Plant Quantitative Genomics in the Spaceflight Environment

A Topical White Paper for Submission to the Biological and Physical Sciences in Space Decadal Survey 2023-2032

Dr. Wolfgang Busch, Salk Institute for Biological Sciences, Harnessing Plants Initiative
wbusch@salk.edu

Dr. Marcio Resende, University of Florida, Department of Horticultural Sciences

Dr. Anna-Lisa Paul, University of Florida, Interdisciplinary Center for Biotechnology Research, Department of Horticultural Sciences

600 character Abstract:

Plants will be necessary for space exploration and colonization to sustain sources for food, feed and fiber. The spaceflight and planetary environments represent novel challenges that lie outside the evolutionary experience of terrestrial organisms. It is imperative that we understand the impact of these environments on plant health and productivity, and that we use these insights to select and engineer plant lines to optimize their adaptability. In addition to tools of genetic engineering, there exists a wealth of genetic variation in natural populations relevant for space exploration.

Plant Quantitative Genomics in the Spaceflight Environment

Wolfgang Busch, Salk Institute for Biological Sciences
Marcio Resende, University of Florida
Anna-Lisa Paul, University of Florida

Plants will be necessary for deep space exploration or space colonization to sustain sources for food, feed and fiber. The spaceflight and planetary environments that comprise this exploration represent novel challenges that lie outside the evolutionary experience of any terrestrial organism. Therefore, it is imperative that first, we understand the impact of these environments on plant health and productivity of plants, and second that we use these insights to select and engineer plant lines to optimize their adaptability. In addition to the molecular genetic tools of engineering, there exists a wealth of genetic variation in natural populations of plants, and many of those traits are relevant for deep space exploration. For instance, even within a species there are significant differences in the degree of tolerance to DNA damaging UV radiation (e.g. Piofczyk et al. 2015) or heavy metal stress (e.g. Li et al. 2019). Both stresses will need to be addressed for plants suited to grow on mars using in situ resources of regolith-based substrates and light.

Plants can be grown on orbit to demonstrate the entire lifecycle.
There is a wealth of plant spaceflight data from a number of species and genotypes, which ranges from observational studies of apparent health and vigor to molecular analyses of transcriptomes and proteomes that categorize genomic responses. Model plant organisms such as Arabidopsis, Mizuna and Brachypodium have contributed the bulk of the fundamental insights into the genes important to the physiological adaptation of plants to spaceflight, even to the level of organ- and cell-specific responses (reviewed in: Barker et al. 2020; Manian et al., 2021). There have also been spaceflight studies in Arabidopsis that have flown mutant genotypes that were chosen based on the patterns of gene expression revealed by earlier experiments (e.g. Califar et al. 2020, Paul et al. 2021). These experiments with mutant genotypes asked the question: can we influence the spaceflight response by manipulating the genome? The answer to that question is yes, and contributes to our understanding of what is important to plants in space with respect to viability, stress responses and productivity for a given plant. It is important to note that novel environments may also include responses from signals that are inappropriately activated or misinterpreted. Thus, it is also important to identify which responses are counterproductive and impose an unnecessary load on the plants as they try to adjust, and then determine whether they can be eliminated to reduce the load on the adaptive process. A refined understanding of the genomic responses of plants to spaceflight can enable genetic manipulation to produce varieties that are better adapted to growth in spaceflight through the elimination of unnecessary responses and through the introduction of new important traits. While there is no single gene that defines the physiological adaptation to spaceflight, there are pathways and metabolic pathways that appear important to almost every genotype that has been assayed. The genes associated with cell-wall remodeling, ROS signaling, and unique application of light sensing pathways suggest that targeting these pathways for investigation may help in the development of "space adapted" plant genotypes. Before genetic manipulation can take place, the genotype must be better correlated to phenotype. Future work will need to focus on which traits to

enhance, which phenotypes to study, manipulate or vary and this can be done by determining the genetic variants of desirable traits and the loci associated with these desirable traits.

Spaceflight experimental platforms

Spaceflight phenotypes can be guided by engineered environmental considerations of the growth hardware. Platforms such as the ISS, suborbital flights, and earth analogs can used to conduct plant growth experiments. Current ISS plant growth hardware includes the Veggie unit and the Advance Plant Habitat (APH), and in the next decade an Advanced Exploration Systems (AES) growth chamber by the name of OhaloIII will be put on orbit specifically for space crop production. Veggie and APH can support up 30-40 of these square 10cm petri plates and OhaloIII will support roughly twice this size. All current and planned units allow for whole plant and seedling experiments, sample preservation and retrieval, and aim use minimal crew time while avoiding microenvironment and other differentials within the confined space.

The use of Quantitative Genomics in terrestrial field agriculture

Plant breeding is a science and a tool that has significantly contributed to the yield increase and general improvement of crops over the last century. Plant breeding is a long and cyclic process that requires the evaluation of populations that have been intercrossed and recombined. Large numbers of plants need to be evaluated to increase the probability of identifying recombination events that create superior individuals. The development of plant breeding as a science leveraged developments over the decades in experimental design, agronomy, quantitative genetics, and more recently, statistical genomics, among other fields. Nowadays, breeding pipelines have been developed for most commercially relevant crops. Plant breeders will then test the plants in a series of target environments for a series of traits of interest. The successful outcome is a variety that is better adapted to the environment where it was selected for, and improved for the traits it was selected for. Importantly, the selection of superior individuals is environment-specific, since the best plants in one location may not present any advantage in a different environment – a phenomena known as genotype by environment interaction.

Quantitative genetics is at the core of plant breeding, since most of the traits selected by breeders are controlled by many genes and quantitatively inherited. Hence, the identification of genes associated with a given quantitative phenotype have had limited impact in plant breeding, since these genes contribute with minor effects to the complex phenotypes. To address this problem, an approach known as genomic selection (GS) was proposed as a tool to accelerate the breeding process and increase genetic gain. Genomic selection is now a well-established tool and its use has led to the development of many commercial cultivars. The fundamental difference between GS and gene mapping approaches is the fact that GS algorithms are not attempting to make inferences about causal genes explaining the trait of interest. Instead, the focus is on achieving the most accurate prediction of the complex phenotype, to guide breeding decisions. Using such prediction models, plant breeders can characterize the DNA of a set of untested individuals, and predict which ones should be moved forward in the breeding process. Moreover, once a GS prediction model is calibrated to predict the plant performance in a given environment of interest, the breeder can select future superior plants prior to field evaluation in the given environment. In principle, it could be possible to establish GS models that can predict the plant performance in spaceflight and planetary environments. However, in order to have these statistical models developed, one would require an initial phenotype screening in the target environment of interest, or a closely related environment, such as earth analogs.

Two additional recent developments can be coupled to GS to improve the breeding of adapted materials adapted to novel environments, and to accelerate genetic gains. The first is the use of computer simulation, which can create many possible iterations of an *in-silico* breeding program to allow process optimization prior to field deployment. Simulation tools are now available to apply quantitative genomics research and allow for future generational predictions. Hence, it is now possible to simulate several decades of a given breeding strategy and estimate the return on the investment or the rate of genetic gain. The second approach is the integration of genomic selection with crop growth models that describe how a crop grows as a function of weather conditions, soil conditions, and other environmental conditions. The integration of these two well established fields have improved our ability to select superior plants in untested environments, as long as we know the environmental parameters of such environments of interest.

The use of GWAS and Arabidopsis

Genetic engineering approaches can be used in addition to traditional or modern plant breeding approaches to change traits in a more accelerated and substantial manner than most breeding approaches can provide. For instance, CRISPR/CaS9 gene editing approaches have allowed for engineering domestication traits such as plant statue, flower and fruit production in wild species (Lemmon et al. 2018, Li et al. 2018) or to enhance quality traits in crops such as omega-3 oil in camelina (Waltz 2018). Utilizing such gene editing approaches simultaneously on several genes or pathways to achieve rapid adaptation of crop species to space conditions appears to be highly promising. However, much like for traditional breeding approaches, it is of fundamental importance to identify the relevant traits for productive growth in the relevant extraterrestrial growth conditions. Once this is achieved, the key genes that determine variation in these traits can be identified and can then serve as prime candidates for genetic engineering approaches. One of the most powerful models for addressing these two challenges is the model plant *Arabidopsis thaliana* (Arabidopsis). The Arabidopsis community has collected more than 7,000 Arabidopsis accessions (natural strains) whose origin ranges from the North of Sweden to Africa and from Japan to the British Aisles. Due to the enormous range of their distribution, Arabidopsis accessions display broad natural genetic and phenotypic variation for almost all traits that have been measured, including plant responses to a diverse set of natural and artificial environments. Importantly, these accessions are isogenic (every seed of an accession contains the same genotype) allowing each measurement of a genotype to be replicated numerous times and in numerous conditions.

High quality SNP and whole genome sequence data for thousands of Arabidopsis accessions allow for efficient genome wide association studies (GWAS) that can pinpoint individual causal genes and their variants that underlie these traits. This has led to the identification of numerous genes and genetic variants underlying quantitative traits in plants (Deja-Muylle et al. 2020). Several of these identified genes are not only relevant for interspecies trait variation, but also determine similar traits in other plant species.

Methods to determine and engineer traits of space crop importance.

To identify traits of interest to space exploration and then engineer these, an integrated, multi-pronged approach can be utilized that iterates smaller-scale screens for traits of interest in spaceflight conditions and larger-scale screens on the ground (Fig. 1). For this, the advantages of the Arabidopsis model can be fully leveraged, as thousands of traits of many of the isogenic and sequenced accessions of Arabidopsis have been catalogued, and many genomic resources for this
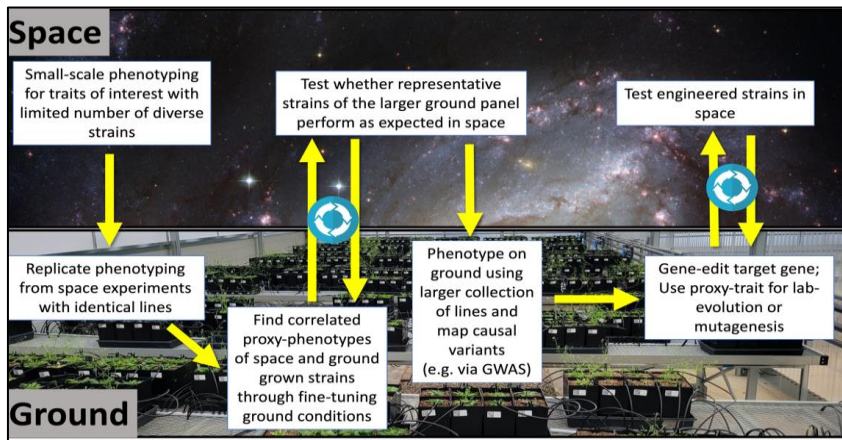
Figure 1. Quantitative genomic spaceflight and ground methodology

model species exist. Arabidopsis accessions display tremendous natural variation that can be mined for potentially adaptive traits for space environments. Acquiring Arabidopsis trait data is easy because many traits can be efficiently monitored during its early growth stages and at multiple time points using time-lapse imaging for thousands of individuals in parallel. Hence, a small-scale screen (i.e. less than 200 Arabidopsis accessions) could be performed under space-flight conditions. Panels of this size have been successfully used to identify nucleotide changes in a few genomic locations that are significantly correlated with trait variation. However, increased detection power will be needed to pinpoint multiple quantitative trait loci (QTLs) that will allow for a more comprehensive collection of genes and their variants that determine space-relevant traits. With a highly powered association panel (e.g. more than 500 accessions) a systems genetics approach can be leveraged in which the genetic association data is jointly analyzed using integrated gene network data (based on interactome, expression and other data) to identify multiple interacting genes that might lead to a better prediction of gene circuits and modules for spaceflight trait engineering. Therefore, an important part of the spaceflight trait engineering strategy will be to identify proxy-spaceflight traits (traits that can be measured on the ground but that are highly correlated with the spaceflight-traits of interest). To achieve this, the same set of accessions that will have been used in space can be subjected to numerous screens on the ground that try to approximate some of the conditions that are relevant during spaceflight (e.g. radiation levels). These conditions can then be refined and fine-tuned to achieve a stronger correlation of the proxy-spaceflight traits to the space-flight related traits. This fine-tuning can be directed by the wealth of trait data sets that already exist for Arabidopsis accessions under different conditions. For instance, if the trait variation observed in spaceflight would correlate with the trait variation of the same set of accessions that had been observed upon exposure to a certain plant hormone, addition of these hormones might be used to increase the correlation of proxy-spaceflight conditions to space conditions. Once a ground-condition has been found that is a good proxy-condition for spaceflight, a larger set of accessions can be tested under proxy-spaceflight conditions on the ground. To test whether the space-proxy ground trait captures the spaceflight traits, one would conduct an additional round of screening in space that includes accessions that hadn't been measured in space but for which the proxy-spaceflight screen predicted certain space responses. If spaceflight experiments confirm the ground predictions, a highly powered set of accessions can be used on the ground to conduct GWAS and map causal genes that underlie spaceflight relevant traits. These causal genes will be genetically engineered or gene edited to enhance the spaceflight relevant traits. Multiplexed gene editing can be used to generate artificial variation in key spaceflight trait determining genes in Arabidopsis, and if these are

5

conserved genes in crop species, to screen the resulting plant lines for enhanced space adaptive traits. Moreover, additional genetic strategies can be pursued under proxy-spaceflight conditions. For instance, utilizing multiparent intercross populations that have the potential to exhibit a larger amount of non-adaptive variation, mutagenesis and selection for spaceflight relevant traits. The engineered plant lines that will have been confirmed to have enhanced traits in proxy-spaceflight conditions will then be tested in space. These steps will most likely involve multiple improvement cycles. Overall, such a strategy promises not only to establish a scalable process for space trait engineering but promises also to provide insight in the molecular and genetic mechanisms that will underpin space-adaptation of plants. This will be highly beneficial to rationally engineer spaceflight adaptive traits in crops the future.

Proposed spaceflight experiment

The ISS contains the capacity for a small-scale phenotyping screen for Arabidopsis root traits, as the APH provides 2 square feet of area and sufficient light and temperature control. One could grow approximately a thousand Arabidopsis seedlings on vertically oriented plates using this area and use a camera set-up to capture root traits. Root growth is sensitive to genotoxic stress such as radiation and is in many instances correlated to plant growth and vigor. Such a number of seedlings would allow for growing more than 100 distinct lines of Arabidopsis with each line being replicated by 10 isogenic individuals, thereby allowing to control for uneven lighting and failure of seedlings to germinate. Such an experimental setup would allow for sampling a notable portion of Arabidopsis diversity, allow for building and calibrating a plant model, and it might even be possible to conduct a GWAS, albeit not overly powered. Because the conditions in the engineered space habitats are highly defined, it can be expected that engineering plant traits and plants for space exploration might be easier than engineering for agricultural environments in which plants are exposed to myriads of constantly fluctuating biotic and abiotic factors.

Representative References

Barker R, Lombardino J, Rasmussen K, and Gilroy S. (2020) Test of Arabidopsis Space Transcriptome: A Discovery Environment to Explore Multiple Plant Biology Spaceflight Exp.. Front Plant Sci **11**:147.

Califar B, Sng NJ, Zupanska A, Paul A-L, and Ferl RJ. (2020) Root Skewing-Associated Genes Impact the Spaceflight Response of Arabidopsis thaliana. Frontiers in Plant Science **11**.

Deja-Muylle A, Parizot B, Motte H, and Beeckman T. (2020) Exploiting natural variation in root system architecture via genome-wide association studies. Journal of experimental botany **71**:2379-2389.

Lemmon ZH, et al. … Lippman ZB. (2018) Rapid improvement of domestication traits in an orphan crop by genome editing. Nature Plants **4**:766-770.

Manian V, Orozco J, Gangapuram H, Janwa H, and Agrinsoni C. (2021) Network Analysis of Gene Transcriptions of Arabidopsis thaliana in Spaceflight Microgravity. Genes (Basel) **12**:337

Li B, et al. … Busch W. (2019) GSNOR provides plant tolerance to iron toxicity via preventing iron-dependent nitrosative and oxidative cytotoxicity. Nature communications **10**:3896.

Li T, et al., … Xu C. (2018) Domestication of wild tomato is accelerated by genome editing. Nat Biotechnol **36**:1160-1163.

Paul A-L, Haveman N, Califar B, and Ferl RJ. (2021) Epigenomic Regulators ELP2 and MET1 Differentially Condition the Spaceflight Response in Arabidopsis. Frontiers in Plant Science **12**.

Piofczyk T, Jeena G, and Pecinka A. (2015) Arabidopsis natural variation reveals connections between UV radiation stress and plant pathogen-like defense responses. Plant Physiol. and Biochem **93**:34-43.

Waltz E. (2018) With a free pass, CRISPR-edited plants reach market in record time. Nat Biotech **36**:6-7.