

Pleiotropy structures plant height and seeds weight scaling in barley despite long history of domestication and breeding selection

Tianhua He¹, Tefera Angessa¹, and Chengdao Li¹

¹Murdoch University

August 28, 2020

Abstract

Size scaling describes the relative growth rates of different body parts of an organism following a positive correlation. The genetic mechanism of the size scaling and how artificial selection influencing the pattern of size scaling remain unexplored. Here we utilise diverse barley panel with genome-wide single nucleotide polymorphisms (SNPs) and the measurement of their plant height and seeds weight to explore the genetic mechanisms that lead to a correlation of the two traits and the influence of domestication and breeding selection on the size scaling. Plant height and seeds weight are heritable and remain positively correlated in domesticated barley regardless of growth type and habit. Genome-wide association studies revealed eight SNPs to be associated with both traits. Linkage decay analysis suggests that a considerable proportion of genetic markers associated with either plant height or seeds weight are closely linked in the chromosome. Common factor analysis revealed twenty SNPs conferring pleiotropic effect on both traits. Genes with multiple functions in plant growth and development are involved in structuring plant height and seeds weight scaling. Pleiotropy forms the genetic bases of plant height and seeds weight scaling in barley. Our results suggest an alternative hypothesis for seeds weight evolution in domestication that the selection in plant size may have constrained variation in seeds weight. Our findings contribute to the understanding of the genetic basis of size scaling and open a new venue for seeking the underlying mechanism of a grand theory on allometric scaling in plants.

Introduction

Plant traits are often correlated, either positively or negatively, which has been ecologically explained to reflect the trade-offs in biological functions and resource allocation, and allometric scaling (Freschet et al., 2015; Agrawal, 2020). The traits could become correlated due to common evolutionary processes such as correlated selection (Armbruster et al., 2014). The plant height and seed size scaling are manifested as a positive correlation between the two traits, and it sits within the broad allometric scaling that describes the relative growth rates of different body parts following a positive correlation of an individual organ versus total body size. Allometric scaling emerges as one of a seemingly universal law in biology from genomes to ecosystems (West & Brown, 2005). Over the past century, many morphological, ecological, and evolutionary size-correlated trends have been observed across organisms and life forms (Kleiber, 1947, 1975; McMahon & Bonner, 1983; Schmidt-Nielsen, 1984; Niklas, 1994; Brown & West, 2000; Moles et al., 2004; Novack-Gottshall, 2008; Vasseur et al., 2018). Various of conceptual frameworks have been proposed when seeking the underlying mechanisms of the broadly observed allometric scaling (e.g. Charnov et al., 1993; Blum, 1977; West et al., 1997; Darveau et al., 2002; Weibel, 2002; Niklas, 2004; Moles et al. 2005; Grubb et al. 2005; Rees & Veranble, 2007; Falster et al. 2008, Veranble and Rees, 2009, Westoby et al., 2009). The mechanism underlying the scaling law and traits correlation requires genetic explanations, but that is less explored in the current research endeavour.

Pleiotropy has been proposed as a significant mechanism leading to trait correlation (Saltz et al., 2017). Pleiotropy is defined as the phenomenon in which a genetic variant influences two or more phenotypic traits. The pleiotropic genetic variant may only have a single function, but it is involved in multiple biological

processes; alternatively, the variant could have multiple functions that are related to different traits (Blow & Hoffmann, 2005; Hine et al., 2014; Saltz et al., 2017). Trait correlation could also be the effect of genetic linkage, e.g. the co-selection of gene variants closely linked to selected loci of interest in the chromosome leading the correlation of trait expression, as physical linkage based correlations can be stable over many generations in species with low recombination rate (Agrawal et al., 2010; Langridge & Fleury, 2011). The recent advancements of high throughput genotyping and large scale phenotyping offers an opportunity to decipher the different genetic basis of trait correlation. Genome-wide association studies (GWAS) reveal single nucleotides polymorphisms (SNPs) that are significantly associated with a particular phenotype (Xiao et al., 2017), therefore allows identifying genetic variants that influence more than one trait. Further, if the physical position on the chromosome of the responsible genetic variants is known, their genetic linkage can be examined. Recently, genomic structural equation modelling (genomic SEM) emerges as a powerful method for deciphering the joint genetic architecture of multiple traits. Genomic SEM models shared genetic architecture across multiple phenotypes with factors representing broad genetic liabilities through common factors analysis (Grotzinger et al., 2019). Therefore, genomic SEM offers an opportunity to not only directly explore pleiotropy as a genetic explanation on trait correlation, but also to describe pleiotropic genetic variants that may have driven the trait correlation.

Moreover, plant height and seed size (weight) are important agronomic traits in cereal crops and pulses. The pattern and evolution of plant height and seeds weight scaling might be affected by direct selection, as domestication and modern breeding often target the two traits in opposite directions. For example, shorter and stiffer stems protect cereal crops against lodging and provide a significant yield improvement, referred as ‘Green Revolution’ in literatures (Langridge, 2014). Larger seeds for consumption might have been one of the selection goals in the domestication and breeding of grains and pulses (Milla et al., 2015). However, it is still largely unknown if the artificial selection such as domestication and intense breeding change the pattern of size scaling. Barley (*Hordeum vulgare* L.) is one of the most important cereal crops in Old World agriculture and has been domesticated 10,000 years ago (Badr et al., 2000). Barley has been subject to extensive genomic study with abundant genomic and phenotypic data resources available (e.g. Mascher et al., 2017; Gonzalez et al., 2018; He et al., 2019; Milner et al. 2019), which provides unprecedented opportunity to explore the genomic basis of size scaling in plants. Here we use the recently generated high-density genome-wide SNP profile for a diverse set of barley samples, and their measurements of plant height and seeds weight to identify the possible genetic mechanisms that may lead to a correlation of plant height and seeds weight to explore the genomic basis of size scaling in plants.

Materials and Methods

Phenotypic and genomic data, and plant height - seeds weight scaling

Ready-to-use phenotypic data for plant height and thousands seeds weight (hereafter seeds weight) and high-density genome-wide SNP dataset for approximately 13,000 barley (*Hordeum vulgare* L.) accessions were obtained from the Federal ex situ Genebank for Agricultural and Horticultural Plant Species (IPK) in Germany. The panel includes both domesticated barley (cultivars and landraces) and its conspecific wild progenitor *H. vulgare* ssp. *spontaneum* (K. Koch) Thell. Plant height from the soil surface to the top of the spike, including awns and seeds weight (in the form of thousand seeds weight) were assessed during seed regeneration using plots of at least 3 m² (Gonzalez et al., 2018). SNP profiles were derived from single plant of the accessions in the IPK barley collection through genotyping-by-sequencing (GBS) method (Milner et al. 2019).

We retained samples with both phenotypic and genotypic data available for further analysis. The retained phenotype and genotype data are subject to further filtering with all samples with <10% missing genotypes and minor allele frequency (MAF) > 0.01. Consequently, we have obtained 133,588 SNPs for 12,828 samples, including wild types, landrace, and cultivars, from 85 countries and regions of all continents with agriculture. The samples also contain different habit (winter-type with vernalisation required for flowering, or spring-type with relaxed vernalisation required for flowering) and growth form (two-rowed, or six-rowed), and contain sufficient variation in life history to capture the general scaling law.

Plant height and seeds weight scaling was first evaluated through bivariate linear model analysis using PAST V3 (Hammer et al., 2001). If the correlation between plant height and seeds weight is determined by shared genomic factors, it would be expected that the two traits are evolutionarily correlated independent of their phylogenetic relationship. We therefore first test the evolutionary correlation of the two traits after controlling phylogenetic relatedness among the samples. To do so, we first used *RAxML* to construct the phylogenetic tree of the 12,828 samples following a maximum likelihood procedure (Stamatakis, 2014). We then implemented a generalized least squares regression analysis, and used phylogenetic generalized ANOVA to test the correlation of the two traits after controlling their phylogenetic relationship using the software package of *Phylocom* (Webb et al., 2008).

Heritability and genetic correlation , genome-wide association studies for plant height and seeds weight

We evaluated the heritability of plant height and seeds weight in barley. We employed a genome-based restricted maximum likelihood method (GREML-LDMS) to estimate the narrow-sense SNP-based heritability (h^2_{SNP}) (Yang et al., 2015). To do so, we computed linkage disequilibrium (LD) scores between SNPs with the block size of 100 kb using GCTA (Yang et al., 2011), then used GREML (a function within GCTA) to calculate the proportion of variance in a phenotype explained by the SNPs following an LD score regression ash^2_{SNP} (Yang et al., 2015). We further estimated the genetic correlation between the two traits following the bivariate GREML procedure using GCTA (Yang et al., 2011).

We further identified SNPs that are associated with either plant height or seed weight through GWAS analysis. We first calculated the first five principal eigenvectors from principal components analysis (PCA) using GCTA (Yang et al., 2011) as covariates in the GWAS model in order to account for population genetic structure. GWAS analysis was conducted using program FaST-LMM that calculates and uses kinship as a realised relationship matrix and following a Factored Spectrally Transformed Linear Mixed Model (Listgarten et al., 2012). We used Bonferroni correction to determine significant SNPs.

We finally evaluated linkage disequilibrium (LD) decay using r^2 parameter between all pairwise SNP comparisons within a genome window of 5 Mb by using PLINK ver 1.9 (Chang et al., 2015) and PopLDdecay (Zhang et al., 2019). We examined the pattern of the distance between immediate neighbouring SNP pairs with one SNPs being significantly associated with plant height, and the another with seeds weight, and evaluated against the global LD decay pattern according to their distance separated in the chromosome.

Pleiotropic effect of SNP on plant height and seed weight

GWAS summary statistic data for both traits were obtained from above GWAS analysis and were used for common factor analysis. Pleiotropic effect of each SNP on plant height and seed weight were estimated with a common factor model (Grotzinger et al. 2019). The common factor model included the two traits plant height and seed weight, assuming each SNP assert effect on both traits. The effect of each SNP on each trait was estimated within a genomic structural equation modelling framework (Grotzinger et al., 2019). The estimate of the SNP effect and model test was implemented with *GenomicSEM* (Grotzinger et al., 2019). As the experimental data currently are not compatible with the hypothesis that every mutation (or gene) affects every trait (Wagner & Zhang 2011), and large data size (e.g. >10,000 in this case) could, meanwhile, cause spurious correlation (Lin et al., 2013; Kaplan et al. 2014), we identified the outliers that are deviated from the general pattern of relationship between the SNP effects on plant height and that on seed seeds as pleiotropic SNPs. A bivariate linear regression between the SNP effects on plant height and that on seeds weight was implemented, and residues of the regression were then obtained. The SNPs with residues of the regression beyond the 95% confident zone as outliers using a Z-score method (Z-score >1.96, or Z-score < -1.96).

Basic summary statistics were carried out using PAST v3 (Hammer et al., 2011). The R package *ggplot2* (Wickham, 2016) was used to plot model results including the kolmogorov-smirnov plots. Regression slopes were compared using estimated marginal means with the R package “*emmeans*” (Lenth et al., 2017). Genomic diversity (π , nucleotide diversity) was calculated using TASSEL v5 (Bradbury et al., 2007). Significance was

taken at $p < 0.05$ for the null hypothesis.

Results

Hosted file

image1.emf available at <https://authorea.com/users/353596/articles/477408-pleiotropy-structures-plant-height-and-seeds-weight-scaling-in-barley-despite-long-history-of-domestication-and-breeding-selection>

Figure 1 Correlation of plant height and seeds weight (thousand seeds weight).

A total of 12,828 samples have been genotyped at $> 90\%$ of filtered SNP loci. There are two growth type, spring-type barley (relaxed vernalisation requirement for flowering) and winter-type barley (vernalisation required for flowering). Winter-type barley is taller than spring-type barley by an average of 8.8 cm, while seeds weight of winter-type barley is bigger than that of spring-type barley by a marginal 1.65 g (Figure 1; Table S1). For both winter barley and spring barley, plant height is correlated with seeds weight with $r = 0.268$ and $p = 0.0001$ for spring barley, and $r = 0.335$ and $p = 0.0001$ for winter barley, and the regression slopes do not differ ($p = 0.680$).

Table S1. Summary of the samples in their two traits plant height and thousand seeds weight

	Height (cm)	Height (cm)	Thousand-seed weight (g)	Thousand-seed weight (g)	Height (cm)	Height (cm)
	Spring type	Winter type	Spring type	Winter type	2-row	6-row
No. accessions	9558	2970	7634	2293	2487	5392
Minimum	37.43	50.44	15.64	15.95	49.37	37.43
Maximum	139.97	145.12	68.44	71.95	145.12	141.0
Mean	96.91	105.72	43.14	44.79	99.58	98.35
S.D.	13.28	15.65	7.33	7.27	12.82	15.43
Median	98.6	107.48	44.03	44.8	100.9	99.12

Hosted file

image2.emf available at <https://authorea.com/users/353596/articles/477408-pleiotropy-structures-plant-height-and-seeds-weight-scaling-in-barley-despite-long-history-of-domestication-and-breeding-selection>

Figure 2. Plant height and seeds weight scaling in wild and cultivated barley. A. pattern of scaling in wild barley, landrace and cultivars. B. Kolmogorov-Smirnov plot showing the phenotypic distribution of plant height in in wild barley, landrace and cultivars. C. Kolmogorov-Smirnov plot showing the phenotypic distribution of seeds weight in in wild barley, landrace and cultivars.

Both barley cultivars and landraces show a positive plant height and seeds weight scaling with $r = 0.422$, $p = 0.0001$ for cultivars (N=3390), and $r = 0.230$, $p = 0.0001$ for landraces (N5740), respectively (Figure 2A). However, plant height and seeds weight scaling is not evident in wild barley (N=292) with $r = 0.0368$ and $p = 0.5233$ (Figure 2A), which is probably due to relatively small number of wild barley samples in the analysis. Significant difference between phenotypic distribution in wild barley, landrace and cultivar were observed. Kolmogorov-Smirnov test for equal distribution for plant height between cultivar and landrace returned a $D = 0.128$ and $p = 0.0001$, and a $D = 0.127$, $p = 0.0001$ for seeds weight. Domestication led to a shift to shorter status, from a median of 107.9 cm in wild barley to 98.8 cm in landraces and 102.2 cm in cultivars (Figure 2B). Meanwhile, the shift of plant height to shorter status was accompanied by a shift to smaller seeds, as it would be expected from a positive plant height and seeds weight scaling. Median seeds weight decreased from 46.5 g in wild barley to 45.3 g in landraces and 43.5 g in cultivars (Figure 2C).

With the filtrations, a total of 133,588 SNPs were obtained for the 12,828 samples. Wild barley has a

relatively higher genomic diversity (nucleotide diversity $\pi = 0.0673$, $N = 292$) than both landraces ($\pi = 0.0535$, $N = 5740$) and cultivars ($\pi = 0.0485$, $N = 3390$). Plant height is highly heritable with SNP-based heritability h_{SNP}^2 of 0.603 ± 0.095 , while SNPs explained less variation in seeds weight with h_{SNP}^2 of 0.322 ± 0.100 . Plant height and seeds weight are genetically correlated ($r_G = 0.272 \pm 0.020$, $p < 0.0001$). Barley accessions with closer genetic relatedness tend to have similar plant height and to have similar seeds weight, as indicated by the phylogenetic signal lambda of 0.822 and $p < 0.001$, while by the lambda of 0.853 and $p < 0.001$ for seeds weight. Generalised least-squares fit by REML indicate that the two traits are phylogenetically correlated ($p < 0.001$), implying that the two traits tend to vary correlated toward the similar direction of the phenotypic spectrum.

Genome-wide association studies revealed 314 SNPs, or 0.23% of the 133,588 SNPs in total, associated with plant height, and 190 (0.14%) with seeds weight. Among them, eight SNPs are associated with both traits (Figure 2). The eight SNPs, forming two clusters in chromosome 2H and 5H (Figure 2), can be traced to at least three functional genes (Table 1), with two genes in cluster 1 (an Expansin B3 gene and an Elongation Factor G gene), and one gene in cluster 2 (Vacuolar protein sorting-associated protein 18 gene).

Hosted file

image3.emf available at <https://authorea.com/users/353596/articles/477408-pleiotropy-structures-plant-height-and-seeds-weight-scaling-in-barley-despite-long-history-of-domestication-and-breeding-selection>

Figure 3 Manhattan plots of genome-wide association study (GWAS) on plant height and thousand seeds weight. The two traits share eight SNPs in two clusters that are significantly associated with the traits.

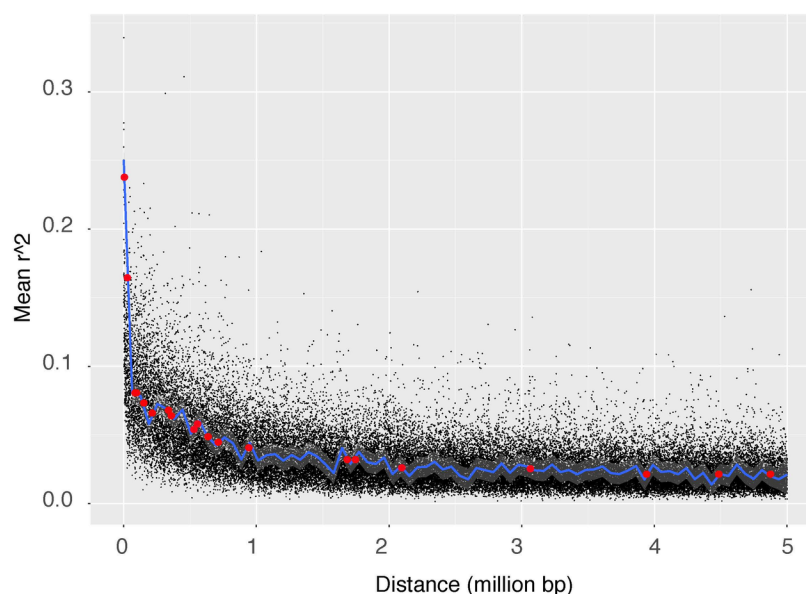


Figure 4. Pattern of linkage decay. Red dots highlight the SNP pairs that are significantly associated with plant height and seed weight.

Linkage disequilibrium decay analysis showed that genes in genomic block within 1 Mb are likely linked in inheritance (if an r^2 of 0.02 and below suggesting random segregation). Of the 23 immediate neighbouring SNPs pairs with one SNP being associated with plant height, and other with seeds weight, 12 pairs were within the distance of 5 Mb in the same chromosome, a further seven pairs within 5 Mb of linkage block

(Figure 4), disproportional higher than expected random distribution on the chromosome (Chi-square $p < 0.05$).

Common factor analysis on the pleiotropic effect of SNPs on both traits revealed as a general trend of positive correlation of SNPs effects on both traits ($r = 0.138$, $p = 0.0001$) (Fig. 2C). Further outlier analysis revealed 20 SNPs with significant pleiotropic effects in both plant height and seed size. Their effect on each trait are significantly correlated with $r = 0.926$ and $p = 0.0001$. (Figure 5). Noticeably, three SNPs, among the 20 found to have pleiotropic effects through common factor analysis, were associated with both trait as revealed in the above GWAS analysis. The SNPs that were identified as being associated with both traits in GWAS analysis and common factor analysis, the variant description and the gene where the SNP is located were shown in Table 1.

Current literature links the responsible SNPs to several genes, including an Hsp70-Hsp90 organising protein gene, an Expansin B3 gene, an Elongation factor G gene, and a Vacuolar protein sorting-associated protein 18 gene (Table 1).

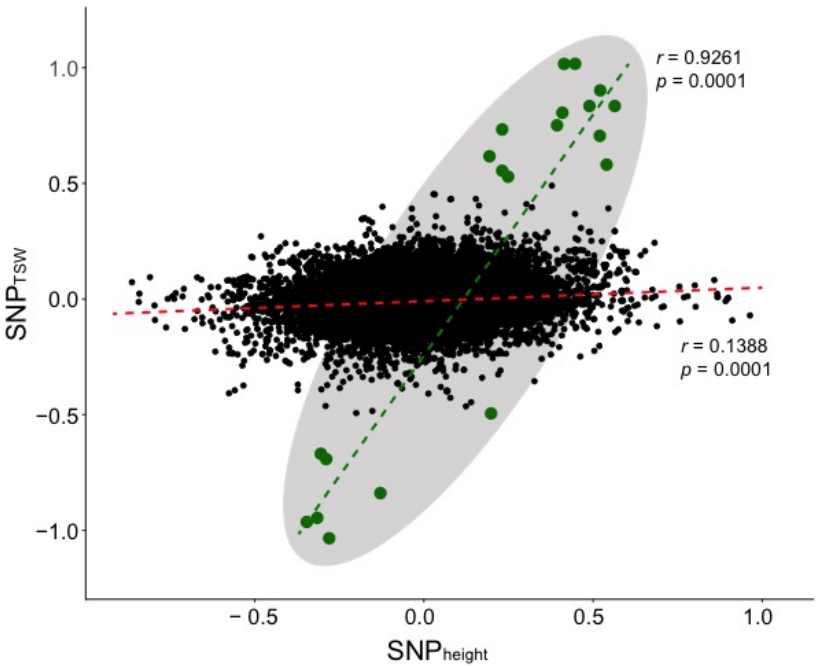


Figure 5. Putative effect of each SNP on plant height and thousand seeds weight. Green dots represent the outliers that confer significant pleiotropic effects on the two traits.

Table 1 SNP loci are contributing to the positive correlation of plant height and seed weight in barley. The light green shaded loci were from joint genetic architecture analysis, the light blue shaded were from GWAS analysis. Red loci highlighted SNPs were detected in both analyses. CHR: chromosome; blank cells indicate relevant information is not available.

CHR	Position	Gene	Variant description	Gene function description
2	632650078			
2	651378029	HORVU2Hr1G092190	missense_variant p.Leu574Phe	Hsp70-Hsp90 organizing protein
2	651535950	HORVU2Hr1G092260	3-prime-UTR_variant	3-prime-UTR_variant

CHR	Position	Gene	Variant description	Gene function description
2	651535959	HORVU2Hr1G092260	3_prime_UTR_variant	3_prime_UTR_variant
2	651676765			
2	651766833	HORVU2Hr1G092280	intergenic_region	intergenic_region
2	652633227			
2	652633251			
2	652633257	HORVU2Hr1G092360	intergenic_region	intergenic_region
2	653986096	HORVU2Hr1G092530	upstream_gene_variant	upstream_gene_variant
2	654161868	HORVU2Hr1G092600	upstream_gene_variant	upstream_gene_variant
2	654165703	HORVU2Hr1G092600	3_prime_UTR_variant	3_prime_UTR_variant
2	654165718			
2	654165739	HORVU2Hr1G092600	3_prime_UTR_variant	3_prime_UTR_variant
2	654786939			
2	654787214			
4	17598761			
2	647258179	HORVU2Hr1G091170	missense_variant p.Ser1Pro	Expansin B3
2	651372029	HORVU2Hr1G092180	synonymous_variant p.Glu123Glu	Elongation factor G
2	651766828	HORVU2Hr1G092280	intergenic_region	intergenic_region
2	652420092	HORVU2Hr1G092340	downstream_gene_variant	downstream_gene_variant
5	593491075			
5	593533461	HORVU5Hr1G093980	HORVU5Hr1G093980	Disease resistance protein
5	593534746	HORVU5Hr1G093980	HORVU5Hr1G093980	Disease resistance protein
5	593561940	HORVU5Hr1G093980	HORVU5Hr1G093980	Vacuolar protein sorting-associated prote

Discussion

The offspring size at independence has been shown to be correlated with adult size across plant species (Moles et al., 2004) and animals (McMahon & Bonner, 1983; Rollinson et al., 2019). Here we show that the size scaling law holds within domesticated crop species when examining the plant height and seeds weights within 12,828 globally collected barley samples. The observed correlation across species between adult size and offspring size at independence (seeds weight in the case of the plant) has been hypothesized as the result of evolutionary coordination between the two traits (Moles et al., 2005; Rees & Venable, 2007, 2009; Westoby et al., 2009), though Grubb et al. (2005) argued for a biomechanically constrained mechanism shaping size scaling. Despite the mechanisms underlying size scaling remain debated from a theoretical perspective (Niklas, 2004), the size components – plant height and seeds weight – are undoubtedly under natural selection or artificially selection in the case of crop plants (Gross & Beckage, 2012). It is a logical hypothesis that the correlation between the two traits has a genetic basis. Using a large dataset with high-density genome-wide SNP map and phenotypes for barley, we demonstrated that two genetic mechanisms might be involved in shaping the plant height and seeds weight scaling. Both plant height and seeds weight are complex traits and are influenced by multiple genes. Multiple genes with significance in plant growth and development assert pleiotropic effects on both plant height and seeds weight contribute to the positive correlation of the two traits. Meanwhile, many of the genes influencing either plant height or seeds weight are closely linked in the chromosome, leading to co-inheritance of the two traits, also contributing to the trait correlation in barley. Together, our results provide direct empirical evidence to the hypothesis that the size scaling in plant has a genetic basis and it may be the result of shared genetic factors controlling both traits.

The plant height and seeds weight scaling hold for the domesticated barley, in both landraces and cultivars, despite thousands of years of domestication and breeding targeting the two traits in opposite directions. Barley breeding tends to select varieties with shorter and stiffer for protection against lodging and benefit of yield improvement (Langridge, 2014). Indeed, we observed an average shorter plant in landrace and cultivar

than in wild barley. On the other hand, larger and plump barley grains are favoured as plump kernels could produce more beer from a given weight of malts (Gupta et al., 2010). Larger seeds would also be a selection goal as they could contribute to yield improvement. However, we observed averagely smaller seeds in both landrace and cultivars than in the wild barley. It seems that plant height and seed weight scaling is genetically constrained and less influenced by direct artificial selection.

The evolution of seeds weight (size) in domestication has been hypothesised be affected by both artificial and natural selection (Milla et al., 2015). However, a comparative study (Kluyver et al., 2013) failed to support the natural selection mechanism that proposes larger seeds were selected in agricultural habitats to allow more effective germination and larger vigorous seedlings with greater fitness (Turnbull & Rees, 1999; Purugganan, 2019). Our results of shared genetic factors controlling the two traits open an alternative hypothesis that the evolution of seed weight in domestication may have been constrained by the selection in plant height (size). Except for major cereal crops, domestication tended to increase aboveground plant size in many crops (Milla et al., 2017), likely as a consequence of selecting larger vegetative part for human consumption. Therefore, larger seeds may have been selected indirectly as the consequence of selecting for larger non-seed part as the two traits are genetically correlated. For cereal crops such as barley, rice and wheat, shorter plant with stiffer stem is advantageous as it resists lodging and increases the relative proportion of photosynthesis product allocating to seeds output. With cereal varieties having shorter plant is selected for improving grain yield, seeds adversely become smaller because of the genetic constrains.

Despite current researches believe that genetic architecture of size scaling and more broadly trait correlations, in general, are polygenic (Saltz et al., 2017). Gardner & Latta (2007) reviewed genetic correlations among quantitative traits and found that an average of only two QTLs (quantitative trait locus) were shared between two correlated traits. The traditional method identifying causal QTLs for a trait, such as GWAS, may have limits because those methods usually rely on linkage decay among causal and non-causal variants to detect associations, and therefore cannot directly establish the number of causal variants (Gianola et al., 2015), which consequently underestimate the pleiotropic genetic variants underlying trait correlations (Saltz et al., 2017). Indeed, using regular GWAS analysis, we identified eight SNPs in two clusters (two QTLs) possibly have pleiotropic effects on plant height and seeds weight in barley. The advanced method in deciphering the genetic architecture of trait correlation, e.g. common factor analysis within the genomic SEM framework, allow us to identify SNPs that may be pleiotropic on influencing plant height and seeds weight, highlighting the power of advancement of analytical methodology.

At least three genes have been revealed to likely play an important role in structuring plant height and seeds weight scaling in barley, an expansins gene, an elongation factor G gene, and an Hsp90 organizing protein gene. It is known that these genes have a function for diverse traits related to plant growth and development. In barley, transcripts of these genes could be found in both grain and shoot (<https://ics.hutton.ac.uk/barleyrtd/index.html>). Expansins enable the local sliding of wall polymers by reducing adhesion between adjacent wall polysaccharides, and has an important role in cell wall re-modelling after cytokinesis. Expansins are required in plant physiological development aspects from germination to fruiting. It is known that expansins influence seed development and seed size, also increase plant height, root mass, number and size of leaves in plants (Chen et al., 2001; Ma et al., 2013; Bae et al., 2014). Elongation factor (*EF*) *G protein* promotes tRNA translocation on the ribosome (Stark, 2000). Liu et al. (2016) reported that overexpressed an EF gene (*MaEF1A*) greatly enhanced plant height, root length, and both rachis and silique length by promoting cell expansion and elongation. Hsp90 organizing protein mediates nuclear-encoded chloroplast preproteins binding to HSP90 prior to chloroplast sorting (Odunuga et al., 2004). Hsp90 is extensively involved in plant growth and development and has a function for diverse traits such as hypocotyl elongation, leaf size, seed mass (Sangster et al., 2008; Delker & Quint 2011). These previous molecular biology studies suggest that the plant height and seed weight scaling may be co-ordinated through multi-functioning genes involved in plant growth and development.

The extensive research into crop genetic improvement has led to the accumulation of extensive data resources from genomics to phenotypes, which offers unprecedented opportunity to explore fundamental biological

questions at the molecular and cellular level. Our results on the genomic mechanism underlying size scaling demonstrated the promise of using crop plants as a model organism in the research of plant biology. On the other hand, allometric relationships between traits can constrain phenotypic variation (Vasseur et al., 2018), which opens promising avenues for crop species with a perspective of targeting optimal crosses based on allometric relationships in parental lines (Garnier & Navas., 2012). For example, a pleiotropic genetic variant with desirable effects on two correlated traits would have direct benefit in plant breeding (Gross & Beckage 2012). In conclusion, plant height and seeds weight scaling could be formed through pleiotropic effect of many genes conferring an effect on both traits, and by the genetic linkage of genes with multiple functions in plant growth and development. Plant height and seed weight scaling is genetically constrained and less influenced by direct artificial selection, which could pose a serious challenge in crop breeding when targeting correlated traits in opposite direction. Though we here only examined plant height and seeds weight scaling within a species, it could be speculated that similar genomic basis may exist to explain the often observed allometric scaling across diverse species. The recent advances at cellular to molecular levels of organization, genomic analysis and large scale phenotyping, and research into heritability and genetic basis of size scaling could open a new venue for a grand unifying theory on allometric scaling in plants.

Data accessibility

The SNP genotype data is publicly available at <https://doi.org/10.5447/IPK/2018/9>. The plant height and seed weight data can be accessed at <https://doi.org/10.5447/IPK/2018/10>.

Author contributions

CL and TH conceived the project. TH and TA collected the data and conducted analysis. TH and CL wrote the paper. CL supervised the project.

Acknowledgement

We thank the Federal ex situ Genebank for Agricultural and Horticultural Plant Species (IPK), Germany, for the access to phenotypic and genomic data. The Project is supported by Grain Research and Development Corporation (UMU00049).

Reference

- Agrawal, A. A., Conner J. K., & Rasmann S. (2010). Trade-offs and adaptive negative correlations in evolutionary ecology. Pages 243–268 in M. Bell, W. Eanes, D. Futuyma, & J. Levinton, editors. *Evolution after Darwin: the first 150 years*. Sinauer Associates, Sunderland, Massachusetts, USA.
- Agrawal, A. A. (2020). A scale-dependent framework for trade-offs, syndromes, and specialization in organismal biology. *Ecology*, 101, e02924.
- Armbruster, W. S., Pélabon, C., Bolstad, G. H., Hansen TF. 2014. Integrated phenotypes: understanding trait covariation in plants and animals. *Philosophical Transactions of Royal Society London Biological Science*, 369, 20130245.
- Bae, J. M., Kwak, M. S., Noh, S. A., Oh, M. J., Kim, Y. S., & Shin, J. S. (2014). Overexpression of sweet potato expansin cDNA (IbEXP1) increases seed yield in Arabidopsis. *Transgenic Research*, 23, 657–667.
- Badr, A., Rabey, H.E., Effgen, S., Ibrahim, H.H., Pozzi, C., Rohde, W., & Salamini, F. (2000). On the origin and domestication history of barley (*Hordeum vulgare*). *Molecular Biology and Evolution* 17, 499–510.
- Blows, M. W., & Hoffmann, A. A. (2005). A reassessment of genetic limits to evolutionary change. *Ecology*, 86, 1371–1384.
- Blum, J. J. (1977). On the geometry of four-dimensions and the relationship between metabolism and body mass. *Journal of Theoretical Biology*, 64, 599–601.
- Bolmgren, K., & Cowan, P. D. (2008). Time-size trade-offs: A phylogenetic comparative study of flowering time, plant height and seed mass in a north-temperate flora. *Oikos*, 117, 424–429.

- Bradbury, P. J., Zhang, Z., Kroon, D. E., Casstevens, T. M., Ramdoss, Y., & Buckler, E. S. (2007). TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics*, *23*, 2633–2635.
- Brown, J. H., & West, G. B. (2000). *Scaling in Biology*. Oxford University Press, New York.
- Chang, C. C., Chow, C. C., Tellier, L. C. A. M., Vattikuti, S., Purcell, S. M., & Lee, J. J. (2015). Second-generation PLINK: rising to the challenge of larger and richer datasets, *Gigascience*, *4*, 7.
- Charnov, E.L. (1993). *Life History Invariants. Some Explorations of Symmetry in Evolutionary Biology*. Oxford University Press, Oxford.
- Chen, F., Dahal, P., & Bradford, K. J. (2001). Two tomato expansin genes show divergent expression and localization in embryos during seed development and germination. *Plant Physiology*, *127*, 928–936.
- Darveau, C. A., Suarez, R. K., Andrews, R. D., & Hochachka, P. W. (2002). Allometric cascade as a unifying principle of body mass effects on metabolism. *Nature*, *417*, 166–170.
- Delker, C., & Quint, M. (2011). Expression level polymorphisms: heritable traits shaping natural variation. *Trends in Plant Science*, *16*, 481–8.
- Falster, D. S., Moles, A. T., & Westoby, M. (2008). A general model for the scaling of offspring size and adult size. *The American Naturalist*, *172*, 299–317.
- Freschet, G. T. Swart E. M., & Cornelissen J. H. (2015). Integrated plant phenotypic responses to contrasting above-and below-ground resources: key roles of specific leaf area and root mass fraction. *New Phytologist*, *206*, 1247–1260.
- Garnier, E. & Navas, M. L. (2012). A trait-based approach to comparative functional plant ecology: concepts, methods and applications for agroecology. A review. *Agronomy for Sustainable Development*, *32*, 365–399.
- Gianola, D., de los Campos, G., Toro, M. A., Naya, H., Schon, C. C., & Sorensen, D. (2015) Do molecular markers inform about pleiotropy? *Genetics*, *201*, 23–29.
- Gonzalez, M. Y., Weise, S., Zhao, Y., Philipp, N., Arend, D., Borner A, ... Schulthess, A. W. (2018). Unbalanced historical phenotypic data from seed regeneration of a barley ex situ collection. *Scientific Data*, *5*, 1.
- Gross, L.J. & Beckage, B. (2012) Toward a metabolic scaling theory of crop systems. *Proceedings of the National Academy of Sciences*, *109*, 15535–6.
- Grotzinger, A. D., Rhemtulla, M., de Vlaming, R., Ritchie, S. J., Mallard, T. T., Hill, W. D., ... Koellinger, P. D. (2019). Genomic structural equation modelling provides insights into the multivariate genetic architecture of complex traits. *Nature Human Behaviour*, *3*, 513.
- Grubb, P. J., Metcalfe, D. J. & Coomes, D. (2005). Comment on ‘a brief history of seed size’. *Science*, *310*, 783A–783A.
- Gupta, M., Abu-Ghannam, N. & Gallagher, E., 2010. Barley for brewing: Characteristic changes during malting, brewing and applications of its by-products. *Comprehensive Reviews in Food Science and Food Safety*, *9*, 318–328.
- Hammer, O., Harper, D. A. T., & Ryan, P.D. (2001). PAST: paleontological statistics software package for education and data analysis. *Paleontologica Electronica*, *4*, 9.
- He, T., Hill, C. B., Angessa, T. T., Zhang, X. Q., Moody, D., Telfer, P., ... Li, C. (2019). Gene-set association and epistatic analyses reveal complex gene interaction networks affecting flowering time in a worldwide barley collection. *Journal of Experimental Botany*, *70*, 5603–5616.
- Hine, E., McGuigan, K. & Blows, M.W. (2014) Evolutionary constraints in high-dimensional trait sets. *American Naturalist*, *184*, 119–131.

- Huxley, J. S. (1924). Constant differential growth-ratios and their significance. *Nature*, *114*, 895.
- Kaplan, R. M., Chambers, D. A., & Glasgow, R. E. (2014). Big data and large sample size: a cautionary note on the potential for bias. *Clinical and translational science*, *7*, 342-346.
- Kermack, K. A. & Haldane, J. B. S. (1950). Organic correlation and allometry. *Biometrika*, *37*, 30–41.
- Kim, T. S., Kim, W. Y., Fujiwara, S., Kim, J., Cha, J. Y., Park, J. H., ... Somers, D. E. (2011). HSP90 functions in the circadian clock through stabilization of the client F-box protein ZEITLUPE. *Proceedings of the National Academy of Sciences*, *108*, 16843-16848.
- Kleiber, M. (1947). Body size and metabolic rate. *Physiological Review*, *27*, 511–541.
- Kleiber, M., (1975) Metabolic turnover rate: a physiological meaning of the metabolic rate per unit body weight. *Journal of Theoretical Biology*, *53*, 199-204.
- Kluyver, T.A., Charles, M., Jones, G., Rees, M., & Osborne, C.P. (2013) Did greater burial depth increase the seed size of domesticated legumes? *Journal of Experimental Botany* , *64*, 4101-4108.
- Langridge, P. (2014). Reinventing the green revolution by harnessing crop mutant resources. *Plant physiology* , *166*, 1682-1683.
- Langridge, P., & Fleury, D. (2011). Making the most of ‘omics’ for crop breeding. *Trends in Biotechnology* , *29* , 33-40.
- Lee, H., Yang, J. Goddard, M. E. Visscher, P. M. Wray, N. R. (2012). Estimation of pleiotropy between complex diseases using single-nucleotide polymorphism-derived genomic relationships and restricted maximum likelihood. *Bioinformatics*, *28*, 2540-2542.
- Lenth, R. (2017). emmeans: Estimated marginal means, aka least-squares means. R package version 0.9.1. <https://CRAN.R-project.org/package=emmeans>
- Lin, M., Lucas, Jr. H.C., & Shmueli, G. (2013). Research commentary—too big to fail: large samples and the p-value problem. *Information Systems Research*, *24*, 906-917.
- Listgarten, J., Lippert, C. Kadie, C. M., Davidson, R.I., Eskin E., & Heckerman, D. (2012) Improved linear mixed models for genome-wide association studies. *Nature Methods*, *9*, 525-526.
- Liu, J. H., Li, Y. C., Zhang, J., Gao, P. Z., Wang, A. B., Zhang, N., ... Jin, Z.Q. (2016). Banana MaEF1A facilitates plant growth and development. *Biologia plantarum*, *60*, 435-42.
- Ma, N., Wang, Y., Qiu, S., Kang, Z., Che, S., Wang, G., Huang, J. (2013). Overexpression of OsEXPA8, a root-specific gene, improves rice growth and root system architecture by facilitating cell extension. *PLoS One* *8*(10) , e75997.
- Mcmahon TA, Bonner JT. 1983. On Size and Life. Scientific American Books, New York.
- Mascher, M., Gundlach, H., Himmelbach, A., Beier, S., Twardziok, S. O., Wicker, T., ... Stern, N. (2017). A chromosome conformation capture ordered sequence of the barley genome. *Nature*, *544*, 427-33.
- Milla, R., & Matesanz, S. (2017) Growing larger with domestication: a matter of physiology, morphology or allocation? *Plant Biology* , *19* , 475-483.
- Milla, R., Osborne, C.P., Turcotte, M.M., & Violle, C. (2015) Plant domestication through an ecological lens. *Trends in Ecology & Evolution* , *30* , 463-469.
- Milner, S. G., Jost, M., Taketa, S., Mazon, E. R., Himmelbach, A., Oppermann, M., ... Schuler, D. (2019). Genebank genomics highlights the diversity of a global barley collection. *Nature Genetics*, *51*, 319-326.
- Moles, A. T., Ackerly, D. D., Webb, C. O., Tweddle, J. C., Dickie, J. B., & Westoby, M. (2005). A brief history of seed size. *Science* , *307* , 576–580.

- Moles, A. T., Ackerly, D. D., Webb, C. O., Tweddle, J. C., Dickie, J. B., Pitman, A. J. & Westoby, M. (2005). Factors that shape seed mass evolution. *Proceedings of the National Academy of Sciences of the United States of America*, *102*, 10540–10544.
- Moles, A. T., Falster, D. S., Leishman, M. R., & Westoby, M. (2004). Small seeded species produce more seeds per square metre of canopy per year, but not per individual per lifetime. *Journal of Ecology*, *92*, 384–396.
- Niklas, K. J. (1994). Plant Allometry. University of Chicago Press, Chicago.
- Niklas, K. J., (2004). Plant allometry: is there a grand unifying theory? *Biological Reviews*, *79* , 871-889.
- Novack-Gottshall, P. M. (2008) Ecosystem-wide body-size trends in Cambrian–Devonian marine invertebrate lineages. *Paleobiology* *34*,210–228.
- Odunuga, O. O., Longshaw, V. M., & Blatch, G. L. (2016). Hop: more than an Hsp70/Hsp90 adaptor protein. *BioEssays* , *26*, 1058–1068.
- R Core Team (2018). R: A language and environment for statistical computing (R Foundation for Statistical Computing).
- Rees, M. & Venable, D. L. (2007) Why do big plants make big seeds? *Journal of Ecology* *95* , 926-936.
- Rollinson, N., Nilsson-Ortman, V., & Rowe, L. (2019). Density-dependent offspring interactions do not explain macroevolutionary scaling of adult size and offspring size. *Evolution*, *73*, 2162-74.
- Saltz, J. B., Hessel, F. C., & Kelly, M.W. (2017) Trait correlations in the genomics era. *Trends in Ecology and Evolution*, *32*, 279–290.
- Sangster, T. A.; Salathia, N.; Undurraga, S.; Milo, R.; Schellenberg, K.; Lindquist, S., & Queitsch, C. (2008). HSP90 affects the expression of genetic variation and developmental stability in quantitative traits *Proceedings of the National Academy of Sciences* , *105*, 2963–2968.
- Sangster, T. A., Salathia, N., Lee, H. N., Watanabe, E., Schellenberg, K., Morneau, K., ... Lindquist, S. (2008). HSP90-buffered genetic variation is common in *Arabidopsis thaliana* . *Proceedings of the National Academy of Sciences* , *105*, 2969-2974.
- Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, *30*,1312-1313.
- Stark, H., Rodnina, M.V., Wieden, H. J., Van Heel, M., & Wintermeyer, W. (2000). Large-scale movement of elongation factor G and extensive conformational change of the ribosome during translocation. *Cell* ,*100* , 301-309.
- Turnbull, L. A., Rees, M., Crawley, M. J. (1999). Seed mass and the competition/colonization trade-off: a sowing experiment. *Journal of Ecology*, *87*, 899-912.
- Vasseur, F., Fouqueau, L., de Vienne, D., Nidelet, T., Violle, C., & Weigel, D. (2019). Nonlinear phenotypic variation uncovers the emergence of heterosis in *Arabidopsis thaliana* . *PLoS Biology* ,*17* (4), e3000214.
- Venable, D. L., & Rees, M., (2009). The scaling of seed size. *Journal of Ecology* , *97*, 27-31.
- Wang, R., Zhang, Y., Kieffer, M., Yu, H., Kepinski, S., Estelle, M. (2016). HSP90 regulates temperature-dependent seedling growth in *Arabidopsis* by stabilizing the auxin co-receptor F-box protein TIR1. *Nature Communications* , *7*, 11677.
- Wagner, G.P., & Zhang, J. (2011). The pleiotropic structure of the genotype–phenotype map: the evolvability of complex organisms. *Nature Reviews Genetics* , *12* , 204-213.
- Webb, C. O., Ackerly, D. D., & Kembel, S. W. (2008). Phylocom: software for the analysis of phylogenetic community structure and trait evolution. *Bioinformatics*, *24*(18), 2098-2100.

- Weibel, E. R. (2002). The pitfalls of power laws. *Nature*, 417, 131-132.
- West, G. B., Brown, J. H., & Enquist, B. J. (1997). A general model for the origin of allometric scaling laws in biology. *Science*, 276, 122-126.
- West, G.B., & Brown, J.H. (2005). The origin of allometric scaling laws in biology from genomes to ecosystems: towards a quantitative unifying theory of biological structure and organization. *Journal of Experimental Biology*, 208, 1575-1592.
- Westoby, M., Leishman, M., & Lord, J. (1996) Comparative ecology of seed size and dispersal. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 351, 1309-1318.
- Westoby, M., Moles, A. T., & Falster, D. S. (2009). Evolutionary coordination between offspring size at independence and adult size. *Journal of Ecology*, 97, 23-26.
- Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis* (Springer-Verlag, New York).
- Xiao, Y., Liu, H., Wu, L., Warburton, M., and Yan, J. (2017). Genome-wide association studies in maize: praise and stargaze. *Molecular Plant*, 10, 359–374.
- Yang, J., Bakshi, A., Zhu, Z., Hemani, G., Vinkhuyzen, A. A. E., Lee, S. H., ... Visscher, P. M. (2015). Genetic variance estimation with imputed variants finds negligible missing heritability for human height and body mass index. *Nature Genetics*, 47, 1114-1120.
- Yang, J., Lee, S. H., Goddard, M. E., Visscher, P. M. (2011). GCTA: a tool for genome-wide complex trait analysis. *American Journal of Human Genetics*, 88, 76-82.
- Zhang, W., Yan, H., Chen, W., Liu, J., Jiang, C., Jiang, H., & Cheng, B. (2014). Genome-wide identification and characterization of maize expansin genes expressed in endosperm. *Molecular Genetics and Genomics*, 289, 1061–1074
- Zhang, C., Dong, S. S., Xu, J. Y., He, W. M., & Yang, T. L. (2019) PopLDdecay: a fast and effective tool for linkage disequilibrium decay analysis based on variant call format files. *Bioinformatics*, 35, 1786–1788.
- Zhou, S., Han, Y., Chen, Y., Kong, X., & Wang, W. (2015). The involvement of expansins in response to water stress during leaf development in wheat. *Journal of Plant Physiology*, 183, 64–74.