Gene-to-trait knowledge graphs show association of plant photoreceptors with physiological and developmental processes that can confer agronomic benefits

Hame Bantei Mawkhiew · Lingaraj Sahoo · Eros V. Kharshiing

Received: 16 February 2021 / Accepted: 9 May 2021 / Published online: 17 May 2021
© The Author(s), under exclusive licence to Springer Nature B.V. 2021

Abstract Global population growth, climate change, altered precipitation rates and cropping patterns are increasingly challenging plant scientists to improve crop productivity for food and non-food applications. Hence, there is a pressing need for identifying candidate genes that can be targeted for breeding future crops having enhanced agronomic benefits. Network mapping utilises available data and creates knowledge graphs that aid in visualising association(s) between the individual data items. Here, we have generated gene-to-trait knowledge graphs of known plant photoreceptors using the KnetMiner gene discovery platform which generates biological networks from literature/data sets available in public databases. The resulting knowledge graphs indicate a close association of photoreceptors with various physiological and developmental processes such as shoot architecture, yield, disease resistance and water use efficiency among others, that can confer agronomically important benefits. Such information can be of assistance to plant biologists in the selection of potential gene targets for improving agronomically beneficial traits in plants. This report highlights the potential of machine learning and knowledge graphs as aids in more efficient knowledge discovery and novel decision-making processes which can also be employed for crop breeding or crop engineering.

Keywords Crop physiology · Gene discovery · Knowledge graphs · Molecular breeding · Photoreceptors · Plant development

Introduction

Climate change will inevitably alter plant developmental patterns which could subsequently impact plant growth and productivity in natural environments. Recent reports have suggested that the impact of climate change has already resulted in a likely reduction of approximately 1 % of the average consumable food calories in major crops (Ray et al. 2019). How plants regulate their developmental plasticity in response to changes in climatic conditions such as increased temperatures, altered precipitation rates and increased frequency of extreme climatic events will be critical to agricultural productivity for meeting the needs of the rising global population. To meet this challenge, there is a need to breed or engineer crop varieties that have better adaptability to their growing environment resulting in improved yields. A better understanding of the various...
mechanisms underpinning traits such as yield, disease resistance, heat or cold tolerance, water use efficiency and others, would therefore considerably improve chances of targeting aspects of plant growth that can potentially confer superior agronomic value to crops.

Light signals are utilised by plants for myriad responses, allowing them to re-program their development according to the light they receive. Plants can perceive light of variable spectra via a suite of photoreceptor proteins encoded by distinct families of genes. These include red/far-red (600–750 nm) absorbing phytochromes (phyA to phyF), which were among the first photoreceptors to be identified in plants, blue/UV-A (390–500 nm) absorbing cryptochromes (cry1, cry2, and cry3), phototropins (phot1 and phot2) and ztl/lpk2 family of blue-light absorbing proteins and the UV-B (280–315 nm) absorbing UVR8 (Jenkins 2014; Christie et al. 2015; Kharshiing et al. 2019; Legris et al. 2019). Except for UVR8, the other families of photoreceptors contain more than one member each with members sharing high degrees of similarity with individual members of the same family. The physiological-genetic analyses of the known photoreceptors indicate complex inter-relationships between them which in turn regulate a wide array of plant responses to incident light, ranging from seed germination to flowering (Kharshiing et al. 2019).

Much of our understanding of how different photoreceptors influence plant development has been gained largely due to research on Arabidopsis and emerging reports indicate an increasing interest in the role(s) photoreceptors can play in an agricultural environment (Wargent and Jordan 2013; González et al. 2015). We now are beginning to realise that these photoreceptors not only function in response to the light environment but also regulate plant adaptive responses to a larger array of environmental stimuli than previously thought. Thus while photoreceptor action in response to light signals involving independent and interdependent action of photoreceptor proteins has been elucidated in numerous studies, utilising the available data to visualise association(s), if any, between the photoreceptors and agronomically relevant traits would add tremendous value towards understanding their functional relevance to improving future agricultural productivity.

Research on plant biology is continually generating vast genetic, genomic and phenotype datasets (Boyle et al. 2017). Harnessing of the available data can tremendously aid our understanding of association of photoreceptor genes with traits of agronomic relevance. This information can successively be utilised by plant biologists to associate the relevant phenotypes with genes or gene products and to determine their role(s) in the traits of interest. Connecting phenotypes derived from experimental manipulation, with the correct genetic loci can further help in the rapid identification of genes associated with agronomically relevant traits. Such knowledge can also be employed for identifying the candidate photoreceptor(s) which can be viable molecular targets for enhancing crop productivity.

Knowledge graphs are gaining popularity for enhancing the efficacy of information discovery from large swathes of data (Fensel et al. 2020). When compared to traditional models, knowledge graphs are more amenable to searching and integrating heterogeneous data, especially when the connectedness of the available data are not previously established (Yoon et al. 2017). In this report, network maps for visualising gene-to-trait relationships were generated using KnetMiner (https://knetminer.com) which is an open-source gene discovery platform that allows for generation of biological networks based on previously published works and can be utilised for any species with available genome sequence (Hanley and Karp 2014; Adamski et al. 2020; Hassani-Pak et al. 2020). Here, the different photoreceptors (such as PHYA, PHYB, CRY1, PHOT1, etc.) were used as individual input search terms for retrieving the plant traits associated with these photoreceptors. From the information extracted by the KnetMiner knowledge graphs, only traits and the related photoreceptor gene(s) were selected for display in the networks shown here. We generated knowledge graphs for evaluating the association of agronomically relevant traits with plant photoreceptors in the widely accepted model plants Arabidopsis (representing dicots) and rice (representing monocots). We anticipate that the information gained from the knowledge graphs in this report could be utilised for identifying the candidate photoreceptor genes for targeted manipulation of these genes towards improving agronomically beneficial traits in other crops. The information gained from this report also highlights the potential of machine learning and knowledge graphs as aids in more efficient knowledge discovery and novel decision-making processes which can also be employed for crop breeding or crop engineering.
**PHYA and PHYB genes can be targeted for yield improvements**

The network maps generated in this study reveal reported associations of the principal photoreceptors for red-light such (PHYA and PHYB) with traits such as drought tolerance, water use efficiency, disease control, plant biomass and flowering time among others which can influence yield of crop plants. In the dicot model, Arabidopsis, PHYA and PHYB mapped to traits such as shoot branching, shoot habit, leaf position, leaf angle, meristem and organ identity (Fig. 1). In crop plants, these traits mentioned can play a major role in determining plant traits such as architecture and yield. Unsurprisingly, PHYB, which is an important photoreceptor that regulates plant phenotypic plasticity, especially under sub-optimal light conditions, has been linked to changes in plant architecture (Li et al. 2012) indicating a genetic control of plant architecture (Yang et al. 2016; Krahmer et al. 2018). In field-grown maize plants, phytochrome B has also been reported to promote plant growth, biomass and grain yield (Wies et al. 2019). Plant architecture such as height, branching, and canopy characteristics, are important agronomic traits that affect crop yield. In fruit crops like tomato, axillary branching from lower regions of the plant often leads to decreased fruit size and quality along with increased disease severity due to the resulting dense foliage in the lower regions. Hence, tomato varieties having decreased branching are of great interest to tomato breeders since the removal of the axillary shoots is both necessary and labor-intensive for ensuring optimal yield. Targeting PHYB has also been reported to interrupt leaf formation in Arabidopsis, by accelerating the vegetative to reproductive transition of the shoot apex which enables it to function as a developmental switch which regulates branching by promoting bud outgrowth (Finlayson et al. 2010). In rice, shoot characteristics like mesocotyl length, coleoptile length and biomass show associations with PHYA and PHYB, while traits such as starchiness are associated with PHYA and PHYB and stem weight was linked only with PHYB (Fig. 1). The knowledge graphs also indicate that agronomically important quantitative trait loci (QTLs) determining grain yield, grain shape, grain length and grain number in rice are also reportedly associated with PHYA and PHYB (Fig. 2).

**Cryptochrome (CRY) and phototropin (PHOT) are potential candidates for disease resistance and enhanced growth**

In plants, the cryptochrome family of photoreceptors participate in several plant processes, from seedling responses to light to the entrainment of the circadian clock (Christie et al. 2015). Our results suggests that CRY1 and CRY2 are associated with traits that can confer superior agronomic performance in plants (Figs. 3 and 4). These include traits such as disease resistance, tolerance to salt stress, shoot architecture and yield among others. Noticeably, while both CRY1 and CRY2 are associated with disease resistance in the Arabidopsis model (Fig. 2), this trait is associated with only CRY2 in rice (Fig. 4). This indicates that CRY2 may be a more viable candidate for disease resistance in crops such as rice and other monocots as compared to CRY1. Conversely, CRY1 may be a better target for alleviating stress as compared to CRY2 since traits such as salt stress, oxidative stress and salt tolerance are associated with CRY1 in Arabidopsis (D’Amico-Damião and Carvalho 2018) while boron tolerance maps with CRY1a in rice. It is also interesting to note that in Arabidopsis, association of disease resistance, shoot architecture and stress tolerance with both CRY1 and CRY2 occurs via their interaction with COP1 (CONSTITUTIVE PHOTOMORPHOGIC 1) which is a core photomorphogenic regulator of plant growth and development.

The phototropin (PHOT) blue-light photoreceptors were initially thought to be involved specifically in plant phototropic responses but have been subsequently shown to regulate many other aspects of plants including movement of chloroplast and stomata (Christie 2007; Sharma et al. 2014). Since then various studies have reported several other roles of PHOTs in plant growth and development (Christie et al. 2015). The network maps obtained in this report indicate that PHOT1 and PHOT2 are also associated with agronomic traits such as biomass, stress tolerance, disease resistance and shoot branching (Fig. 3).

**Photoreceptors and water use efficiency**

Increased temperatures and altered precipitation rates are predicted to occur as a result of climate change. Climate models suggest that unabated greenhouse
emissions, will likely result in an increase in global temperatures coupled with higher contrast in precipitation rates between different regions of the globe (Trenberth 2011). Trenberth (2011) also suggest that certain regions will experience more precipitation while others are likely to receive lesser precipitation than at present. As a result crop plants growing in these regions would be subjected to more challenging conditions than they are currently adapted to. Consequently the crops having superior water use efficiency would thus be better equipped to overcome the challenges of such changes in climatic patterns. The knowledge graphs generated indicate that in Arabidopsis, \( \text{PHYB} \) is directly associated with traits such as water use efficiency and drought tolerance while \( \text{CRY1} \) and \( \text{CRY2} \) are associated with stomatal

\[ \text{Fig. 1} \] Network associations of phytochrome red/far-red light photoreceptor (\( \text{PHY} \)) genes in Arabidopsis (top) and rice (bottom) with various physiological and developmental processes (green pentagons) that can influence agronomic traits in plants. Red triangles indicate quantitative trait loci. The interactions among the different \( \text{PHYs} \) in influencing these processes are also shown (black arrows). Network maps for visualising gene-to-trait relationships were generated using KnetMiner (https://knetminer.com) following the instructions available.
resistance (Fig. 4). Our knowledge graphs generated also identify the role of PHYB in enabling plants to evaluate their surrounding temperature and regulate their growth and development accordingly (Legris et al. 2016; Bianchetti et al. 2020).

Data mining and visualisation tools enable plant researchers to look into diverse biological databases for clues to engineer crops which can be better suited to future climatic conditions. By analysing and visualising specific phenotype to genotype relationships from large biological datasets, researchers can expedite the discovery and design of genetic strategies directed at improving beneficial traits in plants (Hallingbäck et al. 2016). Taking advantage of available genomic resources and vast amount of literature for model plants, knowledge graphs can aid discovery of evidence-based gene-to-trait relationships. Such information can be utilised in breeding applications for systematic testing on crop plants in which the effects mediated by the candidate genes have not been previously reported (Hassani-Pak et al. 2020). Machine learning and knowledge graphs can also aid in more efficient knowledge discovery and novel decision-making processes (Gharibi et al. 2020) which can also be employed for crop breeding or crop engineering.

The knowledge graphs generated in this study help in visualising the association of the different reported photoreceptors with physiological and developmental processes that can confer superior agronomic value to crops. This is not surprising as light has also been shown to induce an extensive reprogramming of gene expression in plants (Petrillo et al. 2014; Kaiserli et al. 2018). In several crop species, agronomic traits such as increased drought tolerance, increased grain yield and accelerated pod maturation have been linked to altered expression of photoreceptor genes (Maphlang and Kharsihing 2017; Fantini and Facella 2020). However, to our knowledge there are no reports yet that utilise available information to illustrate the gene-to-trait networks of the different photoreceptors. The extraction and visual representation of information on the association of genes with physiological and developmental processes that can enhance complex traits of agronomic value can accelerate the design of breeding or genome engineering approaches for...
experimental validation. Such machine-learning assisted gene-to-trait associations can potentially result in faster translation of knowledge from laboratory research model plants into commercial crops to address food and energy demands under changing climate. Altering and/or engineering plant genes that can influence plant developmental responses to changing environments are therefore of particular interest to agricultural scientists or breeders for enhancing crop productivity. While genome engineering in crops has perceived concerns about the safety of integration of edited crops into society (Ishii and Araki

Fig. 3 Association of phototropin (PHOT1 and PHOT2) with plant processes that can confer agronomic benefits in crops. The associated physiological and developmental processes are shown in green pentagons and the interactions of PHOT1 and PHOT2 with other genes in regulating these processes are shown with black arrows. Network maps for visualising gene-to-trait relationships were generated using KnetMiner (https://knetminer.com) following the instructions available
recent reports suggest that genome edited crops pose marginal risk (Lassoued et al. 2019). Coupled with the major revolutions in synthetic biology (Shih et al. 2016), data mining such as reported here, which allows integration and visualisation of connected data can enable association of candidate genes with agronomically beneficial plant processes. The availability of such information will provide plant biologists with enormous possibilities for crop improvement to enhance the agronomic value of domesticated lines for addressing future food and non-food requirements. Given the burgeoning demand for enhanced yield, and nutrition of crops for food security under a changing climate, these novel approaches to identifying candidate genes for targeted breeding would be of enormous benefit to future agriculture.

Acknowledgements This work is partly supported by Eureka Project with grant on NECBH/2019-20/139 under North East Centre for Biological Sciences and Healthcare Engineering (NECBH) Twinning Outreach Programme hosted by Indian Institute of Technology Guwahati (IIITG), Assam funded by Department of Biotechnology (DBT), Ministry of Science and Technology, Govt. of India with number BT/COE/34/SP28408/2018 to EK and grant no. EEQ/2019/0000223 (Science and Engineering Research Board, (Department of Science and Technology, New Delhi) to EK. The funders had no role in the study design; in the collection, analyses and interpretation of data; in the writing of the article; and in the decision to submit the article for publication.

Author Contributions EK conceived and directed the study, HBM and EK acquired data and performed the study, EK analysed the data and wrote the manuscript, EK and LS revised the article. All authors read and approved the manuscript for publication.

Declarations

Conflict of interest The authors also declare that they have no conflict of interest.

References

Adamski NM, Borrill P, Brinton J, et al (2020) A roadmap for gene functional characterisation in crops with large genomes: lessons from polyploid wheat. Elife 9:e55646
Alba R, Cordonnier-Pratt M-M, Pratt LH (2000) Fruit-localized phytochromes regulate lycopene accumulation independently of ethylene production in tomato. Plant Physiol 123:363–370
Ballaré CL (2014) Light regulation of plant defense. Ann Rev Plant Biol 65:335–363
Bianchetti R, DeLuca B, deHarlo A et al (2020) Phytochrome-dependent temperature perception modulates isoprenoid metabolism. Plant Physiol 183:869–882
Botto JF, Alonso-Blanco C, Garzarón I et al (2003) The Cape Verde Islands allele of cryptochrome 2 enhances cotyledon
unfolding in the absence of blue light in Arabidopsis. Plant Physiol 133:1547–1556

Boyle EA, Li YI, Pritchard JK (2017) An expanded view of complex traits: from polygenic to omnigenic. Cell 169:1177–1186

Chen M, Chory J (2011) Phytochrome signaling mechanisms and the control of plant development. Trends Cell Biol 21:664–671

Christie JM (2007) Phototropin blue-light receptors. Annu Rev Plant Biol 58:21–45

Christie JM, Blackwood L, Petersen J, Sullivan S (2015) Plant flavoprotein photoreceptors. Plant Cell Physiol 56:401–413

D’Amico-Damiao V, Carvalho RF (2018) Cryptochrome-related abiotic stress responses in plants. Front Plant Sci 9:1897

Fantini E, Facella P (2020) Cryptochromes in the field: how blue light influences crop development. Physiol Plant 169:336–346

Fensel D, Şimshek U, Angele K et al (2020) Introduction: what is light? Plant Physiol 176:1039–1048

Hanley SJ, Karp A (2014) Genetic strategies for dissecting complex traits in biomass willows (Salix spp.). Tree Physiol 34:1167–1180

Hart JE, Sullivan S, Hermanowicz P et al (2019) Engineering the phototropinophotocycle improves photoreceptor performance and plant biomass production. Proc Nat Acad Sci 116:12550–12557

Legris M, Ince YÇ, Fankhuaser C (2019) Molecular mechanisms underlying phytochrome-controlled morphogenesis in plants. Nat Commun 10:1–15

Legris M, Klose C, Burgie ES et al (2016) Phytochrome B integrates light and temperature signals in Arabidopsis. Science 354:897–900

Li L, Ljung K, Breton G et al (2012) Linking photoreceptor excitation to changes in plant architecture. Genes Dev 26:785–790

Ray DK, West PC, Clark M et al (2019) Climate change has likely already affected global food production. PloS one 14(5):e0217148

Rizzini L, Favory J-J, Cloix C et al (2011) Perception of UV-B by the Arabidopsis UVR8 protein. Science 332:103–106

Santer BD, Taylor KE, Wigley TML et al (1996) A search for human influences on the thermal structure of the atmosphere. Nature 382:39–46

Shalitin D, Yang H, Mockler TC et al (2002) Regulation of Arabidopsis cryptochrome 2 by blue-light-dependent phosphorylation. Nature 417:763–767

Sharma S, Kharshiing E, Srinivas A et al (2014) A dominant mutation in the light-oxygen and voltage 2 domain impairs phototropin1 signaling in tomato. Plant Physiol 164:2030–2044

Shih PM, Vuu K, Mansoori N et al (2016) A robust gene-stacking method utilizing yeast assembly for plant synthetic biology. Nat Commun 7:1–8

Trenberth KE (2011) Changes in precipitation with climate change. Clim Res 47:123–138

Wargent JJ, Jordan BR (2013) From ozone depletion to agriculture: understanding the role of UV radiation in sustainable crop production. New Phytol 197:1058–1076

Wies G, Mantese AI, Casal JJ, Maddonni GÁ (2019) Phytochrome B enhances plant growth, biomass and grain yield in field-grown maize. Ann Bot 123:1079–1088
Yang D, Seaton DD, Krahmer J, Halliday KJ (2016) Photoreceptor effects on plant biomass, resource allocation, and metabolic state. Proc Nat Acad Sci 113:7667–7672
Yoon B-H, Kim S-K, Kim S-Y (2017) Use of graph database for the integration of heterogeneous biological data. Genom Inform 15:19–27

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.