

EDITORIAL

Turning over a new leaf in plant genomics

Mario Caccamo^{*1}, Erich Grotewold²

Plant species have adapted to most of the planet's ecosystems, offering an amazing range of traits, behaviors and morphologies. Unlocking the underlying mechanisms for this diversity is providing novel insights into the molecular hardwiring of plants and how they interact with their environment, furnishing new opportunities to develop better and more sustainable crops. Access to fast and inexpensive sequencing and advances in computational methods, coupled with an increasing demand for food and fuel crops, have provided the basis for the current revolution in plant genomics.

The development of agriculture and the ability to store food laid the foundations of early human civilizations and defined modern society very much as we understand it today. Arguably, the domestication of small grain crops over the past 10,000 years ranks as one of the most important events in human history. The co-evolution of humans and plants, however, goes back to the origins of the primate lineage. One example is the expansion of hominids, which some authors correlate with the direct adoption in the diet of C_4 plants, or indirectly by preying on animals that consumed C_4 plants [1]. The development of salivary amylase, key in the digestion of starch, which emerged as a gene family expansion independently in primates and rodents, provides another example. The number of copies of the amylase genes is highly variable in human populations with different starch diets [2]. But it is not only about food: fiber crops provided clothes and paper; wood from trees became material for shelters, and until the discovery of fossil fuels, was one of the most important sources of energy. We have also used plant extracts as a source of medicines and other useful biochemicals that in many cases cannot yet be synthesized in the laboratory. Plants also have a direct impact on our environment: as autotrophs they are responsible for fixing atmospheric carbon into food and fiber, with RuBisCO, one of the plant enzymes involved in this process, being probably the most abundant protein on earth.

Rich biology, amazing genomics

Plant species have developed diverse strategies to adapt and thrive in all kinds of climates and terrains, resulting in an exquisite repertoire of natural products. Not having the opportunity to travel to obtain food or run away from adverse conditions, plants have evolved to deal with extreme changes in the environment. Annual plants, for instance, can focus the most resource-intensive activities, such as reproduction, at the most favorable time of the year. Other changes, such as day-to-night fluctuations, are linked to the plant circadian cycle and the synchronized expression of specific sets of genes at particular times. Plants also need to cope with more sudden changes such as attacks from pathogens and grazing animals. These strategies are supported by rich and complex metabolic networks that enable plants to synthesize a wide range of compounds (for example, alkaloids, terpenoids and phenolics) [3]. The constant need for adaptation, the prevalence of polyploid species and past whole-genome duplication events are reflected in the larger number of genes observed in plants [4] compared with animals. This diversity has stimulated many of the scientific breakthroughs that have emerged from research in plant sciences, including the discovery of transposable elements in maize by Barbara McClintock back in the 1940s, the characterization of RNA interference [5] and, more recently, the use of TALE nucleases to target and edit plant genomes [6]. The completion of the *Arabidopsis thaliana* genome [7] at the turn of the century followed by the rice genome publications [8,9] have also stirred a growing body of research in plant genomics. The success of *Arabidopsis* as a model organism is rooted in its accessible genetics tied to a relatively simple genome [10-12]. Many plants, especially some of the popular monocot crops and gymnosperm trees, have, however, more complex and larger genomes. The Triticeae tribe within the grasses family includes some of the most important cereal crops, such as barley (*Hordeum vulgare*), with a genome that is 50 times larger than *Arabidopsis* and almost twice the typical mammalian genome [13]. In the same family, bread wheat (*Triticum aestivum*) adds another level of complexity with an allohexaploid genome that, over successive rounds of hybridization, has combined three diploid species with genomes the size of barley into a single huge

*Correspondence: Mario.Caccamo@tgac.ac.uk

¹The Genome Analysis Centre, Norwich Research Park, Norwich NR4 7UH, UK
Full list of author information is available at the end of the article

genome. Another interesting example is the evolution and domestication of the dicot Brassicaceae family that includes *Arabidopsis thaliana* and gave rise to a number of economically important vegetable and oil crops such as broccoli, cabbage and oilseed rape. In a relatively short evolutionary time, this family expanded into species with markedly different external structural and anatomy, offering a unique model for the study of the underlying genetics of morphogenesis [14]. The complexity of these plant genomes has, however, posed an important challenge to the use of sequencing technologies and the downstream computational analyses [15]. Over the past few years, dramatic advances in cytogenetic techniques, sequencing technologies and more sophisticated sequence analysis methods have impacted research in some of the most complex and challenging crops [16-21].

The opportunity of genomics

One of the immediate challenges of plant genomics is to mine more efficiently the vast amount of data that are being generated on a daily basis as more and more plant genomes are being sequenced [22]. There is, however, an opportunity emerging from the integration of the latest technologies with more traditional genetics and the fact that plant genomes are highly plastic. The easy manipulation of plant genomes is not only a convenient tool for research, but it is also the foundation for modern plant breeding. The possibility of generating doubled haploids to obtain fully homozygous individuals, for instance, can significantly speed-up breeding, providing an opportunity to go faster 'from the genome to the field'. Genome-wide association studies are an example of the application of genomics in both model plant species and crops that have helped to identify loci and alleles associated with complex traits [23-26]. The development of high-density molecular markers will be one of the most important tools for informing the design of the breeding programs of the future [27]. However, the most agronomically important traits, such as yield and drought tolerance, involve multiple genes and complex interactions with the environment, requiring more sophisticated breeding strategies such as genomic selection. The application of advanced genomics to improve breeding techniques in grass crops, for instance, will play a key role in securing affordable and nutritious food for an increasing human population.

The reviews, opinions and scientific work presented in this special issue of *Genome Biology* are a testament to the coming of age of plant genomics research and its applications.

Author details

¹The Genome Analysis Centre, Norwich Research Park, Norwich NR4 7UH, UK. ²Center for Applied Plant Sciences (CAPS), The Ohio State University, Columbus, OH 43210, USA.

Published: 27 June 2013

References

1. Peters CR, Vogel JC: Africa's wild C₄ plant foods and possible early hominid diets. *J Hum Evol* 2005, **48**:219-236.
2. Perry GH, Dominy NJ, Claw KG, Lee AS, Fiegler H, Redon R, Werner J, Villanea FA, Mountain JL, Misra R, Carter NP, Lee C, Stone AC: Diet and the evolution of human amylase gene copy number variation. *Nat Genet* 2007, **39**:1256-1260.
3. Kliebenstein DJ, Osbourn A: Making new molecules – evolution of pathways for novel metabolites in plants. *Curr Opin Plant Biol* 2012, **15**:415-423.
4. Sterck L, Rombauts S, Vadepeole K, Rouze P, Van de Peer Y: How many genes are there in plants (... and why are they there)? *Curr Opin Plant Biol* 2007, **10**:199-203.
5. Hamilton AJ, Baulcombe DC: A species of small antisense RNA in posttranscriptional gene silencing in plants. *Science* 1999, **286**:950-952.
6. Bogdanove AJ, Schornack S, Lahaye T: TAL effectors: finding plant genes for disease and defense. *Curr Opin Plant Biol* 2010, **13**:394-401.
7. The Arabidopsis Genome Initiative: Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* 2000, **408**:796-815.
8. International Rice Genome Sequencing Project: The map-based sequence of the rice genome. *Nature* 2005, **436**:793-800.
9. Goff SA, Ricke D, Lan TH, Presting G, Wang R, Dunn M, Glazebrook J, Sessions A, Oeller P, Varma H, Hadley D, Hutchison D, Martin C, Katagiri F, Lange BM, Moughamer T, Xia Y, Budworth P, Zhong J, Miguel T, Paszkowski U, Zhang S, Colbert M, Sun WL, Chen L, Cooper B, Park S, Wood TC, Mao L, Quail P, et al.: Adraft sequence of the rice genome (*Oryza sativa* L. ssp. Japonica). *Science* 2002, **296**:92-100.
10. Kim S, Plagnol V, Hu TT, Toomajian C, Clark RM, Ossowski S, Ecker JR, Weigel D, Nordborg M: Recombination and linkage disequilibrium in *Arabidopsis thaliana*. *Nat Genet* 2007, **39**:1151-1155.
11. Gan X, Stegle O, Behr J, Steffen JG, Drewe P, Hildebrand KL, Lyngsoe R, Schultheiss SJ, Osborne EJ, Sreedharan VT, Kahles A, Bohnert R, Jean G, Derwent P, Kersey P, Belfield EJ, Harberd NP, Kemen E, Toomajian C, Kover PX, Clark RM, Ratsch G, Mott R: Multiple reference genomes and transcriptomes for *Arabidopsis thaliana*. *Nature* 2011, **477**:419-423.
12. Becker C, Hagmann J, Müller J, Koenig D, Stegle O, Borgwardt K, Weigel D: Spontaneous epigenetic variation in the *Arabidopsis thaliana* methylome. *Nature* 2011, **480**:245-249.
13. International Barley Genome Sequencing Consortium, Mayer KF, Waugh R, Brown JW, Schulman A, Langridge P, Platzer M, Fincher GB, Muehlbauer GJ, Sato K, Close TJ, Wise RP, Stein N: A physical, genetic and functional sequence assembly of the barley genome. *Nature* 2012, **491**:711-716.
14. Bailey CD: Toward a global phylogeny of the Brassicaceae. *Mol Biol Evol* 2006, **23**:2142-2160.
15. Schatz M, Witkowski J, McCombie W: Current challenges in *de novo* plant genome sequencing and assembly. *Genome Biol* 2012, **13**:243.
16. Choulet F, Wicker T, Rustenholz C, Paux E, Salse J, Leroy P, Schlub S, Le Paslier MC, Magdelenat G, Gonthier C, Couloux A, Budav H, Breen J, Pumphrey M, Liu S, Kong X, Jia J, Gut M, Brunel D, Anderson JA, Gill BS, Appels R, Keller B, Feuillet C: Megabase level sequencing reveals contrasted organization and evolution patterns of the wheat gene and transposable element spaces. *Plant Cell* 2010, **22**:1686-1701.
17. 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**:705-710.
18. Ling HQ, Zhao S, Liu D, Wang J, Sun H, Zhang C, Fan H, Li D, Dong L, Tao Y, Gao C, Wu H, Li Y, Cui Y, Guo X, Zheng S, Wang B, Yu K, Liang Q, Yang W, Lou X, Chen J, Feng M, Jian J, Zhang X, Luo G, Jiang Y, Liu J, Wang Z, Sha Y, et al.: Draft genome of the wheat A-genome progenitor *Triticum urartu*. *Nature* 2013, **496**:87-90.
19. Jia J, Zhao S, Kong X, Li Y, Zhao G, He W, Appels R, Pfeifer M, Tao Y, Zhang X, Jing R, Zhang C, Ma Y, Gao L, Gao C, Spannagl M, Mayer KF, Li D, Pan S, Zheng F, Hu Q, Xia X, Li J, Liang Q, Chen J, Wicker T, Gou C, Kuang H, He G, Luo Y, et al.: *Aegilops tauschii* draft genome sequence reveals a gene repertoire for wheat adaptation. *Nature* 2013, **496**:91-95.
20. Wang X, Wang H, Wang J, Sun R, Wu J, Liu S, Bai Y, Mun JH, Bancroft I, Cheng F,

- Huang S, Li X, Hua W, Wang J, Wang X, Freeling M, Pires JC, Paterson AH, Chalhoub B, Wang B, Hayward A, Sharpe AG, Park BS, Weisshaar B, Liu B, Li B, Liu B, Tong C, Song C, Duran C, *et al*: **The genome of the mesopolyploid crop species *Brassica rapa***. *Nature Genetics*, 2011, **43**:1035-1039.
21. Nystedt B, Street NR, Wetterbom A, Zuccolo A, Lin YC, Scofield DG, Vezzi F, Delhomme N, Giacomello S, Alexeyenko A, Vicedomini R, Sahlin K, Sherwood E, Elfstrand M, Gramzow L, Holmberg K, Hällman J, Keech O, Klasson L, Koriabine M, Kucukoglu M, Käller M, Luthman J, Lysholm F, Niittylä T, Olson A, Rilakovic N, Ritland C, Rosselló JA, Sena J, *et al*: **The Norway spruce genome sequence and conifer genome evolution**. *Nature* 2013, **497**:1-6.
 22. De Luca V, Salim V, Atsumi SM, Yu F: **Mining the biodiversity of plants: a revolution in the making**. *Science* 2012, **336**:1658-1661.
 23. Huang X, Zhao Y, Wei X, Li C, Wang A, Zhao Q, Li W, Guo Y, Deng L, Zhu C, Fan D, Lu Y, Weng Q, Liu K, Zhou T, Jing Y, Si L, Dong G, Huang T, Lu T, Feng Q, Qian Q, Li J, Han B: **Genome-wide association study of flowering time and grain yield traits in a worldwide collection of rice germplasm**. *Nat Genet* 2011, **44**:32-39.
 24. Abe A, Kosugi S, Yoshida K, Natsume S, Takagi H, Kanzaki H, Matsumura H, Yoshida K, Mitsuoka C, Tamiru M, Innan H, Cano L, Kamoun S, Terauchi R: **Genome sequencing reveals agronomically important loci in rice using MutMap**. *Nat Biotechnol* 2012, **30**:174-178.
 25. Huang X, Wei X, Sang T, Zhao Q, Feng Q, Zhao Y, Li C, Zhu C, Lu T, Zhang Z, Li M, Fan D, Guo Y, Wang A, Wang L, Deng L, Li W, Lu Y, Weng Q, Liu K, Huang T, Zhou T, Jing Y, Li W, Lin Z, Buckler ES, Qian Q, Zhang QF, Li J, Han B: **Genome-wide association studies of 14 agronomic traits in rice landraces**. *Nat Genet* 2010, **42**:961-967.
 26. Segura V, Imsson BJVA, Platt A, Korte A, Seren U, Long Q, Nordborg M: **An efficient multi-locus mixed-model approach for genome-wide association studies in structured populations**. *Nat Genet* 2012, **44**:1-8.
 27. Halpin, C: **Gene stacking in transgenic plants – the challenge for 21st century plant biotechnology**. *Plant Biotechnol J* 2005, **3**:141-155.

doi:10.1186/gb-2013-14-6-403

Cite this article as: Caccamo M, Grotewold E: **Turning over a new leaf in plant genomics**. *Genome Biology* 2013, **14**:403.