Genome **Biology**

EDITORIAL

Plant genomics: sowing the seeds of success

Gemma D Bilsborough*

The exquisite beauty of plants and their fundamental importance to life on Earth has captivated scientists throughout history. During antiquity, Theophrastus began the systematic characterization of plants according to Aristotle's principles of taxonomy. Centuries later, Mendel's studies of the inheritance of traits in pea plants founded the field of genetics, and much of Darwin's work on the evolution of forms by natural selection was supported by experiments on plants. By 2000, the seeds of success were sown in the field of plant genomics with the sequencing of the genome of *Arabidopsis thaliana* [1], a ubiquitous weed that was adopted as the plant model species in the 1990s due to its many desirable traits, such as a short generation time, small diploid genome, and self-compatibility. Prior to this adoption of *Arabidopsis* as a model species, studying a panoply of agronomically and economically important species, such as maize, rice, tomato, petunia and snapdragon, had been favored.

The burgeoning worldwide human population means the pressure to produce sustainable quantities of food and fuel crops, across a relatively constant land area, has never been higher. To improve yields of crops, it will be imperative to understand the genetic mechanisms controlling traits of interest, and genomics approaches will be vital to achieve this. Thus, with the topic of plant genomics at an unprecedented level of importance, we feel it is pertinent to highlight the best of plant research in a special issue.

Crop genomics: growing potential

Research on crops has been experiencing a glorious renaissance over the past few years. Crop species often have large, polyploid genomes with high levels of repetitive sequences, which pose challenges for generating genome assemblies. As Bevan and Uauy [2] describe in their Review article, the recent advances made in next-generation sequencing technologies, combined with the development of bioinformatics approaches, have allowed the assembly of many crop genomes. These improvements have enabled new opportunities for improving existing varieties of crops, in addition to developing new varieties.

Bread wheat is a staple food crop across the world, and has a complex hexaploid genome. To apply effective genomics approaches to improve and breed new bread wheat varieties it will be essential to generate a complete genome sequence. Although whole genome shotgun sequencing allowed the recent assembly of the bread wheat gene set [3], a complete bread wheat genome will be necessary to determine the full effect of genomic variation on wheat phenotypes. Philippe *et al.* [4] describe the application of BAC-by-BAC sequencing to generate a physical map of wheat chromosome 1BL, which has the highest marker density to-date with 11 markers per Mb, and also has high levels of anchoring and contig ordering. It is hoped that applying the methods employed in Philippe *et al.* to the other wheat chromosomes may allow a complete bread wheat genome sequence to be produced in the near future.

Changes in genome structure, in particular copy number variations, may play a role in regulating phenotypic variation in plants. However, the mechanisms regulating the production of copy number variation and the distribution of copy number variants throughout important crop genomes, have remained elusive. Muñoz-Amatriaín *et al.* [5] compare copy number variation distribution between the reference cultivar Morex and 14 barley genotypes (including eight cultivars and six wild barleys) using a barley comparative genomic hybridization array. Interestingly, they identify higher levels of diversity in copy number variation in wild barleys compared with cultivars, which may facilitate future barley breeding programs.

Despite the sequencing of the maize genome in 2009 [6], comparatively little is known about the high levels of genetic diversity that exist between the many maize inbred lines. To address this, Romay *et al.* [7] perform genotyping-by-sequencing of 681,257 SNP markers in the 2,815 maize accessions in the USA national maize inbred seed bank. This resource will be invaluable for the analysis of genetic diversity within the publicly available maize panel using genome-wide association studies and genomic selection approaches.

A key question in maize genomics has been to determine the mechanisms that control the conservation and

*Correspondence: gemma.bilsborough@genomebiology.com

divergence of transcriptomic and epigenomic changes in hybrids. To investigate this, He *et al.* [8] generate transcriptomic and epigenomic maps in the shoots and roots of two maize inbred lines, as well as in their reciprocal hybrids. They show that although there is a positive correlation between gene expression and histone modifications in hybrids and parents, intriguingly there is no significant parental bias in gene expression in maize hybrids, in either shoots or roots.

Model behavior: lessons from *Arabidopsis* and *Brachypodium*

Since *Arabidopsis* was the first plant to have its genome sequenced, it necessarily has the longest history in plant genomics. Despite this, there are many aspects of *Arabidopsis* growth and development that remain unknown. Forward genetic screens are a powerful way of determining the functions of genes in plants, although applying traditional methods to genetically map mutations can be laborious. Identifying the causal mutation of a phenotype can be accelerated by combining genetic mapping with whole genome sequencing, in a process termed mapping-by-sequencing, and James *et al.* [9] describe important new guidelines for mapping-by-sequencing in *Arabidopsis.* They generate a Perl-based simulation tool called Pop-seq, which simulates mapping populations to determine the number of plants and the sequencing depth required to identify candidate mutations in outcrossed and backcrossed mapping populations.

Understanding how cell type specification and differentiation are regulated is an important area of plant biology, and *Arabidopsis* root hair cells represent an excellent system for studying this. Lan *et al.* [10] generate a reference map of transcript abundance, cell-type dependent splicing, and differential accumulation of proteins in *Arabidopsis* root hair cells. By comparing transcriptomic and proteomic profiles, they show that while gene expression is associated with higher abundances of mRNA and protein, decreased gene expression does not necessarily correlate with reduced protein abundance.

Brachypodium distachyon is a model grass species, which is closely related to wheat and barley, and can be used to study grain development: a trait that is sensitive to heat stress. H2A.Z nucleosomes have been shown to regulate transcription in a temperature-dependent manner in *Arabidopsis* as their occupancy in transcriptional start sites of genes reduces with concomitant increases in temperature [11]. However, little is known about the mechanisms controlling temperature-dependent grain development in grasses. Boden *et al.* [12] show that increasing temperature has stronger effects on *B. distachyon* morphology, transcription and H2A.Z nucleosome distribution in reproductive grain development compared with vegetative seedling development. They also find that

transgenic plants with reduced H2A.Z occupancy phenocopy heat-stressed wild-type plants, suggesting that H2A.Z nucleosomes mediate increased thermal sensitivity to heat stress during *B. distachyon* reproductive development.

Sensing change: responses to the environment

Due to their sessile nature, plants must respond to environmental signals, and the sensitivity of plants to their environment has fascinated scientists and poets alike. Shelley eloquently notes in his poem 'The Sensitive Plant' that *Mimosa pudica* leaves can open and close in response to the diurnal cycle as well as in response to touch. However, plants also elicit less conspicuous responses at the level of gene expression and metabolism.

Grapevine (*Vitis vinifera*) is an extensively cultivated fruit crop that exhibits phenotypic plasticity of its berries: this can mean that berries on a single plant mature at different rates, leading to variations in wine quality that are strongly dependent on the environment and vintage. To understand the transcriptomic and metabolomic responses of grapevine berries to environmental differences, Dal Santo *et al.* [13] grew a single clone of *Vitis vinifera* cultivar Corvina in 11 different vineyards across Italy and harvested berries at three developmental stages (veraison, mid-ripening and fully-ripe) across three consecutive years. By performing microarrays and mass spectrometry analyses, they show that the different vineyard environments do not significantly affect gene expression. However, different season climates affect gene expression and metabolism, particularly with regard to genes regulating the synthesis of phenylpropanoid-derived compounds, and early stages of berry development are most responsive to seasonal climate changes. These results will help determine why different terroirs affect the quality of wine, and will aid breeding programs to improve berry and wine quality.

In addition to abiotic stresses, plants must also respond to biotic interactions. As reviewed in Clare *et al.* [14], genome-wide approaches are beginning to be employed to determine the mechanisms controlling plant-pollinator interactions. Turner *et al.* [15] discuss how genomics techniques have also unraveled the striking complexity of plant microbiomes, and Rey and Schornack [16] highlight that determining how beneficial symbiotic microbes and detrimental parasitic microbes interact with plant hosts will be crucial for effective crop breeding programs. In addition, Pais *et al.* [17] discuss progress in understanding how plant parasitic oomycetes, a destructive group of pathogens that cause significant damage to crop yields, infect plant hosts. To investigate the molecular interactions that occur between *Phytophthora capsici* and tomato, Jupe *et al.* [18] compare transcriptional changes that occur simultaneously in the

pathogen and plant over the course of infection. They identify two significant transcriptional changes associated with early infection and the transition from biotrophy to necrotrophy, which may play significant roles in controlling infection in plants and also the life cycle of the pathogen.

The tree of life

Trees are an extremely diverse group of plants that have evolved important roles in many ecosystems, and forest trees are also of great economic importance. Neale *et al.* [19] provide an update on how research on tree genomics has been progressing, and argue that open access data sharing will be essential for allowing the development of forest tree genomics, as well as outlining key research priorities that the community need to focus on.

The cacao tree, *Theobroma cacao*, produces vibrantly colored pods containing beans that are used to create chocolate. There are several varieties of cacao, and the Matina cultivar is a traditional variety with low genetic diversity. Some cacao varieties have undesirable traits such as an unpleasant aroma and flavor caused by high acidity and astringency, which are associated with red pod coloration. Cacao derived from varieties with poor flavor qualities is often blended with cacao derived from other varieties that have more desirable flavors associated with green pod coloration. Unfortunately, the blending of cacao reduces the overall quality of chocolate. Motamayor *et al.* [20] generate a high quality genome sequence of *Theobroma cacao* cultivar Matina to improve the speed and accuracy of cacao breeding. Utilizing the genome sequence in conjunction with a haplotype-based method, they identify the *TcMYB113* gene, which encodes an R2R3 MYB transcription factor, as a likely candidate for controlling red/green pod coloration: higher expression of *TcMYB113* is associated with red pod color. It is hoped that these results will rapidly accelerate breeding programs to improve the quality of chocolate produced from cacao beans.

Concluding remarks

Co-ordinating a special issue is an exciting enterprise, and we are indebted to all our authors for submitting their interesting research, and to our reviewers for taking the time to scrutinize the studies. We would also like to thank our Guest Editors, Mario Caccamo and Erich Grotewold, for their invaluable ideas, advice and support, and for contributing their Editorial discussing why plant genomics is such a timely topic [21]. We are passionate about plants at *Genome Biology*, and we hope that you will share our enthusiasm for this special issue on the blossoming field of plant genomics.

Published: 28 June 2013

References

1. The Arabidopsis Genome Initiative: **Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*.** *Nature* 2000, **408**:796-815.
2. Bevan MW, Uauy C: **Genomics reveals new landscapes for crop improvement.** *Genome Biol* 2013, **14**:206.
3. Brenchley R, Spannagl M, Pfeifer M, Barker GL, D'Amore R, Allen AM, McKenzie N, Kramer M, Kerhornou A, Bolser D, Kay S, Waite D, Trick M, Bancroft I, Gu Y, Huo N, Luo MC, Sehgal S, Gill B, Kianian S, Anderson O, Kersey P, Dvorak J, McCombie WR, Hall A, Mayer KF, Edwards KJ, Bevan MW, Hall N: **Analysis of the bread wheat genome using whole-genome shotgun sequencing.** *Nature* 2012, **491**:705.
4. Philippe R, Paux E, Bertin I, Sourdille P, Choulet F, Laugier C, Šimkova H, Šafař J, Bellec A, Vautrin S, Frenkel Z, Cattonaro F, Magni F, Scalabrin S, Martis M, Mayer K, Korol A, Berges H, Doležel J, Feuillet C: **A high density physical map of chromosome 1BL supports evolutionary studies, map-based cloning and sequencing in wheat.** *Genome Biol* 2013, **14**:R64.
5. Muñoz-Amatriaín M, Eichten SR, Wicker T, Richmond TA, Mascher M, Steuernagel B, Scholz U, Ariyadasa R, Spannagl M, Nussbaumer, T, Mayer KFX, Taudien S, Platzer M, Jeddeloh JA, Springer NM, Muehlbauer GJ, Stein N: **Distribution, functional impact and origin mechanisms of copy number variation in the barley genome.** *Genome Biol* 2013, **14**: R58.
6. Schnable PS, Ware D, Fulton RS, Stein JC, Wei F, Pasternak S, Liang C, Zhang J, Fulton L, Graves TA, Minx P, Reily AD, Courtney L, Kruchowski SS, Tomlinson C, Strong C, Delehaunty K, Fronick C, Courtney B, Rock SM, Belter E, Du F, Kim K, Abbott RM, Cotton M, Levy A, Marchetto P, Ochoa K, Jackson SM, Gillam B, *et al*: **The B73 maize genome: complexity, diversity, and dynamics.** *Science* 2009, **326**:1112–1115.
7. Romay MC, Millard M, Glaubitz JC, Peiffer JA, Swarts KL, Casstevens TM, Elshire RJ, Acharya CB, Mitchell SE, Flint-Garcia S, McMullen MD, Holland JB, Buckler ES, Gardner, C: **Comprehensive genotyping of the USA national maize inbred seed bank.** *Genome Biol* 2013, **14**:R55.
8. He G, Chen B, Wang X, Li X, Li J, He H, Yang M, Lu L, Qi Y, Wang X, Deng XW: **Conservation and divergence of transcriptomic and epigenomic variations in maize hybrids.** *Genome Biol* 2013, **14**:R57.
9. James GV, Patel V, Nordström KJV, Klasen JR, Salomé PA, Weigel D, Schneeberger K: **User guide for mapping-by-sequencing in *Arabidopsis*.** *Genome Biol* 2013, **14**:R61.
10. Lan P, Li W, Lin W-D, Santi S, Schmidt W: **Mapping gene activity of *Arabidopsis* root hairs.** *Genome Biol* 2013, **14**:R67.
11. Kumar SV, Wigge PA: **H2A.Z-containing nucleosomes mediate the thermosensory response in Arabidopsis.** *Cell* 2010, **140**:136–147.
12. Boden SA, Kavanova M, Finnegan EJ, Wigge PA: **Thermal stress effects on grain yield in *Brachypodium distachyon* occur via H2A.Z-nucleosomes.** *Genome Biol.* 2013, **14**:R65.
13. Dal Santo S, Tornielli GB, Zenoni S, Fasoli M, Farina L, Anesi A, Guzzo F, Delledonne M, Pezzotti M: **The plasticity of the grapevine berry transcriptome.** *Genome Biol* 2013, **14**:R54.
14. Clare EL, Schiestl FP, Leitch AR, Chittka L: **The promise of genomics in the study of plant-pollinator interactions.** *Genome Biol* 2013, **14**:207.
15. Turner TR, James EK, Poole PS: **The plant microbiome.** *Genome Biol* 2013, **14**:209.
16. Rey T, Schornack S: **Interactions of beneficial and detrimental root-colonizing filamentous microbes with plant hosts.** *Genome Biol* 2013, **14**:121.
17. Pais M, Win J, Yoshida K, Etherington GJ, Cano LM, Raffaele S, Banfield MJ, Jones A, Kamoun S, Saunders DGO: **From pathogen genomes to host plant processes: the power of plant parasitic oomycetes.** *Genome Biol* 2013, **14**:211.
18. Jupe J, Stam R, Howden AJM, Morris JA, Zhang R, Hedley PE, Huitema E: ***Phytophthora capsici*-tomato interaction features dramatic shifts in gene expression associated with a hemi-biotrophic lifestyle.** *Genome Biol* 2013, **14**:R63.
19. Neale DB, Langley CH, Salzberg SL, Wegrzyn JL: **Open access to tree genomes: the path to a better forest.** *Genome Biol* 2013, **14**:120.
20. Motamayor JC, Mockaitis K, Schmutz J, Haiminen N, Livingstone D, Cornejo OE, Findley SD, Zheng P, Utro F, Royaert S, Saski C, Jenkins J, Podicheti R, Zhao M, Scheffler BE, Stack JC, Feltus FA, Mustiga GM, Amores F, Phillips W, Marelli JP, May GD, Shapiro H, Ma J, Bustamante CD, Schnell RJ, Main D, Gilbert D, Parida L, Kuhn DK: **The genome sequence of the most widely cultivated cacao type and its use to identify candidate genes regulating pod color.** *Genome Biol* 2013, **14**:R53.
21. Caccamo M, Grotewold E: **Turning over a new leaf in plant genomics.** *Genome Biol* 2013, **14**:403.