identified between contrasting N conditions and 24 of these were proposed to be under selection. Overall, six genes were identified in all three experiments (i.e., differentially expressed between subspecies and between contrasting N conditions, and showed evidence of selection).

Selection shaped the expression profiles of genes modulated by nitrogen availability

The 6,991 DEGs found in at least one species when comparing contrasting N conditions included 101 putatively under selection, which are candidates for the adaptive response to N availability. The expression profiles of these selected genes, in all three subspecies under both N conditions, are shown in Supplementary Figure S9. We applied PCA to the normalized read counts in order to determine whether the different genotype groups can be separated based on their gene expression. Initially we incorporated all 6,991 DEGs (Figure 5A,B) before focusing on the subset of 101 DEGs that were also putatively under selection (Figure 5C,D). When considering all DEGs, PC1 did not completely separate the durum wheat genotypes from the other *taxa*, in contrast to the clear separation observed for the SNP data (Figure 1), and this was particularly evident during N starvation (Figure 5A). There was also a moderate degree of overlap between the wild emmer and emmer genotypes along PC2. However, when we focused on the DEGs under selection, PC1 separated the durum wheat genotypes into a densely clustered group (as observed for the SNP data) under both N conditions, and PC2 separated the wild emmer and emmer genotypes more clearly, especially under high-N conditions (Figure 5C,D). For an overview of overall gene expression patterns in the two N conditions, please refer to Supplementary Figure S10, which illustrates the PCA conducted on the complete set of genes (32,358).

The selection signatures (based on Q_{ST}–F_{ST} values) allowed us to identify genes that have putatively diverged between tetraploid wheat subspecies due to selection

pressure. By integrating this information with classical differential expression analysis, which identified genes that are differentially expressed under contrasting N conditions, it was possible to pinpoint 101 candidate genes that are both genetically differentiated and functionally relevant to the environmental factor of interest: N availability during the domestication and diversification of cultivated wheat. Functional annotation (Supplementary Data Set S8) revealed upregulated genes associated with C metabolism as well as some encoding transcription factors and transporters, as well as both upregulated and downregulated genes associated with general stress responses and N metabolism, specifically those encoding enzymes involved in amino acid metabolism such as methionine aminopeptidase, aspartokinase and glutamate dehydrogenase (GDH). The latter is particularly noteworthy because, in addition to its modulation in response to different N conditions and the presence of selection signatures, the *GDH* gene was also upregulated in the comparison between wild emmer and durum wheat under high-N conditions (Supplementary Data Set S7). GDH is a key enzyme involved in N metabolism and N/C balance (Miflin and Habash, 2002). This is supported by the co-localization of quantitative trait loci for GDH activity and physiological traits associated with the flag leaf lamina, such as soluble protein and amino acid content, as well as flag leaf area and dry weight (Fontaine et al., 2009). Selection signatures were also identified in the *GDH* gene when comparing landraces with old and modern durum wheat cultivars (Taranto et al., 2020). Our results confirm that N metabolism has been a key driver during the evolutionary history of wheat, particularly the central role of glutamate in the process of domestication. This was also suggested by a combined transcriptomics and metabolomics study showing that glutamate and γ-aminobutyric acid (mainly synthesized from glutamate) are central to the genotype-specific response of emmer and durum wheat to N starvation (Beleggia et al., 2021).

We have shown that significant changes occurred at the nucleotide and gene expression levels during the domestication of tetraploid wheat, taking into account the environmental variable of N availability. We observed that more nucleotide diversity has been lost during secondary domestication compared to primary domestication. In addition, we found a parallel trend in the loss of gene expression

diversity associated with the domestication process, with a stronger effect due to secondary domestication. The observed loss of expression diversity may be related to N availability in the durum wheat selection environment. Our findings suggest that selection may have operated in different directions during primary and secondary domestication, the former involving changes related to biotic interactions and the latter related to amino acid metabolism.

Despite the limited number of genotypes available for our study, the innovative combination of RNA-Seq analysis and the estimation of quantitative genetic parameters allowed us to develop a pipeline for the identification of selection signatures and phenotypic plasticity in gene expression data based on evolvability and Q_{ST}–F_{ST} scores. Emphasizing the pioneering nature of our work, we opted to introduce stringent and high confidence thresholds for the considered parameters, aiming to present our methodology as a proof of concept. While presenting promising results, we acknowledge the potential for further refinement and adjustment of the methodology in experiments employing larger genotype datasets. The set of genes we identified with underlying selection signatures will facilitate the development of innovative strategies to improve the resource use efficiency and environmental sustainability in crop management.

Materials and Methods

Plant material and experimental design

The study included 32 tetraploid wheat genotypes, comprising 10 accessions of wild emmer (*T. turgidum* ssp. *dicoccoides*), 10 accessions of emmer (*T. turgidum* ssp. *dicoccum*), and 12 accessions of durum wheat (*T. turgidum* ssp. *durum*) (Supplementary Data Set S1). The samples analyzed in our study were selected from a larger experiment conducted in October 2012, as previously described (Gioia et al., 2015) and were chosen as representative of the majority of the diversity within the panel. Briefly, wheat genotypes were grown for 4 weeks under high-N and nitrogen starvation (low-N) conditions in the Phytec Experimental Greenhouse at the Institute of Biosciences and Geosciences (IBG-2), Plant Sciences Institute, Forschungszentrum Jülich GmbH, Germany (50°54′36″ N, 06°24′49″ E). Seeds of uniform size and mass were visually selected, surface sterilized with 1% (*w/v*) NaClO

for 15 min and pre-germinated. After germination, seedlings showing uniform growth (seminal root length, 1–2 cm) were transferred to soil-filled rhizoboxes, which were placed in the automated GROWSCREEN-Rhizo phenotyping system available at IBG-2. We used a Type 0 manually sieved peat soil (Nullerde Einheitserde; Balster Einheitserdewerk, Frondenberg, Germany), which provided low nutrient availability (ammonium N and nitrate N concentrations of < 1.0 and < 1.0 mg l^{-1}, respectively). All plants were watered twice daily with 400 ml tap water and were supplied three times per week with 200 ml modified Hoagland solution (Hoagland and Arnon, 1950), adapted for optimal N and N starvation conditions. Stock solution contained 5 mM KNO_3, 5 mM $Ca(NO_3)_2$, 2 mM $MgSO_4$, 1 mM KH_2PO_4, and trace elements. For N starvation conditions, KNO_3 and $Ca(NO_3)_2$ were replaced with K_2SO_4 and $CaCl_2 \cdot 6(H_2O)$, respectively. The experiment was carried out under natural lighting in the greenhouse, with an air temperature of 18–24 °C and a relative humidity of 40–60%. For each N treatment, we used two replicates of each genotype with two plants per replicate (four plants per genotype in total). After 4 weeks, leaves were pooled from two plants of the same genotype growing in the same rhizobox. Accordingly, four independent biological replicates (two replicates per N condition) were produced for each genotype, with the exception of wild emmer IG 46504, PI 233288, PI 466991, PI 538656, emmer MG 5293/1, and durum wheat Creso, Pedroso and Trinakria, for which only three replicates were available, and emmer Molise Sel. Colli and durum wheat Simeto, for which eight replicates were available. The tissues were immediately frozen in liquid N_2 and stored at −80 °C. Further details of the experiment and growth conditions are provided elsewhere (Gioia et al., 2015).

RNA extraction and sequencing

RNA was extracted from 100 mg of frozen ground leaves per replicate using the Spectrum Plant Total RNA kit (Sigma-Aldrich, St Louis, MO, USA) followed by treatment with RNase-free DNase using the On-Column DNase I Digestion Set (Sigma-Aldrich). RNA integrity and purity were assessed by agarose gel electrophoresis and a Bioanalyzer 2100, respectively (Agilent/Bonsai Technologies, Santa Clara, CA, USA). Only RNA samples with an RNA integrity number > 8.0 were considered suitable for analysis.

Library construction and RNA sequencing were carried out using the Illumina mRNA-Seq platform at the Montpellier Genomix sequencing facility (http://www.mgx.cnrs.fr)

as previously described (David et al., 2014). Briefly, RNA samples were processed using TruSeq RNA sample preparation kits v2 (Illumina, San Diego, CA, USA). Libraries were quantified by RT-qPCR using the KAPA Library Quantification Kit for Illumina Sequencing Platforms (Roche, Basel, Switzerland), followed by quality control using a DNA 100 Chip on a Bioanalyzer 2100. Cluster generation and sequencing were carried out using the Illumina HiSeq 2000 instrument and TruSeq PE Cluster Kit v3, following the Illumina PE_Amp_Lin_Block_V8.0 recipe, and Illumina TruSeq PE Cluster v3-cBot-HS kits with the 2 × 100 cycles, paired-end, indexed protocol, respectively (David et al., 2014).

RNA-Seq library processing and mapping

We pre-processed 128 raw paired-end RNA-Seq libraries (David et al., 2014). Cutadapt (Martin, 2011) was then used to remove adaptor sequences and trim the end of reads with low quality scores (parameter -q 20) while keeping reads with a minimum length of 35 bp. Reads with a mean quality score < 30 were discarded, and orphan reads (whose mates were discarded in the previous filtering steps) were also removed (David et al., 2014). The final quality of trimmed and filtered reads was assessed using FastQC (Andrews, 2014).

The bread wheat (*Triticum aestivum* cv. Chinese Spring) genome assembly IWGSC RefSeq v2.1, along with the corresponding genome annotation, were downloaded from the IWGSC data repository hosted by URGI-INRAE (https://wheat-urgi.versailles.inra.fr/) and used as a reference to map each cleaned library to the A and B subgenomes. To validate our choice of reference genome, reads were also mapped to the available tetraploid wheat reference genomes: wild emmer accession Zavitan (https://www.ncbi.nlm.nih.gov/datasets/genome/GCA_002162155.3/) and durum wheat cultivar Svevo (https://www.interomics.eu/durum-wheat-genome). Pairwise genetic distances between the three reference genomes were computed using Mash v2.3 (Ondov et al., 2016).

STAR v2.7.0e (Dobin et al., 2013) was used for read mapping with the --quantMode TranscriptomeSAM and --quantTranscriptomeBan Singleend options. The output alignments were translated into transcript coordinates (in addition to alignments in genomic coordinates), allowing insertions, deletions and soft-clips in the transcriptomic alignments. The transcriptomic alignments were used as inputs for

salmon v1.6.0 (Patro et al., 2017) to quantify gene expression. Raw read counts were computed for all genes in each sample and, to filter out weakly-expressed transcripts, only genes with at least 1 count per million (CPM) in at least 10 samples (of the same subspecies) were retained. This was calculated separately in each of the three subspecies and the raw counts of the filtered genes in each subspecies were then combined for downstream analysis (Supplementary Data Set S3).

Variant identification

Variants were called by applying BCFtools v1.15 (previously SAMtools) (Danecek et al., 2021) to the alignment bam files. The "*bcftools mpileup*" command was used to determine the genotype likelihoods at each genomic position, with a minimum alignment quality of 20 and a minimum base quality of 30. The actual calls were obtained using the "*bcftools call*" command. The resulting VCF file was filtered using the "*bcftools view*" command, removing indels and keeping only sites covered by at least three reads in all genotypes. Subsequently, only biallelic SNPs with maximum values of 50% missingness and a 1% minor allele frequency were retained. To identify private and shared SNPs among the different subspecies, every possible comparison of the three subsampled VCF files (wild emmer, emmer and durum wheat) was carried out using the "*bcftools isec*" command.

Population genetics analysis

Variants were filtered (one SNP per 500 kb) using the VCFtools v0.1.17 *--thin 500000* option (Danecek et al., 2011) and then converted to ped format with PLINK v1.90p (Purcell et al., 2007). PLINK was also used to compute genetic distances between individuals with the *--distance-matrix* flag. The output matrix was used as input for PCA with the *cmdscale* function of R v4.2.1 (R Core Team, 2022).

Genetic diversity statistics, including nucleotide diversity (π and θ) (Tajima, 1983; Watterson, 1975) were computed on the alignment bam files for each subspecies, from the folded site frequency spectra using ANGSD (Korneliussen et al., 2014). First, the *doSaf* function was used to estimate per-site allele frequencies (Saf) then *realSFS* was used to get the site frequency spectra. The loss of diversity statistic (Vigouroux et al., 2002) was used to test the impact of primary and secondary domestication on the molecular diversity of the three subspecies. For primary domestication, the statistic was computed as $[1 - (x_{emmer}/x_{wild})]$, where x_{emmer} and x_{wild}

are the diversities in emmer and wild emmer, respectively, measured using π and θ. If x_{emmer} was higher than x_{wild}, then the parameter was calculated as [(x_{wild}/x_{emmer}) – 1]. The loss of diversity due to secondary domestication in durum wheat versus emmer was calculated as [1 − (x_{durum}/x_{emmer})], where x_{durum} and x_{emmer} are the diversities in durum wheat and emmer, respectively. If x_{durum} was higher than x_{emmer}, then the parameter was calculated as [(x_{emmer}/x_{durum}) – 1].

We calculated F_{ST} for each pair of populations using ANGSD (Korneliussen et al., 2014). Saf and 2D site frequency spectra were calculated as for nucleotide diversity, and then the *fst index* function was used to obtain the global estimate. To get an F_{ST} value for each gene in our dataset, we used the *fst print* function, which prints the posterior expectation of genetic variance between populations (called A), and total expected variance (called B) for every locus. We then computed the weighted F_{ST} as the ratio of the summed As and summed Bs for every gene region, using an *ad hoc* R script.

Expression profiles, heritability and Q_{ST} analysis

Raw read counts of the 32,358 genes were normalized using the *vst* method allowing the additive coefficient of variation (CV_A) (standard deviation/mean) to be calculated for the two N conditions in every subspecies, averaging the biological replicates of every genotype. The statistical loss approach (Vigouroux et al., 2002) was then applied to test the loss of expression diversity in the different groups, as previously reported (Bellucci et al., 2014). The statistical significance of the differences between each CV_A value and the percentage loss of expression diversity was determined using the Mann-Whitney U-test in R v4.2.1 (R Core Team, 2022) with the function *wilcox.test*.

To compute heritability, the raw counts of each subspecies under each condition were first normalized using the trimmed mean M-values normalization method in the R package edgeR (Robinson et al., 2010) and the voom normalization method in the R package limma (Smyth, 2005). To determine the variance component of each factor and heritability, the following model was considered:

$$Y_{ijkl} = S_i + G_{j(i)} + N_k + (S \times N)_{ik} + (G \times N)_{jk(i)} + \varepsilon_{l(ijk)},$$

where Y_{ijkl} is the normalized gene expression level, S_i is the species factor, $G_{j(i)}$ is the genotype factor nested in species, N_k is the N-level factor, $(S \times N)_{ik}$ is the

interaction between species and N levels, $(G \times N)_{jk(i)}$ is the interaction between genotypes and N levels, and $\varepsilon_{l(ijk)}$ is the residual error. All factors were treated as random effects in the model except the intercept, which was a fixed effect. The linear mixed models were fitted using the *lmer* function in R package lme4 based on the normalized data of each transcript (Bates et al., 2015). The heritability (H^2) was calculated as $H^2 = \frac{V_S+V_G}{V_A}$, where $V_A = V_S + V_G + V_N + \frac{V_{S\times N}}{n} + \frac{V_{G\times N}}{n} + \frac{V_\varepsilon}{n}$, V_S is the variance of species, V_G is the variance of genotype, V_N is the variance of N level, $V_{S\times N}$ is the variance of species and N level interaction, $V_{G\times N}$ is the variance of genotype and N level interaction, V_ε is the residual variance, and n is the number of N levels. $V_{S\times N}$ and $V_{G\times N}$ represent the genotype × environment interaction variance components at the species and genotype (nested in species) levels, respectively.

Q_{ST} was calculated between pairs of the three subspecies under low-N and high-N conditions separately. The wild emmer *vs* emmer comparison revealed the effects of primary domestication, the emmer *vs* durum wheat comparison revealed the effects of secondary domestication, and the wild emmer *vs* durum wheat comparison revealed the cumulative effect of domestication. To this end, the model can be reduced to $Y_{ijl} = S_i + G_{j(i)} + \varepsilon_{l(ij)}$ at each N level. The Q_{ST} value was calculated as $Q_{ST} = \frac{V_S}{V_S+V_G}$, that is, the ratio of between-species and within-species variance.

Q_{ST} distributions were used to perform a "selection scan" on a restricted number of genes. First, genes were filtered for $H^2 \geq 0.7$ and, in order not to lose genes whose expression was strongly influenced by N availability, the species × environment (*S*×*N*) variance component was also evaluated (i.e., every species subgroup × N condition), retaining those genes meeting the threshold *S*×*N* ≥ 0.2 (Supplementary Figure S3). Successively, we obtained six different Q_{ST} value distributions (Q_{ST} WILD EMMER *VS* EMMER, Q_{ST} EMMER *VS* DURUM WHEAT and Q_{ST} WILD EMMER *VS* DURUM WHEAT, each for high-N and low-N conditions) and we retained the 5% right-hand tail of every distribution. Finally, we compared F_{ST} and Q_{ST} values for every gene, discarding F_{ST} values < 0.01. We confirmed that every retained gene satisfied the condition $Q_{ST} > F_{ST}$, allowing it to be classed as undergoing directional selection. For these selected genes, SNPs were annotated using SnpEff v.5.1d (Cingolani et al., 2012).

Differential expression analysis

764 Differential gene expression was assessed by analyzing the pre-processed raw
765 count dataset (32,358 genes). We identified DEGs by comparing (*i*) two conditions
766 (i.e., high-N and low-N levels) within each subspecies, and (*ii*) pairs of the three
767 subspecies under the same N levels, which considered the genotypes nested in
768 species. For the two scenarios, we used three different approaches to detect DEGs:
769 one linear model-based approach implemented in the R package limma (Smyth,
770 2005), and two Poisson model-based approaches implemented in the R packages
771 edgeR (Robinson et al., 2010) and DESeq2 (Love et al., 2014). In all approaches,
772 the normalization of raw counts was applied by default in the package before
773 differential analysis. To reduce the number of false positives, the intersection of
774 DEGs resulting from the three approaches was retained (Zhang et al., 2014;
775 Rapaport et al., 2013) and the significance threshold was set to an adjusted p-value
776 < 0.001. The DEGs between high and low N levels in at least one subspecies were
777 used for PCA following the *DESeq2* approach (Love et al., 2014), first using all the
778 DEGs, then repeating the analysis on the DEGs considered to be under selection. At
779 each step, counts were normalized using the *vst* method before the *plotPCA* function
780 was applied to define principal components 1 and 2 for the two N levels separately.
781 The expression patterns of the DEGs considered to be under selection were plotted
782 from normalized read counts using R package pheatmap (Kolde, 2019).

783 **GO enrichment analysis**

784 Enriched terms in the DEGs and genes under selection were identified using agriGO
785 v.2.0 (Tian et al., 2017). All annotated genes of bread wheat were used as
786 background and the following parameters were set: hypergeometric test, multiple
787 hypothesis test adjustment according to the Hochberg FDR procedure at significance
788 level < 0.05, and minimum number of mapping entries = 3.

789 **Accession Numbers**

790 The RNA-Seq libraries generated and analyzed in this study have been deposited in
791 the Sequence Read Archive (SRA) of the National Center of Biotechnology
792 Information (NCBI) under BioProject number PRJNA1015013.

793 The nucleotide sequences of all discussed genes are available on the GarinGenes
794 Database (https://wheat.pw.usda.gov/cgi-bin/GG3/browse.cgi) and can be accessed

using the Gene model IDs provided in the Supplementary Data Sets (e.g., *TraesCS2A03G0941300*).

Data availability

Scripts used in this study are available at Github https://github.com/PapaLab/transcriptomics_triticum.git.

Supplementary Data

Supplementary Figure S1: Folded site frequency spectra (SFS) of single nucleotide polymorphisms (SNPs) in the three wheat *taxa*.

Supplementary Figure S2: Principal component analysis (PCA) of 32 wheat genotypes based on single-nucleotide polymorphisms (SNPs) using different reference genomes.

Supplementary Figure S3: Workflow of gene expression selection scanning.

Supplementary Figure S4: Comparison of *QST* and *FST* estimates of the 5,868 genes showing $H^2 \geq 0.7$ or $S \times N \geq 0.2$.

Supplementary Figure S5: Gene Ontology (GO) categories of genes under selection.

Supplementary Figure S6: Genome-wide distribution of differentially expressed genes (DEGs) in the comparison between contrasting N conditions within each subspecies.

Supplementary Figure S7: GO classification of DEGs in the comparison between contrasting N conditions within each subspecies.

Supplementary Figure S8: DEGs between subspecies.

Supplementary Figure S9: Expression profiles of the 101 DEGs putatively under selection in the three wheat *taxa* in high-N (N+) and low-N (N–) conditions.

Supplementary Figure S10: Principal component analysis (PCA) of 32 wheat genotypes based on expression data of all 32,358 genes in each subspecies.

822 **Supplementary Table S1:** Mean CV_A in gene expression for the three wheat *taxa*
823 and loss of expression diversity for two gene subgroups (6,991 differentially
824 expressed and 25,367 non-differentially expressed genes).

825 **Supplementary Table S2:** Synonymous and non-synonymous mutations and non-
826 synonymous/synonymous ratios of two gene groups: 967 genes under selection and
827 967 randomly selected genes.

828 **Supplementary Data Set S1A-C:** List of the 128 samples and reads mapping
829 results using three different reference genomes.

830 **Supplementary Data Set S2:** Results of single-nucleotide polymorphism (SNP)
831 calling and nucleotide diversity estimates, using three different reference genomes.

832 **Supplementary Data Set S3:** Raw read counts of the 32,358 genes in the 128
833 RNA-Seq samples.

834 **Supplementary Data Set S4:** List of the 967 genes retained from the "selection
835 scan".

836 **Supplementary Data Set S5A-D:** List of GO "Biological process" and "Molecular
837 function" subcategories for upregulated and downregulated DEGs under different
838 nitrogen N conditions for each subspecies.

839 **Supplementary Data Set S6:** Functional annotations of the DEGs between N
840 conditions in each subspecies. Genes with the top 5% |log2FC| values are shown.

841 **Supplementary Data Set S7:** Functional annotations of the DEGs between
842 subspecies under all N conditions.

843 **Supplementary Data Set S8:** Functional annotation of the 101 genes selected by
844 the integration of selection signatures and differential expression analysis between N
845 conditions.

846

847

848 **Funding information:** This work was financially supported by Università Politecnica
849 delle Marche, by the PON a3 PlASS Project. PlASS – Platform for Agrofood Science
850 and Safety and by the Transnational Access capacities of the European Plant

Phenotyping Network (EPPN, grant no. 284443) funded by the FP7 Research Infrastructures Program of the European Union.

H.T. and Z.N. were funded by the European Union's Horizon 2020 research and innovation program through grant no. 862201.

We acknowledge the CINECA award under the ISCRA initiative, for the availability of high-performance computing resources and support.

Author contributions: R.P. and R.B. conceived and designed the study. T.G. and F.F. carried out the experiments. J.L.D. carried out the library preparation and sequencing. A.P. performed the RNA-Seq analysis. A.P., H.T. and Z.N. performed the bioinformatics analysis and analyzed the data. C.D.Q., A.R.L. and M.R. provided technical support for RNA-Seq analysis. A. P., R.B. and R.P. wrote the paper. Z.N., J.L.D., U.S., V.D.V., G.F., E.Bi., L.N., E.Be., N.P., S.M. and P.D.V. reviewed and contributed to the editing of the manuscript. All authors have read and approved the manuscript.

Declaration of interests: The authors declare no competing interests.

Tables

				Loss of nucleotide diversity (%)		
	Wild emmer	**Emmer**	**Durum wheat**	Lpd	Lsd	Both
π	0.0050	0.0045	0.0037	11.4	16.8	26.3
θ	0.0047	0.0040	0.0029	15.3	27.2	38.3

Table 1: Nucleotide diversity estimates and diversity loss for the three wheat *taxa*. Diversity loss is shown during primary domestication (wild emmer to emmer, Lpd), secondary domestication (emmer to durum wheat, Lsd) and both processes (wild emmer to durum wheat), based on average π and θ estimates of nucleotide diversity.

				Loss of expression diversity (%)		
	Wild emmer	**Emmer**	**Durum wheat**	Lpd	Lsd	Both
CV_A high N	0.062	0.056	0.048	9.1	14.5*	22.3
CV_A low N	0.076	0.063	0.056	17.6*	11.1*	26.7

Table 2: Mean additive coefficient of variation (CV_A) in gene expression and loss of expression diversity for the three wheat *taxa*. Diversity loss is shown during primary domestication (wild emmer to emmer, Lpd), secondary domestication (emmer to durum wheat, Lsd) and both processes (wild emmer to durum wheat), based on averaged CV_A values calculated for all 32,358 genes. *$p < 0.001$, Mann–Whitney U-test for difference between Lpd and Lsd within each N condition and difference between high N and low N within Lpd and within Lsd.

Figure 1: Principal component analysis (PCA) of 32 wheat genotypes based on single-nucleotide polymorphisms (SNPs). The first two principal components (PC1 and PC2) are shown. The three colors represent different *taxa*. Labels show the accession name of each genotype.

Figure 2: Density plots of the additive coefficient of variation (CV_A) in the three wheat *taxa*. Comparison of the estimated density functions of the CV_A in gene expression, calculated using all 32,358 genes. **A** Low-N conditions. **B** High-N conditions. Dashed lines represent the averaged CV_A value, colored according to the different *taxa*.

Figure 3: F_{ST} and Q_{ST} distributions. **A** Boxplots showing the gene locus F_{ST} distribution for every subspecies pairwise comparison. **B** Boxplots showing the gene expression Q_{ST} distribution for every subspecies pairwise comparison under low-N and high-N conditions, represented by empty and hatched grayscale bars, respectively. The borders of the box represent the 25th and 75th percentiles. The horizontal line in the middle of the box represents the median. Whiskers extend to the minimum and maximum values, unless a point exceeds 1.5 times the interquartile range, in which case the whisker extends to this value and values beyond are plotted as individual points (outliers).

Figure 4: Venn diagrams showing differentially expressed genes (DEGs) when comparing high-N and low-N conditions within each subspecies. A Total set of DEGs. **B** Upregulated DEGs only. **C** Downregulated DEGs only.

Figure 5: Principal component analysis (PCA) of 32 wheat genotypes based on expression data of differentially expressed genes (DEGs) when comparing high-N and low-N conditions within each subspecies. A,B Plots based on all 6,991 DEGs (not filtered): **A** low-N conditions and **B** high-N conditions. **C,D** Plots based on 101 DEGs that are also putatively under selection: **C** low-N conditions and **D** high-N conditions. Samples are represented by *taxon*-specific colored dots. Labels show the accession name of each genotype.

References

Anderson, O.D., Coleman-Derr, D., Gu, Y.Q., and Heath, S. (2010). Structural and transcriptional analysis of plant genes encoding the bifunctional lysine ketoglutarate reductase saccharopine dehydrogenase enzyme. BMC Plant Biol. **10**: 113.

Andrews, S. (2014). FastQC A Quality Control tool for High Throughput Sequence Data -Babraham Bioinformatics.

Arruda, P. and Barreto, P. (2020). Lysine Catabolism Through the Saccharopine Pathway: Enzymes and Intermediates Involved in Plant Responses to Abiotic and Biotic Stress. Front. Plant Sci. **11**: 587.

Aykanat, T., Thrower, F.P., and Heath, D.D. (2011). Rapid evolution of osmoregulatory function by modification of gene transcription in steelhead trout. Genetica **139**: 233–242.

Barneix, A.J. (2007). Physiology and biochemistry of source-regulated protein accumulation in the wheat grain. J. Plant Physiol. **164**: 581–590.

Bates, D., Mächler, M., Bolker, B., and Walker, S. (2015). Fitting Linear Mixed-Effects Models Using **lme4**. J. Stat. Softw. **67**: 1–48.

Beleggia, R., Omranian, N., Holtz, Y., Gioia, T., Fiorani, F., Nigro, F.M.,

Pecchioni, N., De Vita, P., Schurr, U., David, J.L., Nikoloski, Z., and Papa, R. (2021). Comparative Analysis Based on Transcriptomics and Metabolomics Data Reveal Differences between Emmer and Durum Wheat in Response to Nitrogen Starvation. Int. J. Mol. Sci. **22**: 4790.

Beleggia, R., Rau, D., Laido, G., Platani, C., Nigro, F., Fragasso, M., De Vita, P., Scossa, F., Fernie, A.R., Nikoloski, Z., and Papa, R. (2016). Evolutionary metabolomics reveals domestication-associated changes in tetraploid wheat kernels. Mol. Biol. Evol. **33**: 1740–1753.

Bellucci, E. et al. (2014). Decreased Nucleotide and Expression Diversity and Modified Coexpression Patterns Characterize Domestication in the Common Bean. Plant Cell **26**: 1901–1912.

Bitocchi, E., Rau, D., Bellucci, E., Rodriguez, M., Murgia, M.L., Gioia, T., Santo, D., Nanni, L., Attene, G., and Papa, R. (2017). Beans (Phaseolus ssp.) as a model for understanding crop evolution. Front. Plant Sci. **8**: 722.

Bogaard, A. et al. (2013). Crop manuring and intensive land management by Europe's first farmers. Proc. Natl. Acad. Sci. U. S. A. **110**: 12589–12594.

Bonnin, I., Prosperi, J.M., and Olivieri, I. (1996). Genetic Markers and Quantitative Genetic Variation in Medicago Truncatula (Leguminosae): A Comparative Analysis of Population Structure. Genetics **143**: 1795.

Bradshaw, A.D. (1965). Evolutionary Significance of Phenotypic Plasticity in Plants. Adv. Genet. **13**: 115–155.

Burgarella, C., Berger, A., Glémin, S., David, J., Terrier, N., Deu, M., and Pot, D. (2021). The Road to Sorghum Domestication: Evidence From Nucleotide Diversity and Gene Expression Patterns. Front. Plant Sci. **12**: 666075.

Chapman, M.A., Pashley, C.H., Wenzler, J., Hvala, J., Tang, S., Knapp, S.J., and Burke, J.M. (2008). A Genomic Scan for Selection Reveals Candidates for Genes Involved in the Evolution of Cultivated Sunflower (Helianthus annuus). Plant Cell **20**: 2931.

Cingolani, P., Platts, A., Wang, L.L., Coon, M., Nguyen, T., Wang, L., Land, S.J., Lu, X., and Ruden, D.M. (2012). A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of

963 Drosophila melanogaster strain w1118; iso-2; iso-3. Fly (Austin). **6**: 80.

964 **Civáň, P., Ivaničová, Z., and Brown, T.A.** (2013). Reticulated Origin of
965 Domesticated Emmer Wheat Supports a Dynamic Model for the Emergence of
966 Agriculture in the Fertile Crescent. PLoS One **8**: e81955.

967 **Curci, P.L., Aiese Cigliano, R., Zuluaga, D.L., Janni, M., Sanseverino, W., and**
968 **Sonnante, G.** (2017). Transcriptomic response of durum wheat to nitrogen
969 starvation. Sci. Rep. **7**: 1–14.

970 **Danecek, P., Auton, A., Abecasis, G., Albers, C.A., Banks, E., DePristo, M.A.,**
971 **Handsaker, R.E., Lunter, G., Marth, G.T., Sherry, S.T., McVean, G., and**
972 **Durbin, R.** (2011). The variant call format and VCFtools. Bioinformatics **27**:
973 2156.

974 **Danecek, P., Bonfield, J.K., Liddle, J., Marshall, J., Ohan, V., Pollard, M.O.,**
975 **Whitwham, A., Keane, T., McCarthy, S.A., Davies, R.M., and Li, H.** (2021).
976 Twelve years of SAMtools and BCFtools. Gigascience **10**: 1–4.

977 **David, J., Holtz, Y., Ranwez, V., Santoni, S., Sarah, G., Ardisson, M., Poux, G.,**
978 **Choulet, F., Genthon, C., Roumet, P., and Tavaud-Pirra, M.** (2014).
979 Genotyping by sequencing transcriptomes in an evolutionary pre-breeding
980 durum wheat population. Mol. Breed. **34**: 1531–1548.

981 **Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut,**
982 **P., Chaisson, M., and Gingeras, T.R.** (2013). STAR: ultrafast universal RNA-
983 seq aligner. Bioinformatics **29**: 15–21.

984 **Donald, C.M.** (1968). The breeding of crop ideotypes. Euphytica **17**: 385–403.

985 **Fontaine, J.X., Ravel, C., Pageau, K., Heumez, E., Dubois, F., Hirel, B., and Le**
986 **Gouis, J.** (2009). A quantitative genetic study for elucidating the contribution of
987 glutamine synthetase, glutamate dehydrogenase and other nitrogen-related
988 physiological traits to the agronomic performance of common wheat. Theor.
989 Appl. Genet. **119**: 645–662.

990 **Fréville, H., Montazeaud, G., Forst, E., David, J., Papa, R., and Tenaillon, M.I.**
991 (2022). Shift in beneficial interactions during crop evolution. Evol. Appl. **15**: 905–
992 918.

993 **Frizzi, A., Huang, S., Gilbertson, L.A., Armstrong, T.A., Luethy, M.H., and**

Malvar, T.M. (2008). Modifying lysine biosynthesis and catabolism in corn with a single bifunctional expression/silencing transgene cassette. Plant Biotechnol. J. **6**: 13–21.

Garcia-Gonzalez, F., Simmons, L.W., Tomkins, J.L., Kotiaho, J.S., and Evans, J.P. (2012). COMPARING EVOLVABILITIES: COMMON ERRORS SURROUNDING THE CALCULATION AND USE OF COEFFICIENTS OF ADDITIVE GENETIC VARIATION. Evolution (N. Y). **66**: 2341–2349.

Gioia, T., Nagel, K.A., Beleggia, R., Fragasso, M., Ficco, D.B.M., Pieruschka, R., De Vita, P., Fiorani, F., and Papa, R. (2015). Impact of domestication on the phenotypic architecture of durum wheat under contrasting nitrogen fertilization. J. Exp. Bot. **66**: 5519–5530.

Hansen, T.F., Pélabon, C., and Houle, D. (2011). Heritability is not Evolvability. Evol. Biol. 2011 383 **38**: 258–277.

Haudry, A., Cenci, A., Ravel, C., Bataillon, T., Brunel, D., Poncet, C., Hochu, I., Poirier, S., Santoni, S., Glémin, S., and David, J. (2007). Grinding up Wheat: A Massive Loss of Nucleotide Diversity Since Domestication. Mol. Biol. Evol. **24**: 1506–1517.

Hawkesford, M.J. and Griffiths, S. (2019). Exploiting genetic variation in nitrogen use efficiency for cereal crop improvement. Curr. Opin. Plant Biol. **49**: 35–42.

Hoagland, D.R. and Arnon, D.I. (1950). The water-culture method for growing plants without soil. Circ. Calif. Agric. Exp. Stn. **347**.

Houle, D. (1992). Comparing evolvability and variability of quantitative traits. Genetics **130**: 195–204.

Houmard, N.M., Mainville, J.L., Bonin, C.P., Huang, S., Luethy, M.H., and Malvar, T.M. (2007). High-lysine corn generated by endosperm-specific suppression of lysine catabolism using RNAi. Plant Biotechnol. J. **5**: 605–614.

Howarth, J.R. et al. (2008). Co-ordinated expression of amino acid metabolism in response to N and S deficiency during wheat grain filling. J. Exp. Bot. **59**: 3675–3689.

Kavi Kishor, P.B., Suravajhala, P., Rathnagiri, P., and Sreenivasulu, N. (2022). Intriguing Role of Proline in Redox Potential Conferring High Temperature

Stress Tolerance. Front. Plant Sci. **13**: 867531.

Kohn, M.H., Shapiro, J., and Wu, C.I. (2008). Decoupled differentiation of gene expression and coding sequence among Drosophila populations. Genes Genet. Syst. **83**: 265–273.

Kolde, R. (2019). pheatmap: Pretty Heatmaps. R Packag. version 1.0.12.

Korneliussen, T.S., Albrechtsen, A., and Nielsen, R. (2014). ANGSD: Analysis of Next Generation Sequencing Data. BMC Bioinformatics **15**: 1–13.

Laidò, G., Mangini, G., Taranto, F., Gadaleta, A., Blanco, A., Cattivelli, L., Marone, D., Mastrangelo, A.M., Papa, R., and De Vita, P. (2013). Genetic Diversity and Population Structure of Tetraploid Wheats (Triticum turgidum L.) Estimated by SSR, DArT and Pedigree Data. PLoS One **8**: e67280.

Laidò, G., Marone, D., Russo, M.A., Colecchia, S.A., Mastrangelo, A.M., Vita, P. De, and Papa, R. (2014). Linkage Disequilibrium and Genome-Wide Association Mapping in Tetraploid Wheat (Triticum turgidum L.). PLoS One **9**: e95211.

Laitinen, R.A.E. and Nikoloski, Z. (2019). Genetic basis of plasticity in plants. J. Exp. Bot. **70**: 795–804.

Leinonen, T., McCairns, R.J.S., O'Hara, R.B., and Merilä, J. (2013). QST -FST comparisons: Evolutionary and ecological insights from genomic heterogeneity. Nat. Rev. Genet. **14**: 179–190.

Lemmon, Z.H., Bukowski, R., Sun, Q., and Doebley, J.F. (2014). The Role of cis Regulatory Evolution in Maize Domestication. PLOS Genet. **10**: e1004745.

Levy, A.A. and Feldman, M. (2022). Evolution and origin of bread wheat. Plant Cell **34**: 2549–2567.

Liu, W., Chen, L., Zhang, S., Hu, F., Wang, Z., Lyu, J., Wang, B., Xiang, H., Zhao, R., Tian, Z., Ge, S., and Wang, W. (2019). Decrease of gene expression diversity during domestication of animals and plants. BMC Evol. Biol. **19**: 1–11.

Long, X., Liu, Q., Chan, M., Wang, Q., and Sun, S.S.M. (2013). Metabolic engineering and profiling of rice with increased lysine. Plant Biotechnol. J. **11**: 490–501.

Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. **15**: 550.

Ludemann, C.I., Gruere, A., Heffer, P., and Dobermann, A. (2022). Global data on fertilizer use by crop and by country. Sci. Data 2022 91 **9**: 1–8.

Luo, M.C., Yang, Z.L., You, F.M., Kawahara, T., Waines, J.G., and Dvorak, J. (2007). The structure of wild and domesticated emmer wheat populations, gene flow between them, and the site of emmer domestication. Theor. Appl. Genet. **114**: 947–959.

Maccaferri, M. et al. (2019). Durum wheat genome highlights past domestication signatures and future improvement targets. Nat. Genet. 2019 515 **51**: 885–895.

Martín-Robles, N., Lehmann, A., Seco, E., Aroca, R., Rillig, M.C., and Milla, R. (2018). Impacts of domestication on the arbuscular mycorrhizal symbiosis of 27 crop species. New Phytol. **218**: 322–334.

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal **17**: 10.

Michaletti, A., Naghavi, M.R., Toorchi, M., Zolla, L., and Rinalducci, S. (2018). Metabolomics and proteomics reveal drought-stress responses of leaf tissues from spring-wheat. Sci. Reports 2018 81 **8**: 1–18.

Miflin, B.J. and Habash, D.Z. (2002). The role of glutamine synthetase and glutamate dehydrogenase in nitrogen assimilation and possibilities for improvement in the nitrogen utilization of crops. J. Exp. Bot. **53**: 979–987.

Oleksiak, M.F., Churchill, G.A., and Crawford, D.L. (2002). Variation in gene expression within and among natural populations. Nat. Genet. 2002 322 **32**: 261–266.

Oliveira, H.R., Jacocks, L., Czajkowska, B.I., Kennedy, S.L., and Brown, T.A. (2020). Multiregional origins of the domesticated tetraploid wheats. PLoS One **15**: e0227148.

Ondov, B.D., Treangen, T.J., Melsted, P., Mallonee, A.B., Bergman, N.H., Koren, S., and Phillippy, A.M. (2016). Mash: Fast genome and metagenome distance estimation using MinHash. Genome Biol. **17**: 1–14.

Oostra, V., Saastamoinen, M., Zwaan, B.J., and Wheat, C.W. (2018). Strong phenotypic plasticity limits potential for evolutionary responses to climate change. Nat. Commun. 2018 91 **9**: 1–11.

Page, A., Gibson, J., Meyer, R.S., and Chapman, M.A. (2019). Eggplant Domestication: Pervasive Gene Flow, Feralization, and Transcriptomic Divergence. Mol. Biol. Evol. **36**: 1359–1372.

Patro, R., Duggal, G., Love, M.I., Irizarry, R.A., and Kingsford, C. (2017). Salmon provides fast and bias-aware quantification of transcript expression. Nat. Methods 2017 144 **14**: 417–419.

Plett, D.C., Holtham, L.R., Okamoto, M., and Garnett, T.P. (2018). Nitrate uptake and its regulation in relation to improving nitrogen use efficiency in cereals. Semin. Cell Dev. Biol. **74**: 97–104.

Porter, S.S. and Sachs, J.L. (2020). Agriculture and the Disruption of Plant–Microbial Symbiosis. Trends Ecol. Evol. **35**: 426–439.

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A.R., Bender, D., Maller, J., Sklar, P., De Bakker, P.I.W., Daly, M.J., and Sham, P.C. (2007). PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. Am. J. Hum. Genet. **81**: 559.

R Core Team (2022). R: A language and environment for statistical computing. R Foundation Statistical Computing.

Rapaport, F., Khanin, R., Liang, Y., Pirun, M., Krek, A., Zumbo, P., Mason, C.E., Socci, N.D., and Betel, D. (2013). Comprehensive evaluation of differential gene expression analysis methods for RNA-seq data. Genome Biol **14**: R95.

Roberge, C., Guderley, H., and Bernatchez, L. (2007). Genomewide Identification of Genes Under Directional Selection: Gene Transcription QST Scan in Diverging Atlantic Salmon Subpopulations. Genetics **177**: 1011.

Robinson, M.D., McCarthy, D.J., and Smyth, G.K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics **26**: 139–40.

Rockström, J. et al. (2009). A safe operating space for humanity. Nat. 2009 4617263 **461**: 472–475.

Sauvage, C., Rau, A., Aichholz, C., Chadoeuf, J., Sarah, G., Ruiz, M., Santoni, S., Causse, M., David, J., and Glémin, S. (2017). Domestication rewired gene expression and nucleotide diversity patterns in tomato. Plant J. **91**: 631–645.

Savary, S., Willocquet, L., Pethybridge, S.J., Esker, P., McRoberts, N., and Nelson, A. (2019). The global burden of pathogens and pests on major food crops. Nat. Ecol. Evol. 2019 33 **3**: 430–439.

Smyth, G.K. (2005). limma: Linear Models for Microarray Data. Bioinforma. Comput. Biol. Solut. Using R Bioconductor: 397–420.

Sultana, N., Islam, S., Juhasz, A., Yang, R., She, M., Alhabbar, Z., Zhang, J., and Ma, W. (2020). Transcriptomic Study for Identification of Major Nitrogen Stress Responsive Genes in Australian Bread Wheat Cultivars. Front. Genet. **11**: 583785.

Swanson-Wagner, R., Briskine, R., Schaefer, R., Hufford, M.B., Ross-Ibarra, J., Myers, C.L., Tiffin, P., and Springer, N.M. (2012). Reshaping of the maize transcriptome by domestication. Proc. Natl. Acad. Sci. U. S. A. **109**: 11878–11883.

Tajima, F. (1983). Evolutionary relationship of DNA sequences in finite populations. Genetics **105**: 437–460.

Taranto, F., D'Agostino, N., Rodriguez, M., Pavan, S., Minervini, A.P., Pecchioni, N., Papa, R., and De Vita, P. (2020). Whole Genome Scan Reveals Molecular Signatures of Divergence and Selection Related to Important Traits in Durum Wheat Germplasm. Front. Genet. **11**: 217.

Tian, T., Liu, Y., Yan, H., You, Q., Yi, X., Du, Z., Xu, W., and Su, Z. (2017). agriGO v2.0: a GO analysis toolkit for the agricultural community, 2017 update. Nucleic Acids Res. **45**: W122–W129.

Valente, J., Gerin, F., Mini, A., Richard, R., Gouis, J. Le, Prigent-Combaret, C., and Moënne-Loccoz, Y. (2023). Symbiotic Variations among Wheat Genotypes and Detection of Quantitative Trait Loci for Molecular Interaction with Auxin-Producing Azospirillum PGPR. Microorg. 2023, Vol. 11, Page 1615 **11**: 1615.

Vigouroux, Y., McMullen, M., Hittinger, C.T., Houchins, K., Schulz, L., Kresovich, S., Matsuoka, Y., and Doebley, J. (2002). Identifying genes of

agronomic importance in maize by screening microsatellites for evidence of selection during domestication. Proc. Natl. Acad. Sci. U. S. A. **99**: 9650–9655.

Watterson, G.A. (1975). On the number of segregating sites in genetical models without recombination. Theor. Popul. Biol. **7**: 256–276.

Whitehead, A. and Crawford, D.L. (2006a). Neutral and adaptive variation in gene expression. Proc. Natl. Acad. Sci. U. S. A. **103**: 5425–5430.

Whitehead, A. and Crawford, D.L. (2006b). Variation within and among species in gene expression: raw material for evolution. Mol. Ecol. **15**: 1197–1211.

Whitlock, M.C. and Guillaume, F. (2009). Testing for Spatially Divergent Selection: Comparing QST to FST. Genetics **183**: 1055–1063.

Wood, S.W. and Cowie, A. (2004). A Review of Greenhouse Gas Emission Factors for Fertiliser Production.

Zhang, Z.H., Jhaveri, D.J., Marshall, V.M., Bauer, D.C., Edson, J., Narayanan, R.K., Robinson, G.J., Lundberg, A.E., Bartlett, P.F., Wray, N.R., and Zhao, Q.-Y. (2014). A comparative study of techniques for differential expression analysis on RNA-Seq data. PLoS One **9**: e103207.

Zhu, T., Wang, L., Rodriguez, J.C., Deal, K.R., Avni, R., Distelfeld, A., McGuire, P.E., Dvorak, J., and Luo, M.-C. (2019). Improved Genome Sequence of Wild Emmer Wheat Zavitan with the Aid of Optical Maps. G3 Genes, Genomes, Genet. **9**: 619–624.

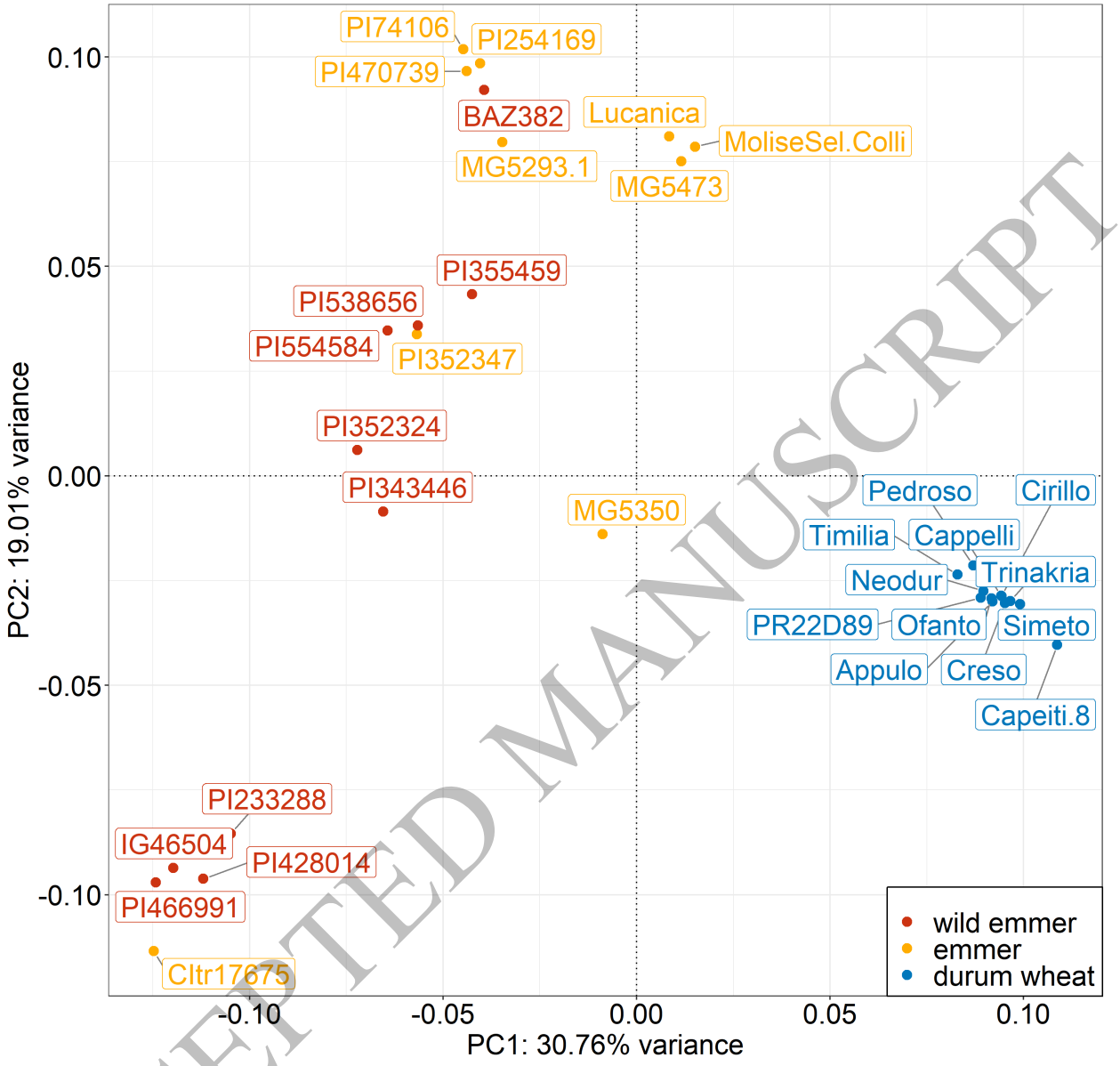

Figure 1: Principal component analysis (PCA) of 32 wheat genotypes based on single-nucleotide polymorphisms (SNPs). The first two principal components (PC1 and PC2) are shown. The three colors represent different *taxa*. Labels show the accession name of each genotype.

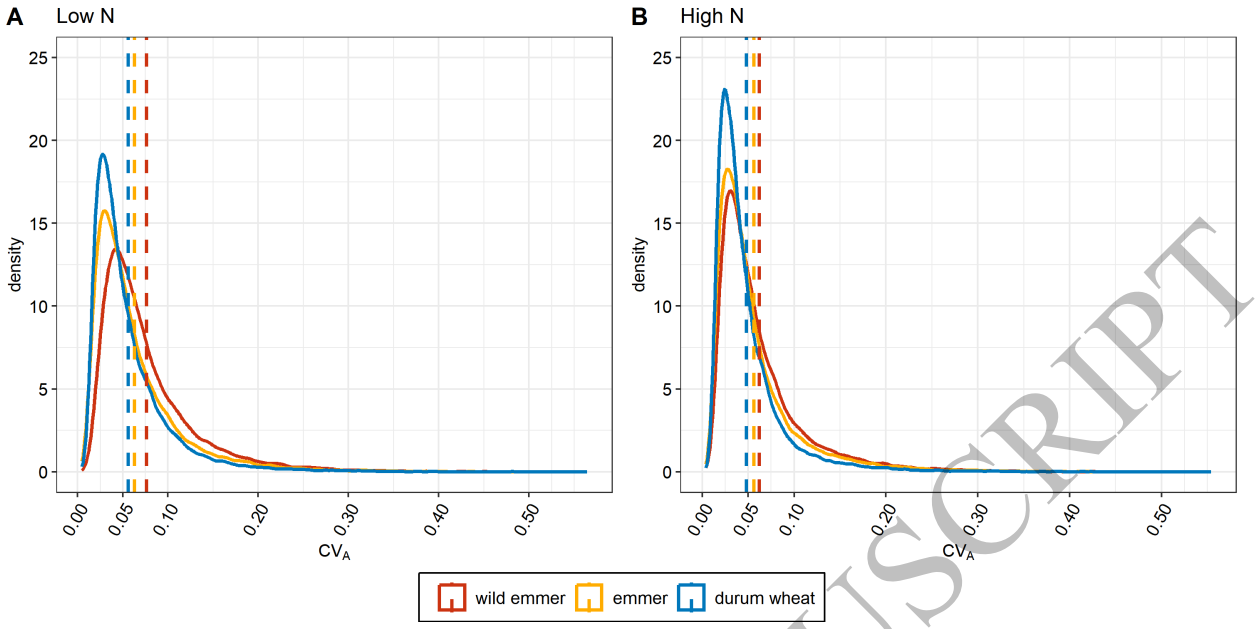

Figure 2: Density plots of the additive coefficient of variation (CV_A) in the three wheat *taxa*. Comparison of the estimated density functions of the CV_A in gene expression, calculated using all 32,358 genes. **A** Low-N conditions. **B** High-N conditions. Dashed lines represent the averaged CV_A value, colored according to the different *taxa*.

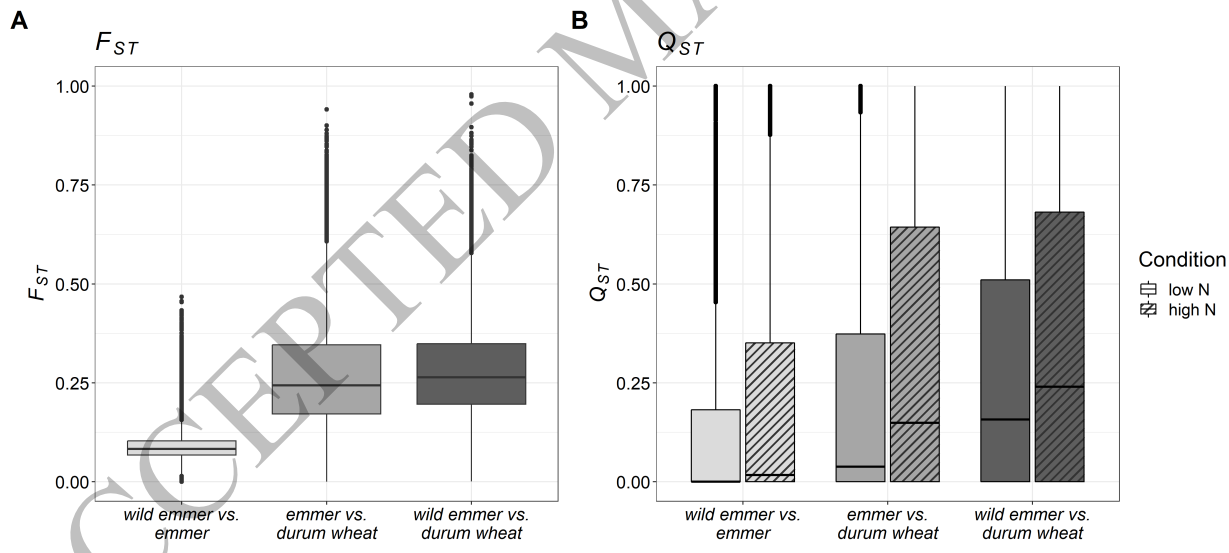

Figure 3: F_{ST} and Q_{ST} distributions. **A** Boxplots showing the gene locus F_{ST} distribution for every subspecies pairwise comparison. **B** Boxplots showing the gene expression Q_{ST} distribution for every subspecies pairwise comparison under low-N and high-N conditions, represented by empty and hatched grayscale bars, respectively. The borders of the box represent the 25th and 75th percentiles. The horizontal line in the middle of the box represents the median. Whiskers extend to the minimum and maximum values, unless a

point exceeds 1.5 times the interquartile range, in which case the whisker extends to this value and values beyond are plotted as individual points (outliers).

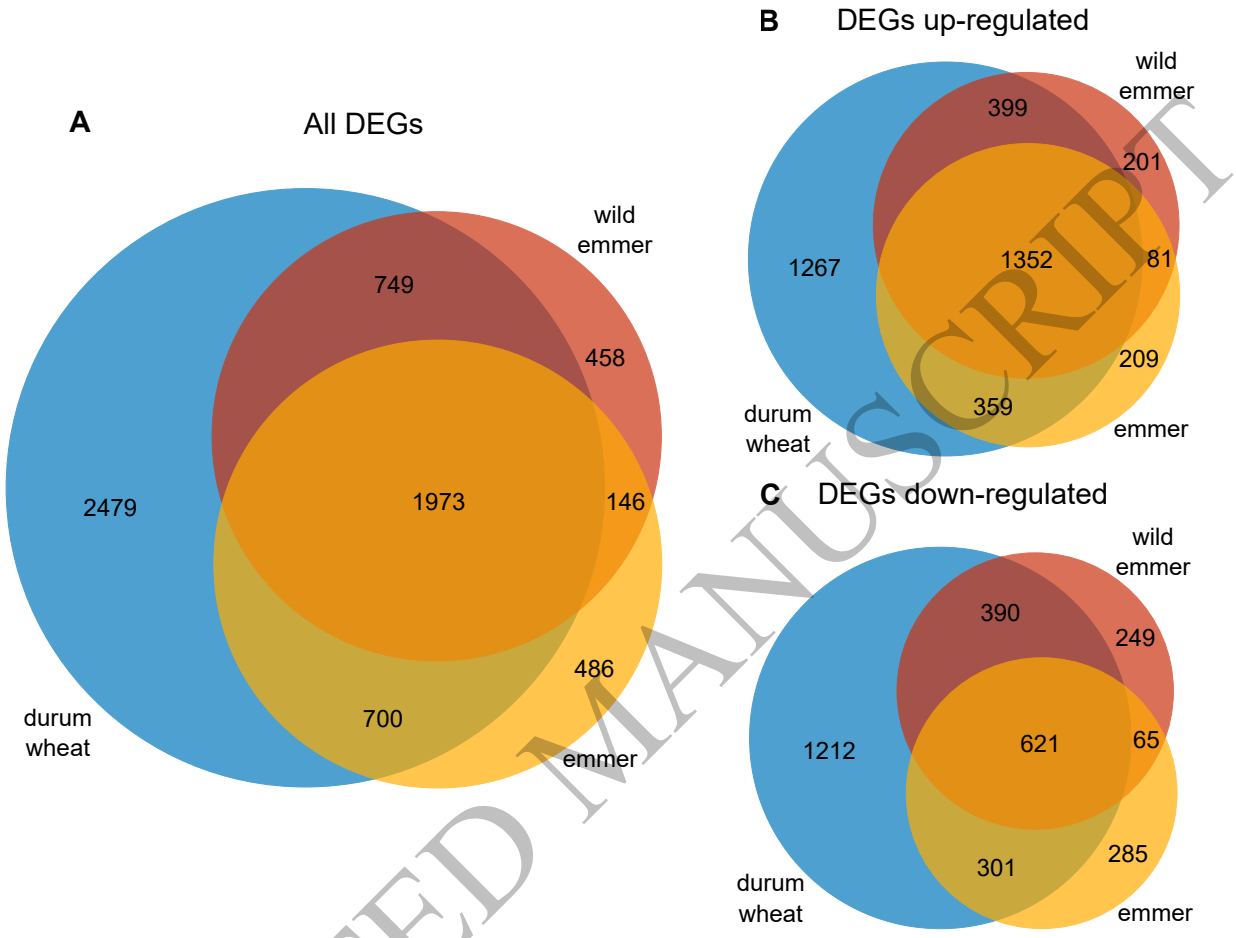

Figure 4: Venn diagrams showing differentially expressed genes (DEGs) when comparing high-N and low-N conditions within each subspecies. A Total set of DEGs. **B** Upregulated DEGs only. **C** Downregulated DEGs only.

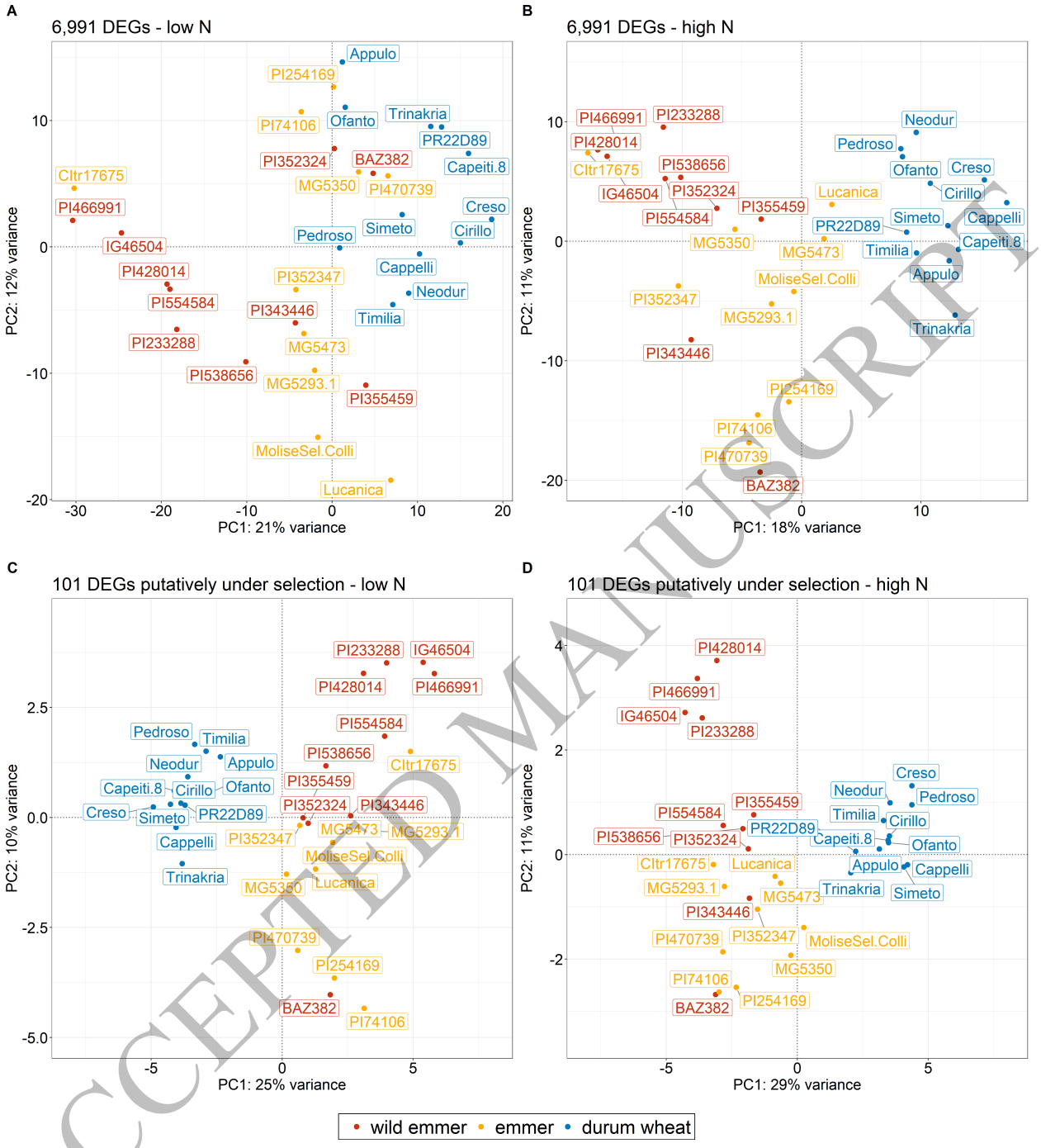

Figure 5: Principal component analysis (PCA) of 32 wheat genotypes based on expression data of differentially expressed genes (DEGs) when comparing high-N and low-N conditions within each subspecies. A,B Plots based on all 6,991 DEGs (not filtered): **A** low-N conditions and **B** high-N conditions. **C,D** Plots based on 101 DEGs that are also putatively under selection: **C** low-N conditions and **D** high-N conditions. Samples are represented by *taxon*-specific colored dots. Labels show the accession name of each genotype.

Parsed Citations

Anderson, O.D., Coleman-Derr, D., Gu, Y.Q., and Heath, S. (2010). Structural and transcriptional analysis of plant genes encoding the bifunctional lysine ketoglutarate reductase saccharopine dehydrogenase enzyme. BMC Plant Biol. 10: 113.
Google Scholar: Author Only Title Only Author and Title

Andrews, S. (2014). FastQC A Quality Control tool for High Throughput Sequence Data -Babraham Bioinformatics.

Arruda, P. and Barreto, P. (2020). Lysine Catabolism Through the Saccharopine Pathway: Enzymes and Intermediates Involved in Plant Responses to Abiotic and Biotic Stress. Front. Plant Sci. 11: 587.
Google Scholar: Author Only Title Only Author and Title

Aykanat, T., Thrower, F.P., and Heath, D.D. (2011). Rapid evolution of osmoregulatory function by modification of gene transcription in steelhead trout. Genetica 139: 233–242.
Google Scholar: Author Only Title Only Author and Title

Barneix, A.J. (2007). Physiology and biochemistry of source-regulated protein accumulation in the wheat grain. J. Plant Physiol. 164: 581–590.
Google Scholar: Author Only Title Only Author and Title

Bates, D., Mächler, M., Bolker, B., and Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. J. Stat. Softw. 67: 1–48.
Google Scholar: Author Only Title Only Author and Title

Beleggia, R., Omranian, N., Holtz, Y., Gioia, T., Fiorani, F., Nigro, F.M., Pecchioni, N., De Vita, P., Schurr, U., David, J.L., Nikoloski, Z., and Papa, R. (2021). Comparative Analysis Based on Transcriptomics and Metabolomics Data Reveal Differences between Emmer and Durum Wheat in Response to Nitrogen Starvation. Int. J. Mol. Sci. 22: 4790.
Google Scholar: Author Only Title Only Author and Title

Beleggia, R., Rau, D., Laido, G., Platani, C., Nigro, F., Fragasso, M., De Vita, P., Scossa, F., Fernie, A.R., Nikoloski, Z., and Papa, R. (2016). Evolutionary metabolomics reveals domestication-associated changes in tetraploid wheat kernels. Mol. Biol. Evol. 33: 1740–1753.
Google Scholar: Author Only Title Only Author and Title

Bellucci, E. et al. (2014). Decreased Nucleotide and Expression Diversity and Modified Coexpression Patterns Characterize Domestication in the Common Bean. Plant Cell 26: 1901–1912.
Google Scholar: Author Only Title Only Author and Title

Bitocchi, E., Rau, D., Bellucci, E., Rodriguez, M., Murgia, M.L., Gioia, T., Santo, D., Nanni, L., Attene, G., and Papa, R. (2017). Beans (Phaseolus ssp.) as a model for understanding crop evolution. Front. Plant Sci. 8: 722.
Google Scholar: Author Only Title Only Author and Title

Bogaard, A. et al. (2013). Crop manuring and intensive land management by Europe's first farmers. Proc. Natl. Acad. Sci. U. S. A. 110: 12589–12594.
Google Scholar: Author Only Title Only Author and Title

Bonnin, I., Prosperi, J.M., and Olivieri, I. (1996). Genetic Markers and Quantitative Genetic Variation in Medicago Truncatula (Leguminosae): A Comparative Analysis of Population Structure. Genetics 143: 1795.
Google Scholar: Author Only Title Only Author and Title

Bradshaw, A.D. (1965). Evolutionary Significance of Phenotypic Plasticity in Plants. Adv. Genet. 13: 115–155.
Google Scholar: Author Only Title Only Author and Title

Burgarella, C., Berger, A., Glémin, S., David, J., Terrier, N., Deu, M., and Pot, D. (2021). The Road to Sorghum Domestication: Evidence From Nucleotide Diversity and Gene Expression Patterns. Front. Plant Sci. 12: 666075.
Google Scholar: Author Only Title Only Author and Title

Chapman, M.A., Pashley, C.H., Wenzler, J., Hvala, J., Tang, S., Knapp, S.J., and Burke, J.M. (2008). A Genomic Scan for Selection Reveals Candidates for Genes Involved in the Evolution of Cultivated Sunflower (Helianthus annuus). Plant Cell 20: 2931.
Google Scholar: Author Only Title Only Author and Title

Cingolani, P., Platts, A., Wang, L.L., Coon, M., Nguyen, T., Wang, L., Land, S.J., Lu, X., and Ruden, D.M. (2012). A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. Fly (Austin). 6: 80.
Google Scholar: Author Only Title Only Author and Title

Civáň, P., Ivaničová, Z., and Brown, T.A. (2013). Reticulated Origin of Domesticated Emmer Wheat Supports a Dynamic Model for the Emergence of Agriculture in the Fertile Crescent. PLoS One 8: e81955.
Google Scholar: Author Only Title Only Author and Title

Curci, P.L., Aiese Cigliano, R., Zuluaga, D.L., Janni, M., Sanseverino, W., and Sonnante, G. (2017). Transcriptomic response of

durum wheat to nitrogen starvation. Sci. Rep. 7: 1–14.
Google Scholar: Author Only Title Only Author and Title

Danecek, P., Auton, A., Abecasis, G., Albers, C.A., Banks, E., DePristo, M.A., Handsaker, R.E., Lunter, G., Marth, G.T., Sherry, S.T., McVean, G., and Durbin, R. (2011). The variant call format and VCFtools. Bioinformatics 27: 2156.
Google Scholar: Author Only Title Only Author and Title

Danecek, P., Bonfield, J.K., Liddle, J., Marshall, J., Ohan, V., Pollard, M.O., Whitwham, A., Keane, T., McCarthy, S.A., Davies, R.M., and Li, H. (2021). Twelve years of SAMtools and BCFtools. Gigascience 10: 1–4.
Google Scholar: Author Only Title Only Author and Title

David, J., Holtz, Y., Ranwez, V., Santoni, S., Sarah, G., Ardisson, M., Poux, G., Choulet, F., Genthon, C., Roumet, P., and Tavaud-Pirra, M. (2014). Genotyping by sequencing transcriptomes in an evolutionary pre-breeding durum wheat population. Mol. Breed. 34: 1531–1548.
Google Scholar: Author Only Title Only Author and Title

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29: 15–21.
Google Scholar: Author Only Title Only Author and Title

Donald, C.M. (1968). The breeding of crop ideotypes. Euphytica 17: 385–403.
Google Scholar: Author Only Title Only Author and Title

Fontaine, J.X., Ravel, C., Pageau, K., Heumez, E., Dubois, F., Hirel, B., and Le Gouis, J. (2009). A quantitative genetic study for elucidating the contribution of glutamine synthetase, glutamate dehydrogenase and other nitrogen-related physiological traits to the agronomic performance of common wheat. Theor. Appl. Genet. 119: 645–662.
Google Scholar: Author Only Title Only Author and Title

Fréville, H., Montazeaud, G., Forst, E., David, J., Papa, R., and Tenaillon, M.I. (2022). Shift in beneficial interactions during crop evolution. Evol. Appl. 15: 905–918.
Google Scholar: Author Only Title Only Author and Title

Frizzi, A., Huang, S., Gilbertson, L.A., Armstrong, T.A., Luethy, M.H., and Malvar, T.M. (2008). Modifying lysine biosynthesis and catabolism in corn with a single bifunctional expression/silencing transgene cassette. Plant Biotechnol. J. 6: 13–21.
Google Scholar: Author Only Title Only Author and Title

Garcia-Gonzalez, F., Simmons, L.W., Tomkins, J.L., Kotiaho, J.S., and Evans, J.P. (2012). COMPARING EVOLVABILITIES: COMMON ERRORS SURROUNDING THE CALCULATION AND USE OF COEFFICIENTS OF ADDITIVE GENETIC VARIATION. Evolution (N. Y). 66: 2341–2349.
Google Scholar: Author Only Title Only Author and Title

Gioia, T., Nagel, K.A., Beleggia, R., Fragasso, M., Ficco, D.B.M., Pieruschka, R., De Vita, P., Fiorani, F., and Papa, R. (2015). Impact of domestication on the phenotypic architecture of durum wheat under contrasting nitrogen fertilization. J. Exp. Bot. 66: 5519–5530.
Google Scholar: Author Only Title Only Author and Title

Hansen, T.F., Pélabon, C., and Houle, D. (2011). Heritability is not Evolvability. Evol. Biol. 2011 383 38: 258–277.
Google Scholar: Author Only Title Only Author and Title

Haudry, A., Cenci, A., Ravel, C., Bataillon, T., Brunel, D., Poncet, C., Hochu, I., Poirier, S., Santoni, S., Glémin, S., and David, J. (2007). Grinding up Wheat: A Massive Loss of Nucleotide Diversity Since Domestication. Mol. Biol. Evol. 24: 1506–1517.
Google Scholar: Author Only Title Only Author and Title

Hawkesford, M.J. and Griffiths, S. (2019). Exploiting genetic variation in nitrogen use efficiency for cereal crop improvement. Curr. Opin. Plant Biol. 49: 35–42.
Google Scholar: Author Only Title Only Author and Title

Hoagland, D.R. and Arnon, D.I. (1950). The water-culture method for growing plants without soil. Circ. Calif. Agric. Exp. Stn. 347.
Google Scholar: Author Only Title Only Author and Title

Houle, D. (1992). Comparing evolvability and variability of quantitative traits. Genetics 130: 195–204.
Google Scholar: Author Only Title Only Author and Title

Houmard, N.M., Mainville, J.L., Bonin, C.P., Huang, S., Luethy, M.H., and Malvar, T.M. (2007). High-lysine corn generated by endosperm-specific suppression of lysine catabolism using RNAi. Plant Biotechnol. J. 5: 605–614.
Google Scholar: Author Only Title Only Author and Title

Howarth, J.R. et al. (2008). Co-ordinated expression of amino acid metabolism in response to N and S deficiency during wheat grain filling. J. Exp. Bot. 59: 3675–3689.
Google Scholar: Author Only Title Only Author and Title

Kavi Kishor, P.B., Suravajhala, P., Rathnagiri, P., and Sreenivasulu, N. (2022). Intriguing Role of Proline in Redox Potential Conferring High Temperature Stress Tolerance. Front. Plant Sci. 13: 867531.
Google Scholar: Author Only Title Only Author and Title

Kohn, M.H., Shapiro, J., and Wu, C.I. (2008). Decoupled differentiation of gene expression and coding sequence among Drosophila populations. Genes Genet. Syst. 83: 265–273.
Google Scholar: Author Only Title Only Author and Title

Kolde, R. (2019). pheatmap: Pretty Heatmaps. R Packag. version 1.0.12.
Google Scholar: Author Only Title Only Author and Title

Korneliussen, T.S., Albrechtsen, A., and Nielsen, R. (2014). ANGSD: Analysis of Next Generation Sequencing Data. BMC Bioinformatics 15: 1–13.
Google Scholar: Author Only Title Only Author and Title

Laidò, G., Mangini, G., Taranto, F., Gadaleta, A., Blanco, A., Cattivelli, L., Marone, D., Mastrangelo, A.M., Papa, R., and De Vita, P. (2013). Genetic Diversity and Population Structure of Tetraploid Wheats (Triticum turgidum L.) Estimated by SSR, DArT and Pedigree Data. PLoS One 8: e67280.
Google Scholar: Author Only Title Only Author and Title

Laidò, G., Marone, D., Russo, M.A., Colecchia, S.A., Mastrangelo, A.M., Vita, P. De, and Papa, R. (2014). Linkage Disequilibrium and Genome-Wide Association Mapping in Tetraploid Wheat (Triticum turgidum L.). PLoS One 9: e95211.
Google Scholar: Author Only Title Only Author and Title

Laitinen, R.A.E. and Nikoloski, Z. (2019). Genetic basis of plasticity in plants. J. Exp. Bot. 70: 795–804.
Google Scholar: Author Only Title Only Author and Title

Leinonen, T., McCairns, R.J.S., O'Hara, R.B., and Merilä, J. (2013). QST -FST comparisons: Evolutionary and ecological insights from genomic heterogeneity. Nat. Rev. Genet. 14: 179–190.
Google Scholar: Author Only Title Only Author and Title

Lemmon, Z.H., Bukowski, R., Sun, Q., and Doebley, J.F. (2014). The Role of cis Regulatory Evolution in Maize Domestication. PLOS Genet. 10: e1004745.
Google Scholar: Author Only Title Only Author and Title

Levy, A.A. and Feldman, M. (2022). Evolution and origin of bread wheat. Plant Cell 34: 2549–2567.
Google Scholar: Author Only Title Only Author and Title

Liu, W., Chen, L., Zhang, S., Hu, F., Wang, Z., Lyu, J., Wang, B., Xiang, H., Zhao, R., Tian, Z., Ge, S., and Wang, W. (2019). Decrease of gene expression diversity during domestication of animals and plants. BMC Evol. Biol. 19: 1–11.
Google Scholar: Author Only Title Only Author and Title

Long, X., Liu, Q., Chan, M., Wang, Q., and Sun, S.S.M. (2013). Metabolic engineering and profiling of rice with increased lysine. Plant Biotechnol. J. 11: 490–501.
Google Scholar: Author Only Title Only Author and Title

Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 15: 550.
Google Scholar: Author Only Title Only Author and Title

Ludemann, C.I., Gruere, A., Heffer, P., and Dobermann, A. (2022). Global data on fertilizer use by crop and by country. Sci. Data 2022 91 9: 1–8.
Google Scholar: Author Only Title Only Author and Title

Luo, M.C., Yang, Z.L., You, F.M., Kawahara, T., Waines, J.G., and Dvorak, J. (2007). The structure of wild and domesticated emmer wheat populations, gene flow between them, and the site of emmer domestication. Theor. Appl. Genet. 114: 947–959.
Google Scholar: Author Only Title Only Author and Title

Maccaferri, M. et al. (2019). Durum wheat genome highlights past domestication signatures and future improvement targets. Nat. Genet. 2019 515 51: 885–895.
Google Scholar: Author Only Title Only Author and Title

Martín-Robles, N., Lehmann, A., Seco, E., Aroca, R., Rillig, M.C., and Milla, R. (2018). Impacts of domestication on the arbuscular mycorrhizal symbiosis of 27 crop species. New Phytol. 218: 322–334.
Google Scholar: Author Only Title Only Author and Title

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal 17: 10.
Google Scholar: Author Only Title Only Author and Title

Michaletti, A., Naghavi, M.R., Toorchi, M., Zolla, L., and Rinalducci, S. (2018). Metabolomics and proteomics reveal drought-stress responses of leaf tissues from spring-wheat. Sci. Reports 2018 81 8: 1–18.

Google Scholar: Author Only Title Only Author and Title

Miflin, B.J. and Habash, D.Z. (2002). The role of glutamine synthetase and glutamate dehydrogenase in nitrogen assimilation and possibilities for improvement in the nitrogen utilization of crops. J. Exp. Bot. 53: 979–987.
Google Scholar: Author Only Title Only Author and Title

Oleksiak, M.F., Churchill, G.A., and Crawford, D.L. (2002). Variation in gene expression within and among natural populations. Nat. Genet. 2002 322 32: 261–266.
Google Scholar: Author Only Title Only Author and Title

Oliveira, H.R., Jacocks, L., Czajkowska, B.I., Kennedy, S.L., and Brown, T.A. (2020). Multiregional origins of the domesticated tetraploid wheats. PLoS One 15: e0227148.
Google Scholar: Author Only Title Only Author and Title

Ondov, B.D., Treangen, T.J., Melsted, P., Mallonee, A.B., Bergman, N.H., Koren, S., and Phillippy, A.M. (2016). Mash: Fast genome and metagenome distance estimation using MinHash. Genome Biol. 17: 1–14.
Google Scholar: Author Only Title Only Author and Title

Oostra, V., Saastamoinen, M., Zwaan, B.J., and Wheat, C.W. (2018). Strong phenotypic plasticity limits potential for evolutionary responses to climate change. Nat. Commun. 2018 91 9: 1–11.
Google Scholar: Author Only Title Only Author and Title

Page, A., Gibson, J., Meyer, R.S., and Chapman, M.A. (2019). Eggplant Domestication: Pervasive Gene Flow, Feralization, and Transcriptomic Divergence. Mol. Biol. Evol. 36: 1359–1372.
Google Scholar: Author Only Title Only Author and Title

Patro, R., Duggal, G., Love, M.I., Irizarry, R.A., and Kingsford, C. (2017). Salmon provides fast and bias-aware quantification of transcript expression. Nat. Methods 2017 144 14: 417–419.
Google Scholar: Author Only Title Only Author and Title

Plett, D.C., Holtham, L.R., Okamoto, M., and Garnett, T.P. (2018). Nitrate uptake and its regulation in relation to improving nitrogen use efficiency in cereals. Semin. Cell Dev. Biol. 74: 97–104.
Google Scholar: Author Only Title Only Author and Title

Porter, S.S. and Sachs, J.L. (2020). Agriculture and the Disruption of Plant–Microbial Symbiosis. Trends Ecol. Evol. 35: 426–439.
Google Scholar: Author Only Title Only Author and Title

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A.R., Bender, D., Maller, J., Sklar, P., De Bakker, P.I.W., Daly, M.J., and Sham, P.C. (2007). PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. Am. J. Hum. Genet. 81: 559.
Google Scholar: Author Only Title Only Author and Title

R Core Team (2022). R: A language and environment for statistical computing. R Foundation Statistical Computing.
Google Scholar: Author Only Title Only Author and Title

Rapaport, F., Khanin, R., Liang, Y., Pirun, M., Krek, A., Zumbo, P., Mason, C.E., Socci, N.D., and Betel, D. (2013). Comprehensive evaluation of differential gene expression analysis methods for RNA-seq data. Genome Biol 14: R95.
Google Scholar: Author Only Title Only Author and Title

Roberge, C., Guderley, H., and Bernatchez, L. (2007). Genomewide Identification of Genes Under Directional Selection: Gene Transcription QST Scan in Diverging Atlantic Salmon Subpopulations. Genetics 177: 1011.
Google Scholar: Author Only Title Only Author and Title

Robinson, M.D., McCarthy, D.J., and Smyth, G.K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26: 139–40.
Google Scholar: Author Only Title Only Author and Title

Rockström, J. et al. (2009). A safe operating space for humanity. Nat. 2009 4617263 461: 472–475.
Google Scholar: Author Only Title Only Author and Title

Sauvage, C., Rau, A., Aichholz, C., Chadoeuf, J., Sarah, G., Ruiz, M., Santoni, S., Causse, M., David, J., and Glémin, S. (2017). Domestication rewired gene expression and nucleotide diversity patterns in tomato. Plant J. 91: 631–645.
Google Scholar: Author Only Title Only Author and Title

Savary, S., Willocquet, L., Pethybridge, S.J., Esker, P., McRoberts, N., and Nelson, A. (2019). The global burden of pathogens and pests on major food crops. Nat. Ecol. Evol. 2019 33 3: 430–439.
Google Scholar: Author Only Title Only Author and Title

Smyth, G.K. (2005). limma: Linear Models for Microarray Data. Bioinforma. Comput. Biol. Solut. Using R Bioconductor: 397–420.
Google Scholar: Author Only Title Only Author and Title

Sultana, N., Islam, S., Juhasz, A., Yang, R., She, M., Alhabbar, Z., Zhang, J., and Ma, W. (2020). Transcriptomic Study for Identification of Major Nitrogen Stress Responsive Genes in Australian Bread Wheat Cultivars. Front. Genet. 11: 583785.
Google Scholar: Author Only Title Only Author and Title

Swanson-Wagner, R., Briskine, R., Schaefer, R., Hufford, M.B., Ross-Ibarra, J., Myers, C.L., Tiffin, P., and Springer, N.M. (2012). Reshaping of the maize transcriptome by domestication. Proc. Natl. Acad. Sci. U. S. A. 109: 11878–11883.
Google Scholar: Author Only Title Only Author and Title

Tajima, F. (1983). Evolutionary relationship of DNA sequences in finite populations. Genetics 105: 437–460.
Google Scholar: Author Only Title Only Author and Title

Taranto, F., D'Agostino, N., Rodriguez, M., Pavan, S., Minervini, A.P., Pecchioni, N., Papa, R., and De Vita, P. (2020). Whole Genome Scan Reveals Molecular Signatures of Divergence and Selection Related to Important Traits in Durum Wheat Germplasm. Front. Genet. 11: 217.
Google Scholar: Author Only Title Only Author and Title

Tian, T., Liu, Y., Yan, H., You, Q., Yi, X., Du, Z., Xu, W., and Su, Z. (2017). agriGO v2.0: a GO analysis toolkit for the agricultural community, 2017 update. Nucleic Acids Res. 45: W122–W129.
Google Scholar: Author Only Title Only Author and Title

Valente, J., Gerin, F., Mini, A., Richard, R., Gouis, J. Le, Prigent-Combaret, C., and Moënne-Loccoz, Y. (2023). Symbiotic Variations among Wheat Genotypes and Detection of Quantitative Trait Loci for Molecular Interaction with Auxin-Producing Azospirillum PGPR. Microorg. 2023, Vol. 11, Page 1615 11: 1615.
Google Scholar: Author Only Title Only Author and Title

Vigouroux, Y., McMullen, M., Hittinger, C.T., Houchins, K., Schulz, L., Kresovich, S., Matsuoka, Y., and Doebley, J. (2002). Identifying genes of agronomic importance in maize by screening microsatellites for evidence of selection during domestication. Proc. Natl. Acad. Sci. U. S. A. 99: 9650–9655.
Google Scholar: Author Only Title Only Author and Title

Watterson, G.A. (1975). On the number of segregating sites in genetical models without recombination. Theor. Popul. Biol. 7: 256–276.
Google Scholar: Author Only Title Only Author and Title

Whitehead, A. and Crawford, D.L. (2006a). Neutral and adaptive variation in gene expression. Proc. Natl. Acad. Sci. U. S. A. 103: 5425–5430.
Google Scholar: Author Only Title Only Author and Title

Whitehead, A. and Crawford, D.L. (2006b). Variation within and among species in gene expression: raw material for evolution. Mol. Ecol. 15: 1197–1211.
Google Scholar: Author Only Title Only Author and Title

Whitlock, M.C. and Guillaume, F. (2009). Testing for Spatially Divergent Selection: Comparing QST to FST. Genetics 183: 1055–1063.
Google Scholar: Author Only Title Only Author and Title

Wood, S.W. and Cowie, A. (2004). A Review of Greenhouse Gas Emission Factors for Fertiliser Production.

Zhang, Z.H., Jhaveri, D.J., Marshall, V.M., Bauer, D.C., Edson, J., Narayanan, R.K., Robinson, G.J., Lundberg, A.E., Bartlett, P.F., Wray, N.R., and Zhao, Q.-Y. (2014). A comparative study of techniques for differential expression analysis on RNA-Seq data. PLoS One 9: e103207.
Google Scholar: Author Only Title Only Author and Title

Zhu, T., Wang, L., Rodriguez, J.C., Deal, K.R., Avni, R., Distelfeld, A., McGuire, P.E., Dvorak, J., and Luo, M.-C. (2019). Improved Genome Sequence of Wild Emmer Wheat Zavitan with the Aid of Optical Maps. G3 Genes, Genomes, Genet. 9: 619–624.
Google Scholar: Author Only Title Only Author and Title