



REVIEW

Time to crop: jumping from biological models to crop biotechnology

Vagner Augusto Benedito^{1*}

Received 30 January 2007

Accepted 09 April 2007

ABSTRACT - *Model organisms are useful systems to understand the complexity of mechanisms behind life. In the past, each aspect of biology possessed its own model species, but with the advent of molecular biology, it was necessary to investigate with great intensity a few chosen species aiming at having a whole idea on how biological mechanisms function altogether and Arabidopsis thaliana was elected as the prime working model in plant science. With the completion of its genome sequencing, other plant models rose as workable systems at the molecular level and now one can see a blossoming of plant species, including many crops, that have at least part of their genome sequence unveiled. Plant breeding is one of the fields that will benefit enormously from these studies, since knowing the genome sequence of a given species can lead to link the Mendelian genetics to its molecular bases, making breeding programs to advance more quickly.*

Key-words: evolution, functional genetics, genomics, molecular breeding, physical mapping.

INTRODUCTION

Among ~300,000 known land plant species, just 30 crop species dominate food production presently (Hedges 2002, Ruane and Sonnino 2006) and with an ever-growing population to be fed, plant biotechnology will occupy a more and more distinguished position in breeding programs aiming at not solely higher yields, but also introducing novel characteristics beyond their genetic pools.

Works on model species in several kingdoms do not only reveal biological aspects of closely related species,

but an integration of such studies in a broad perspective also allows understanding how life evolved, the universal laws governing cell biology and, the integrative and derivative aspects of living species, which will ultimately enable scientists and breeders to manipulate genomes towards their goals.

It is understandable that the molecular era of biology investigation demanded firstly a stronger focus in the simplest model among the plant species to make possible an integration of all information that would be generated. *Arabidopsis thaliana* came out as a workable option for genetic studies and it was chosen later as the

¹ Plant Biology Division, The Samuel Roberts Noble Foundation, 2510 Sam Noble Parkway, 73401, Ardmore, Oklahoma, USA. *Email: vbenedito@noble.org

first plant for genome sequencing. Since *Arabidopsis* is not commercially important, it was chosen as a model solely for its inherent characteristics in order to speed up the achievements of proposed goals. It was never forgotten, however, that one of the most important aspects of unraveling plant biology through model species is the application of this knowledge to breed species that serve humanity as food and other resources, such as energy precursors, fiber, wood and other raw materials.

Some other workable models are now following to have a better representation among plant species. Even though investigation on plant model species still goes on, and it will be so for a long time due to the complex nature of plant biology, the time has come to profusely apply this generated knowledge into crop species. Whereas genome sequencing of some crops are being currently performed and others are being proposed (Paterson 2006), many aspects of plant biology revealed by *Arabidopsis* as well as rice projects can already be applied into crop biotechnology investigation.

Arabidopsis as a contemporary model for plant biology

Taking into account that a simple model would allow understanding the bases the biological mechanisms of plants, several characteristics made the temperate weed species *Arabidopsis thaliana* (thale cress) as the most viable alternative for genetics and molecular investigation: its rapid life cycle (5-6 weeks from seed to seed), small size (20-cm tall), prolificacy (thousands of seeds per plant), simple genetics and genome (diploid, autogamous species with 125 Mb dispersed in five chromosomes presenting very little repetitive DNA), easy cultivation and lab manipulation (crossings, genetic transformation through floral dipping and chemical selection), and an extensive catalogue of mutants.

Plant molecular biology reached its maturity with the publication of the *Arabidopsis* genome sequence and structure at the end of 2000 (*Arabidopsis* Genome Initiative 2000). Now, the main challenge is to assess the biological function of each one of its 25,500 predicted genes, which is expected to be accomplished by 2010. *Arabidopsis*, thus, is located at the center of predictive and comparative plant biology (Bevan and Walsh 2004, 2005) aiming at serving as a reference to other plant species. Moreover, sequencing efforts of other ecotypes will also help to enlighten the mechanisms of natural variation at the genomic level (Jander et al. 2002) which shall lead to a better

understanding of how genotypes evolve within a single species.

Along with functional genomics, proteomics is a fast growing field which certainly will help to elucidate functional and biochemical properties of plant proteins, leading us to an era of a still-infant structural biology. Coupling protein shape and chemical properties with its mode of action is still a difficult task (especially for proteins with unknown functions) but this is seen as the next frontier to be conquered by molecular biologists (Agrawal and Rakwal 2006).

The genome sequencing of *Arabidopsis* has opened up an avenue of techniques and bioinformatics methods that accelerated studies in other plant species. Obviously, *Arabidopsis* does not cover all aspects of plant biology and other models are necessary for contrasting characteristics and expanding the knowledge on plant biology, which will ultimately contribute to crop breeding, since the most important aspects studied in plant biology, even in model species and regardless the research focus, are related to agricultural importance, such as plant-pathogen interactions, stress tolerance, photosynthesis, and mineral nutrition.

Other models are following the trail

While the *Arabidopsis* genome project was being accomplished, other projects were raised intending to cover distinctive aspects of plant biology. Rice was an obvious choice, since, besides being an important staple food worldwide, as a monocot, it is distantly related (belonging to another taxonomic class), it is diploid and autogamous, with a relative short generation time (3-6 months), a small genome (~430 Mb divided within 12 chromosomes), and presents established genetic transformation procedures. In this way, the results gathered in this sequencing project can be directly applied into molecular breeding programs of this species. The draft sequences of two rice subspecies (the tropical *indica* cv. Nipponbare and the temperate *japonica* cv. 93-11) were published in 2002 (Goff et al. 2002, Yu et al. 2002). The International Rice Genome Sequencing Project (IRGSP), after releasing the first draft of *japonica* sequence, improved both the genome sequencing and its annotation (IRGSP 2005). Two private biotechnology companies, Monsanto and Syngenta, had also independently developed rice sequencing projects and later joined the IRGSP (Buell 2002, Vij et al. 2006). The identification of single-nucleotide polymorphisms (SNP)

and simple sequence repeats (SSR) assists in identifying reliable markers for important agricultural traits and is highly significant for direct application in breeding programs. Moreover, to help functional analysis of the 37,500 predicted rice genes, the research community can rely upon several mutant collections (Miyao et al. 2003, Hirochika et al. 2004), saturated genetic maps, EST sequences as well as Affymetrix chips for gene expression profiling.

Apart the 90% of Arabidopsis proteins that share homology with rice, 2,859 rice genes could not be related to the Arabidopsis genome (IRGSP 2005) and are thought to have evolved beyond recognition and possibly gained novel functions in the last 135-250 million years ago (MYA) that separate the last common ancestor between these two species (Schneider et al. 2004). Rice has over Arabidopsis several distinctive aspects: 1) it is a monocot species; 2) it forms caryopses as dry fruits in a spike; 3) it presents the development of a large starchy endosperm in seeds; 4) its flowering is induced by long nights whereas the opposite occurs in Arabidopsis; 5) it does not require vernalization for germination; 6) it is able to establish symbiotic relationships with mycorrhiza whereas Arabidopsis is not. Rice belongs to the grass family (Poaceae, which encompasses ~10,000 species and is reckoned to be evolved in the last 65 million years), thus, serving as a better model platform for the ~80,000 monocot species currently distributed in over 100 families, and especially for cereals (such as maize, millet, sorghum, barley, wheat, oat), with which rice share very close genetic relationship (Figure 1, Song et al. 2002, Bennetzen and Ma 2003).

In the same avenue opened by Arabidopsis and rice, two legume sequencing projects are underway: *Medicago truncatula* and *Lotus japonica* (Young et al. 2005). Both belong to the subfamily Papilionoidea, but to different tribes, they are temperate legumes and thought to have shared a last common ancestor circa 40 MYA (Doyle and Luckow 2003). Legume species, besides a high importance in human and animal nutrition, also feature symbiosis with nitrogen-fixing bacteria and root nodulation. Instead of information redundancy, comparative genomics between these two model legume genomes will bring novel insights on plant genome evolution and organization, and especially on how rhizobia symbiosis evolved, as well as the genetic synteny among crop legumes that will lead to a better understanding and manipulation of legume breeding, since these two species also show contrasting traits on nodule development, flowering determination,

pod morphology, and growth habit. Genome annotation and functional projects are going along the sequencing (Tadege et al. 2005, Udvardi et al. 2005, Benedito et al. 2006, Town 2006). An oligonucleotide chip (Affymetrix Medicago GeneChip®) based on expressed sequences and genome information was recently released and also encompasses probes for *Medicago sativa* (alfalfa) and *Sinorhizobium meliloti*, their nitrogen-fixing symbiont. It was recently demonstrated that *M. truncatula* probe sets can also be applied for alfalfa transcriptional profiling (Tefaye et al. 2006). A recent announcement stated that the soybean genome sequencing is also underway. Soybean is an important tropical legume estimated to have diverged 50 MYA from *M. truncatula* (Mudge et al. 2005).

Another taxonomic order that is gaining momentum is the Solanales, which includes important crop species as well tobacco and many medicinal, toxic and weedy species. Moreover, the Solanales order is divergent from the Brassicales and closer to Asterales and even Gentianales (Figure 1), allowing a knowledge expansion within dicot species. Tomato (recently renamed *Solanum lycopersicum*) is already an established model species for fleshy, climacteric fruit development (White 2002) and hormone physiology, and shows a good potential for expanding to other aspects of plant development with the use of a dwarf variety called MicroTom (Meissner et al. 1997) that shows a miniature size (8-cm tall) and relative short life cycle (12 weeks from seed to seed). Its *in vitro* regeneration and genetic transformability has also been improved recently (Lima et al. 2004, Dan et al. 2006, Sun et al. 2006). Over Arabidopsis, MicroTom is more suitable for studies involving mycorrhiza symbiosis, disease resistance and interspecific genetic studies within the Solanaceae crops, because of a close synteny within this family. There are also large EST collections and public databases, commercial oligonucleotide microarray (Affymetrix), introgression lines (Canady et al. 2005), saturated genetic maps and ongoing efforts for genome sequencing (The International Tomato Sequencing Project; Sol Genomics Network: Mueller et al. 2005). Moreover, expanding the knowledge of molecular developmental mechanisms in tomato will help to extend our evolutionary comprehension and make predictions more accurately on crops belonging to the Asterid subclass, that is thought to have split from Rosid subclass (of which Arabidopsis belongs) circa 100-150 MYA (Yang et al. 1999).

The potato research goes in the same pace as for tomato, with a genome sequencing project in progress

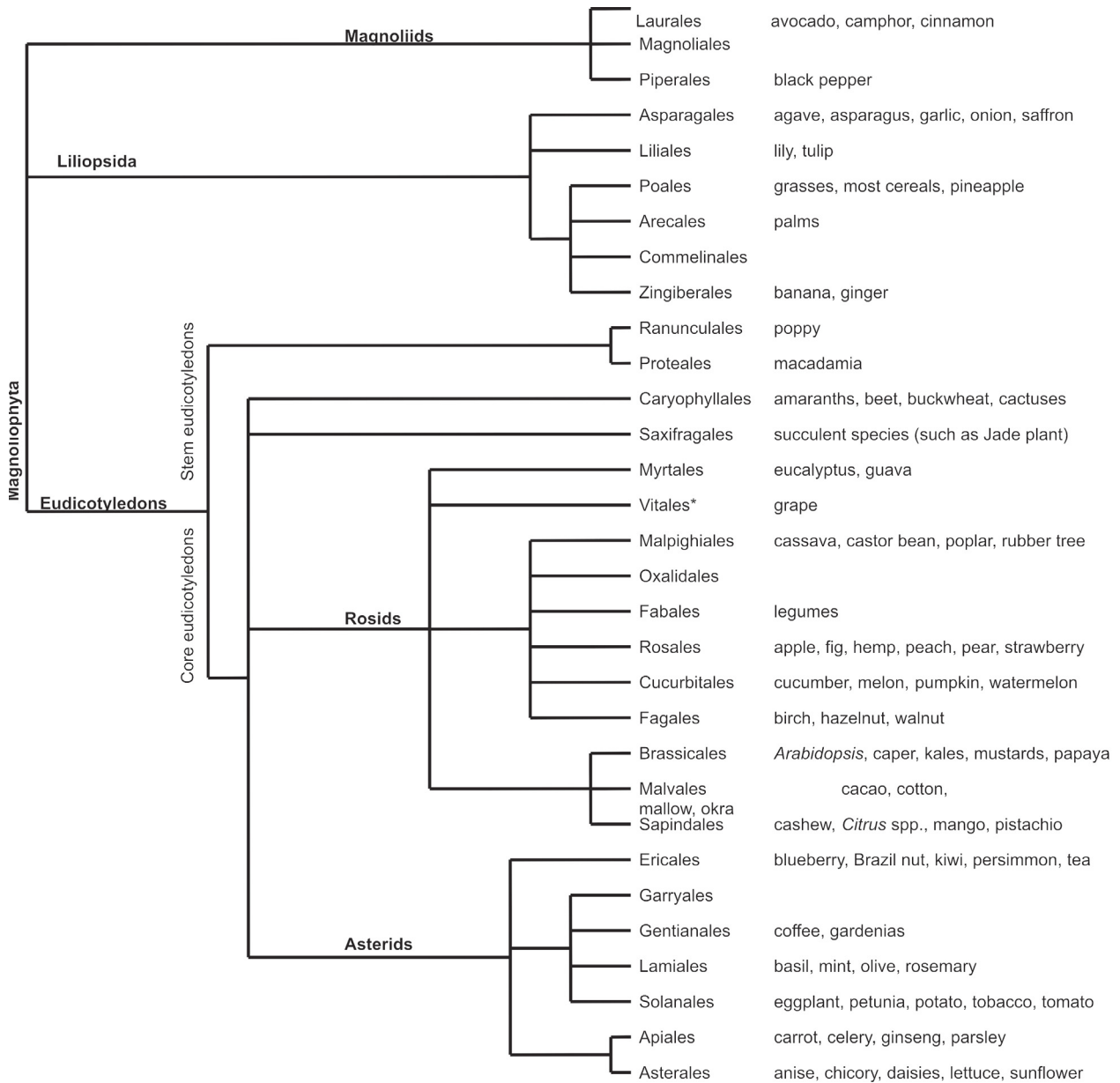


Figure 1. Phylogenetic relationship of model species and cultivated crops. The tree was constructed with the ‘Taxonomy Common Tree’ tool from the National Center for Biotechnology Information (<http://www.ncbi.nlm.nih.gov/Taxonomy/CommonTree/wwwcmt.cgi>) and visualized by the TreeView software (Page 1996). (*) the order Vitales is non-existent in the classification system used by the Angiosperm Phylogeny Group, which places the Vitaceae family under the Rosids into a general group called “*rosids incertae sedis*”, meaning “uncertain place among the Rosids”

(The NSF Potato Genome Project) and every genetic aspect investigated intensively (van Os et al. 2006, Mullins et al. 2006). Here, however, more attention is given to agronomical aspects of potato cultivation, such as resistance or tolerance to diseases as well as tuberization, for which potato is considered a model organism. *Petunia hybrida*, another Solanaceae, is also a reference species for plant biology, since many studies on flavonoid biosynthesis, self-incompatibility, and meristem development were carried out in this species and a whole range of genetic resources and experimental procedures is available (Angenent et al. 2005, Gerats and Vandenbussche 2005).

Cotton (Malvaceae family) is being considered for genome sequencing and is a model for trichome development, since it bears very long seed epidermal fibers. Recently, it has been demonstrated that a MYB transcription factor expressed in developing cotton seeds was able to genetically complement an Arabidopsis mutant (*gl1*) defective for a transcription factor involved in trichome development (Wang et al. 2004). It was also recently revealed a relatively conserved synteny between cotton and Arabidopsis (Rong et al. 2005). Moreover, the Malvales order is a sister clade of the Sapindales, which includes important fruit species.

The maize genome sequencing project is much more modest than the Arabidopsis and rice ones. Yet, maize features important characteristics not covered by these two model species, such as C_4 photosynthesis and monoecious sexuality that will certainly be very useful for crop improvement and to broaden our understanding on cereal species. The gaps in understanding C_4 photosynthesis was attributed to the lack of a good genetic model with short life span (Brown et al. 2005), since maize and *Flaveria bidentis* (an Asteraceae) are considered the current model species for the C_4 physiology, which emphasizes the importance of good models in fundamental research. The genus *Cleome* (a Capparaceae, the caper family) has been claimed as a suitable model for C_4 photosynthesis studies because it encloses both C_3 and C_4 species (allowing a close comparative analysis) and is the closest C_4 species related to Arabidopsis (Brassicales order). Unfortunately, there are not yet much genetic resources, mutants or an efficient transformation protocol for this species (Brown et al. 2005).

Sorghum bicolor, a C_4 grass that diverged circa 15 MYA from maize and 50 MYA from rice, has also been target of attention. In a way, it can also work as a model for

closely related species, such as sugarcane, from which it diverged just 3 MYA. Sorghum has highly resolved genetic, physical and cytological maps (Mullet et al. 2001), largely important for comparative genomics and breeding programs. Agronomically, one of the most important traits of sorghum is its tolerance to environmental stresses, such as drought and heat. The discovery of the molecular background behind adaptation to arid conditions (genes and their allelic variation) would allow transferring these traits to other crops.

Sugarcane has a very complex genetics with a large genome (7,440 Mb) split into a high number of chromosomes ($n=40$) featuring aneuploid polyploidy and interspecies crossings in modern varieties (100 d" 2n e" 130, Grivet and Arruda 2001). Nevertheless, sugarcane is syntenic to grasses in general (Freeling 2001) and has been subjected to extensive genetic mapping through molecular markers, EST project and microarrays, and an oligonucleotide chip is recently available. As the demand for clean energy increases in the world, sugarcane is seen as a valuable alternative for ethanol production and this can boost the molecular research of this crop in the near future (Marris 2006).

Expressed Sequence Tag (EST) projects are being generally performed for many crop species in order to reveal sequences and build up clone banks. These data are deposited in public databases, most of all, at the Computational Biology and Functional Genomics Laboratory (but until recently at The Institute for Gene Resources, TIGR) and the National Center for Biotechnology Information (NCBI). What was once a great scientific challenge, thanks to the projects with model species and the advances from them, presently an EST project requires just a little more than resource allocation to be accomplished.

Although we can see nowadays a flourishing of species considered for genomics investigation (Table 1), there are still many gaps of knowledge in plant biology to be fulfilled (Jackson et al. 2006) and it will only be accomplished when a broader range of plant species evenly distributed among diverse taxonomic groups is explored molecularly.

Novel approaches are imperative to understand crop biology

Where as certain research fields such as molecular markers or tissue culture and genetic transformation are

Table 1. Genome characteristics and state of genomic research on model and crop species

Species	Genome size (Mb)	Chromosome set (ploidy)	Availability of genomic tools	Online information
<i>Arabidopsis thaliana</i>	125	n=5 (2x)	Genome sequence completed; Affymetrix chip; Mutant collections; >350,000 ESTs	arabidopsis.org
Apple	750	n=17 (2x)	local EST projects	genomics.msu.edu/fruitdb/analyses/apple.shtml
Banana	600	n=11 (2x to 4x)	Genome sequence underway; Germplasm collections; local EST projects	musagenomics.org/index.php?page=home
Barley	5,000	n=7 (2x)	Affymetrix chip; mutant collections; >370,000 ESTs	barleyworld.org/northamericanbarley.php
Cassava	760	n=18 (2x)	Genome sequence underway; Mutant collections; local EST projects	jgi.doe.gov/sequencing/why/CSP2007/cassava.html
Castor bean	400	n=10 (2x)	Genome sequence underway; local EST projects;	msc.tigr.org/r_communis/index.shtml
Cocoa	400	n=10 (2x)	~6,000 ESTs	cirad.fr/presentation/programmes/biotrop/resultats/biositecirad/cocoa.htm
Coffee	1,300	n=11 (2 or 4x)	BAC library; local EST projects	coffeedna.net/index.php?url=group
Common bean	630	n=11 (2x)	>20,000 ESTs; Phaseomics Project	lge.ibi.unicamp.br/cafe phaseolus.net
Cotton	2100	n=26 (2 or 4x)	BAC sequencing; >210,000 ESTs	cottongenomecenter.ucdavis.edu/pub/template.asp genome.clemson.edu/projects/cotton
<i>Eucalyptus globulus</i>	600	n=11(2x)	Genome sequence underway; local EST projects	ieugc.up.ac.za/index.htm
Grape	500	n=19 (2x)	International Grape Genome Program; Affymetrix chip; ~140,000 ESTs	vitaceae.org
Lettuce	2,600	n=9 (2x)	Compositae Genome Project; ~70,000 ESTs	compgenomics.ucdavis.edu/index.htm; compositdb.ucdavis.edu
<i>Lotus japonicus</i>	470	n=6 (2x)	Genome sequence underway; mutant collection; ~110,000 ESTs	kazusa.or.jp/lotus; shigen.nig.ac.jp/lotusjaponicus/index_e.html; comparative-legumes.org
Maize	2,400	n=10 (2x)	Affymetrix chip; BAC sequencing; mutant collection; ~570,000 ESTs	maizegdb.org
<i>Medicago truncatula</i>	470	n=8 (2x)	Genome sequence underway; Affymetrix chip; mutant collections; ~230,000 ESTs	tigr.org/tigr-scripts/medicago/IMGAG/imgag_annotator.pl medicago.toulouse.inra.fr/Mt/ESTtigr.org/tdb/e2k1/mta1 noble.org/medicago
Orange (<i>Citrus silensis</i>)	380	n=9 (2x)	Genome sequence underway – International Citrus Genome Consortium; Affymetrix chip; mutant collections; >220,000 ESTs	int-citrusgenomics.org
Papaya	372	n=9 (2x)	Genome sequence underway	cgpbr.hawaii.edu/papaya
Pepper (<i>Capsicum annuum</i>)	2,700	n=12 (2x)	BAC library; ~30,000 ESTs	plant.pdrc.re.kr/ks200201/pepper.html
Poplar	480	n=19(2x)	Genome Sequence underway; Affymetrix chip; >260,000 ESTs	genome.jgi-psf.org/Poptr1/Poptr1.home.html
Potato	840	n=12 (4x)	Genome sequence underway; mutant collections; ~200,000 ESTs	tigr.org/tdb/potato/profiling_service2.shtml
Rice	430	n=12(2x)	Genome sequence completed; Affymetrix chip; mutant collections; >270,000 ESTs	shigen.nig.ac.jp/rice/oryzabase/top/top.jsp rgp.dna.affrc.go.jp irri.org
Rye	9,100	n=7(2x)	>9,000 ESTs	bafz.de/primer/start.php3
Sorghum	760	n=10 (2x)	Genome sequence underway; Affymetrix chip; mutant collections; ~200,000 ESTs	jgi.doe.gov/sequencing/why/CSP2006/sorghum.html
Soybean	1,200	n=20 (2x)	Genome sequencing announced; Affymetrix chip; mutant collections; >330,000 ESTs	shigen.nig.ac.jp/bean/glycinesoja/top/top.jsp?language=en soybean.ccgb.umn.edu
Sugarcane	7,440	n=40 (complex)	Affymetrix chip; germplasm collections; >250,000 ESTs	sucest.lad.ic.unicamp.br/en/
Sunflower	3,000	n=17 (2x)	Mutant collections; ~60,000 ESTs	compgenomics.ucdavis.edu/overview.htm cgpdb.ucdavis.edu
Tomato	950	n=12 (2x)	Partial genome sequence underway; Affymetrix chip; mutant collection; >160,000 ESTs	sgn.cornell.edu/about/tomato_sequencing.pl genome.clemson.edu/projects/other/tomato tgrc.ucdavis.edu
Wheat	16,000	n=21 (2x)	Partial genome sequence underway; Affymetrix chip; mutant collection; ~600,000 ESTs	wheat.pw.usda.gov/NSF/htmlversion.html

not bounded to a single species or phylogenetic group, basic aspects of plant biology are generally first investigated in dedicated model species. This approach led to impressive findings in the last decades on the mechanisms governing plant development, but at the same time, let gaps in our knowledge on how evolution drove the immense diversity of traits found among plant species and the variety of developmental mechanisms underneath.

In general, conserved biological mechanisms found in model species are likely to be assessed in crop species as well, whereas unique crop traits will have to be subjected to experimentation in order to understand the programs governing the development of such characteristics. These studies will be more difficult as divergent the feature is from the traits found in model or species already characterized molecularly. For that, new areas are being developed to congregate knowledge from several fields altogether and pose a more holistic view on plant biology.

Evolutionary Developmental Biology is a growing research field tackling how traits evolved during evolution and how genetic changes reprogram developmental pathways causing phenotype modification (Irish and Benfey 2004). Despite being a good starting point, sometimes gene sequence homology between crop and model species can lead to wrong assumptions and a close investigation about functional characteristics of gene products as well as their temporal expression profiling is necessary before concluding on functional orthology and cellular roles. Heterologous systems are being useful to tackle genetic functional analysis in crops difficult to manipulate molecularly (Benedito et al. 2004, Senthil-Kumar et al. 2006, Senthil-Kumar and Udayakumar 2006).

Functional Genomics is central to plant molecular biology and key for crop molecular breeding, since it aims at assessing the roles of gene products within the cell and relating them to plant phenotypes. For that, allelic variation (natural or induced) is important to demonstrate degrees of gene activity instead of solely a black-or-white picture derived from null and wild-type alleles.

One aspect of Comparative Genetics is the investigation on the synteny (genetic collinearity in chromosome fragments) among species, what is a key point to establish their phylogenetic relationship. For example, it was already demonstrated microsynteny among dicot and monocot species, which was confirmed through genome sequence comparison between *Arabidopsis* and rice (Salse et al. 2002). The genetics of *Arabidopsis* has

also been contrasted against cotton and revealed considerable synteny (Bowers et al. 2003). Nevertheless, extensive synteny between *Arabidopsis* and the legumes pea and alfalfa failed to be demonstrated (Delseny 2004), although microsynteny between *Arabidopsis* and *M. truncatula* was already shown (Zhu et al. 2003). Genome sequencing also revealed that even a simple genome such as of *Arabidopsis* could be evolved from duplication events, leading to numerous redundant genes (*Arabidopsis* Genome Initiative 2000).

Comparative mapping between model and crop species has the potential to speed up the assessment of orthology and collinearity in complex genomes. Apart studies carried out in related species, *Arabidopsis* and rice sequence information was used to anchor markers, such as single-nucleotide polymorphisms (SNPs), simple sequence repeats (SSRs) or conserved orthologous sets (COS), from ESTs in homologous loci in several crop species, such as sugar beet, potato, sunflower, cabbages, wheat, maize and peach (Jung et al. 2006; Varshney et al. 2005; Delseny 2004; van Buuren et al. 2002). A strategy to use Comparative Genetics to unravel single-nucleotide polymorphisms (SNPs) between model species and orphan crops was recently proposed (Feltus et al. 2006). However, difficulties may be faced when the markers used lies on redundant syntenic blocks, as reported in a work comparing the *Arabidopsis* and potato genomes (Gebhardt et al. 2003).

Natural variation is usually the most important resource in breeding programs. Molecular markers can be used to evaluate the genetic diversity existent in germplasm collections (Ruane and Sonnino 2006). The understanding of molecular mechanisms behind natural variation is still sparing and some efforts are being directed on several ecotypes of *Arabidopsis thaliana* and species within the same genus (Koornneef et al. 2004, Tonsor et al. 2005, Alonso-Blanco et al. 2006, Hannah et al. 2006). Related species such as *Arabidopsis lyrata*, *A. halleri* and *A. arenosa* will certainly be useful to better understand mechanisms such as self-compatibility, heavy metal tolerance and inter-specific hybridization (Claus and Koch 2006) that are also important aspects for crop cultivation.

Whereas genomes of the most important crops have been intensively studied, making them easier for comparison with model genomes, the so-called orphan crops (defined as the less studied or yet neglected species) still need input of molecular data to make comparison possible. This work is as more difficult as the crop is more

distantly related to a better studied species and has a more complex genome.

PERSPECTIVES

The unveiling of whole genome sequences and the advances in bioinformatics programs brought to a better comprehension of how plant genome is organized, evolved, diverged and is related to one another. Many genetic pathways have also been revealed at the molecular

level from studies carried out in model species. We expect that with the accomplishment of functional projects in Arabidopsis, functional analyses in crop species will be made easier and faster, and the differential existent between crops and models will be ripe for fine molecular evaluation. From these studies, we expect to get a better idea of each crop characteristics and be able to work better on single traits as the Molecular Crop Breeding walks towards its maturity.

A hora da colheita: de modelos biológicos à biotecnologia de plantas

RESUMO - Modelos biológicos são sistemas úteis para entender a complexidade de mecanismos atrás da vida. No passado, cada aspecto da biologia possuía sua própria espécie modelo. Com o advento da biologia molecular, se fez necessária a investigação com maior intensidade de apenas algumas espécies escolhidas com o intuito de uma compreensão geral do funcionamento dos mecanismos biológicos. A espécie *Arabidopsis thaliana* foi eleita como o primeiro modelo de trabalho na ciência molecular de plantas. Com o término do seqüenciamento de seu genoma, outros modelos vegetais emergiram como sistemas viáveis no nível molecular e hoje se vê muitas espécies vegetais, incluindo aquelas de interesse agrônômico, que estão tendo seu genoma revelado, ao menos parcialmente. O melhoramento de plantas é um dos campos que serão beneficiados enormemente com esses estudos, uma vez que o conhecimento da seqüência do genoma de uma dada espécie pode ajudar a ligar a genética mendeliana às suas bases moleculares, fazendo com que os programas de melhoramento avancem mais rapidamente.

Palavras-chaves: evolução, genética funcional, genômica, mapeamento físico, melhoramento molecular

REFERENCES

- Agrawal GK and Rakwal R (2006) Rice proteomics: a cornerstone for cereal food crop proteomes. **Mass Spectrometry Reviews** 25: 1-53.
- Alonso-Blanco C, Koornneef M and van Ooijen JW (2006) QTL analysis. **Methods in Molecular Biology** 323: 79-99.
- Angenent GC, Stuurman J, Snowden KC and Koes R (2005) Use of *Petunia* to unravel plant meristem functioning. **Trends in Plant Science** 10: 243-250.
- Arabidopsis Genome Initiative (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. **Nature** 408: 796-815.
- Benedito VA, Dai X, He J, Zhao PX and Udvardi MK (2006) Functional genomics of plant transporters in legume nodules. **Functional Plant Biology** 33: 731-736.
- Benedito VA, Visser PB, Angenent GC and Krens FA (2004) The potential of virus-induced gene silencing for speeding up functional characterization of plant genes. **Genetics and Molecular Research** 30: 323-341.
- Bennetzen JL and Ma J (2003) The genetic collinearity of rice and other cereals on the basis of genomic sequence analysis. **Current Opinion in Plant Biology** 6: 128-133.
- Bevan M and Walsh S (2004) Positioning Arabidopsis in plant biology. A key step toward unification of plant research. **Plant Physiology** 135: 602-606.
- Bevan M and Walsh S (2005) The *Arabidopsis* genome: a foundation for plant research. **Genome Research** 15: 1632-1642.
- Bowers JE, Chapman BA, Rong J and Paterson AH (2003) Unravelling angiosperm genome evolution by phylogenetic analysis of chromosomal duplication events. **Nature** 422: 433-438.
- Brown NJ, Parsley K and Hibberd JM (2005) The future of C₄ research – maize, *Flaveria* or *Cleome*? **Trends in Plant Science** 10: 215-221.
- Buell CR (2002) Obtaining the sequence of the rice genome and lessons learned along the way. **Trends in Plant Science** 7: 538-542.

- Canady MA, Meglic V and Chetelat RT (2005) A library of *Solanum lycopersicoides* introgression lines in cultivated tomato. **Genome** **48**: 685-697.
- Claus MJ and Koch MA (2006) Poorly known relatives of *Arabidopsis thaliana*. **Trends in Plant Science** **11**: 449-459.
- Dan Y, Yan H, Munyikwa T, Dong J, Zhang Y and Armstrong CL (2006) MicroTom – a high-throughput model transformation system for functional analyses. **Plant Cell Reports** **25**: 432-441.
- Delseny M (2004) Re-evaluating the relevance of ancestral shared synteny as a tool for crop improvement. **Current Opinion in Plant Biology** **7**: 126-131.
- Doyle JJ and Luckow MA (2003) The rest of the iceberg: legume diversity and evolution in a phylogenetic context. **Plant Physiology** **131**: 900-910.
- Feltus FA, Singh HP, Lohithaswa HC, Schulze SR, Silva TD and Paterson AH (2006) A comparative genomics strategy for targeted discovery of single-nucleotide polymorphisms and conserved-noncoding sequences in orphan crops. **Plant Physiology** **140**: 1183-1191.
- Freeling M (2001) Grasses as a single genetic system: reassessment. **Plant Physiology** **125**: 1191-1197.
- Gebhardt C, Walkemeier B, Henselewski H, Barakat A, Delseny M and Stuker K (2003) Comparative mapping between potato (*Solanum tuberosum*) and *Arabidopsis thaliana* reveals structurally conserved domains and ancient duplications in the potato genome. **Plant Journal** **34**: 529-541.
- Gerats T and Vandenbussche M (2005) A model system for comparative research: Petunia. **Trends in Plant Science** **10**: 251-256.
- Goff SA, Ricke D, Lan TH et al. (2002) A draft sequence of the rice genome (*Oryza sativa* L. ssp. *japonica*). **Science** **296**: 92-100.
- Grivet L and Arruda P (2001) Sugarcane genomics: depicting the complex genome of an important tropical crop. **Current Opinion in Plant Biology** **5**: 122-127.
- Hannah MA, Wiese D, Freund S, Fiehn O, Heyer AG and Hincha DK (2006) Natural genetic variation of freezing tolerance in *Arabidopsis*. **Plant Physiology** **142**: 98-112.
- Hedges SB (2002) The origin and evolution of model organisms. **Nature Reviews Genetics** **3**: 838-849.
- Hirochika H, Guiderdoni E, An G, Hsing YI, Eun MY, Han CD, Upadhyaya N, Ramachandran S, Zhang Q, Pereira A, Sundaresan V and Leung H (2004) Rice mutant resources for gene discovery. **Plant Molecular Biology** **54**: 325-334.
- IRGSP - International Rice Genome Sequencing Project (2005) The map-based sequence of the rice genome. **Nature** **436**: 793-800.
- Irish VF and Benfey PN (2004) Beyond *Arabidopsis*. Translational Biology meets Evolutionary Developmental Biology. **Plant Physiology** **135**: 611-614.
- Jackson S, Rounsley S and Purugganan M (2006) Comparative sequencing of plant genomes: choices to make. **Plant Cell** **18**: 1100-1104.
- Jander G, Norris SR, Rounsley SD, Bush DF, Levin IM and Last RL (2002) *Arabidopsis* map-based cloning in the post-genome era. **Plant Physiology** **129**: 440-450.
- Jung S, Main D, Staton M, Cho I, Zhebentyayeva T, Arus P and Abbott A (2006) Synteny conservation between the *Prunus* genome and both the present and ancestral *Arabidopsis* genomes. **BMC Genomics** **7**: 81-98.
- Koornneef M, Alonso-Blanco C and Vreugdenhil D (2004) Naturally occurring genetic variation in *Arabidopsis thaliana*. **Annual Review of Plant Biology** **55**: 141-172.
- Lima JE, Carvalho RF, Tulmann Neto A, Figueira A and Peres LEP (2004) Micro-MsK: a tomato genotype with miniature size, short life cycle, and improved *in vitro* shoot regeneration. **Plant Science** **167**: 753-757.
- Marris E (2006) Sugar cane and ethanol: drink the best and drive the rest. **Nature** **444**: 670-672.
- Meissner R, Jacobson Y, Melame S, Levyatuv S, Shalev G, Ashri A, Elkind Y and Levy A (1997) A new model system for tomato genetics. **Plant Journal** **12**: 1465-1472.
- Miyao A, Tanaka K, Murata K, Sawaki H, Takeda S, Abe K, Shinozuka Y, Onosato K and Hirochika H (2003) Target site specificity of the *Tos17* retrotransposon shows a preference for insertion within genes and against insertion in retrotransposon-rich regions of the genome. **Plant Cell** **15**: 1771-1780.
- Mudge J, Cannon S, Kalo P, Oldroyd GED, Roe BA, Town CD and Young ND (2005) Highly syntenic regions in the genomes of soybean, *Medicago truncatula*, and *Arabidopsis thaliana*. **BMC Plant Biology** **5**: 15.
- Mueller LA, Solow TH, Taylor N, Skwarecki B, Buels R, Binns J, Lin C, Wright MH, Ahrens R, Wang Y, Herbst EV, Keyder ER, Menda N, Zamir D and Tanksley SD (2005) The SOL Genomics Network: a comparative resource for Solanaceae biology and beyond. **Plant Physiology** **138**: 1310-1317.
- Mullet JE, Klein RR and Klein P (2001) *Sorghum bicolor* – an important species for comparative grass genomics and a source of beneficial genes for agriculture. **Current Opinion in Plant Biology** **5**: 118-121.
- Mullins E, Milbourne D, Petti C, Doyle-Prestwich BM and Meade C (2006) Potato in the age of biotechnology. **Trends in Plant Science** **11**: 254-260.
- Page RDM (1996) TREEVIEW: an application to display phylogenetic trees on personal computers. **Computational Applied Bioscience** **12**: 357-358.
- Paterson AH (2006) Leafing through the genomes of our major crop plants: strategies for capturing unique information. **Nature Reviews Genetics** **7**: 174-184.

- Rong J, Bowers JE, Schulze SR, Waghmare VN, Rogers CJ, Pierce GJ, Zhang H, Estill JC and Paterson AH (2005) Comparative genomics of *Gossypium* and *Arabidopsis*: unraveling the consequences of both ancient and recent polyploidy. **Genome Research** **15**: 1198-1210.
- Ruane J and Sonnino A (2006) **The role of biotechnology in exploring and protecting agricultural genetic resources**. FAO, Rome, 187p.
- Salse J, Piegu B, Cooke R and Delseny M (2002) Synteny between *Arabidopsis thaliana* and rice at the genome level: a tool to identify conservation in the ongoing rice genome sequencing project. **Nucleic Acids Research** **30**: 2316-2328.
- Schneider H, Schuettpelz E, Pryer KM, Cranfill R, Magallon S and Lupia R (2004) Ferns diversified in the shadow of angiosperms. **Nature** **428**: 553-557.
- Senthil-Kumar M and Udayakumar M (2006) High-throughput virus-induced gene-silencing approach to assess the functional relevance of a moisture stress-induced cDNA homologous to *lea4*. **Journal of Experimental Botany** **57**: 2291-2302.
- Senthil-Kumar M, Govind G, Kang L, Mysore KS and Udayakumar M (2006) Functional characterization of *Nicotiana benthamiana* homologs of peanut water deficit-induced genes by virus-induced gene silencing. **Planta** [ahead of publication; DOI: 10.1007/s00425-006-0367-0].
- Song R, Llaca V and Messing J (2002) Mosaic organization of orthologous sequences in grass genomes. **Genome Research** **12**: 1549-1555.
- Sun HJ, Uchii S, Watanabe S and Ezura H (2006) A highly efficient transformation protocol for MicroTom, a model cultivar for tomato functional genomics. **Plant and Cell Physiology** **47**: 426-431.
- Tadege M, Ratet P and Mysore KS (2005) Insertional mutagenesis: a Swiss Army knife for functional genomics of *Medicago truncatula*. **Trends in Plant Science** **10**: 229-235.
- Tesfaye M, Silverstein KAT, Bucciarelli B, Samac DA and Vance CP (2006) The Affymetrix *Medicago* GeneChip® array is applicable for transcript analysis of alfalfa (*Medicago sativa*). **Functional Plant Biology** **33**: 783-788.
- Tonsor SJ, Alonso-Blanco C and Koornneef M (2005) Gene function beyond the single trait: natural variation, gene effects, and evolutionary ecology in *Arabidopsis thaliana*. **Plant Cell and Environment** **28**: 2-20.
- Town CD (2006) Annotating the genome of *Medicago truncatula*. **Current Opinion in Plant Biology** **9**: 122-127.
- Udvardi MK, Tabata S, Parniske M and Stougaard J (2005) *Lotus japonicus*: legume research in the fast lane. **Trends in Plant Science** **10**: 222-228.
- van Buuren ML, Salvi S, Morgante M, Serhani B and Tuberosa R (2002) Comparative genomic mapping between a 754 kb region flanking DREB1A in *Arabidopsis thaliana* and maize. **Plant Molecular Biology** **48**: 741-750.
- van Os H, Andrzejewski S, Bakker E, Barrena I, Bryan GJ, Caromel B, Ghareeb B, Isidore E, de Jong W, van Koert P, Lefebvre V, Milbourne D, Ritter E, van der Voort JN, Rousselle-Bourgeois F, van Vliet J, Waugh R, Visser RG, Bakker J and van Eck HJ (2006) Construction of a 10,000-marker ultradense genetic recombination map of potato: providing a framework for accelerated gene isolation and a genomewide physical map. **Genetics** **173**: 1075-1087.
- Varshney RK, Graner A and Sorrells ME (2005) Genomics-assisted breeding for crop improvement. **Trends in Plant Science** **10**: 621-630.
- Vij S, Gupta V, Kumar D, Vydianathan R, Raghuvanshi S, Khurana P, Khurana JP, Tyagi AK (2006) Decoding the rice genome. **Bioessays** **28**: 421-432.
- Wang S, Wang JW, Yu N, Li CH, Luo B, Gou JY, Wang LJ and Chen XY (2004) Control of plant trichome development by a cotton fiber MYB gene. **Plant Cell** **16**: 2323-2334.
- White PJ (2002) Recent advances in fruit development and ripening: an overview. **Journal of Experimental Botany** **53**: 1995-2000.
- Yang YW, Lai KN, Tai PY and Li WH (1999) Rates of nucleotide substitution in angiosperm mitochondrial DNA sequences and dates of divergence between *Brassica* and other angiosperm lineages. **Journal of Molecular Evolution** **48**: 597-604.
- Young ND, Cannon SB, Sato S, Kim D, Cook DR, Town CD, Roe BA and Tabata S (2005) Sequencing the genespaces of *Medicago truncatula* and *Lotus japonicus*. **Plant Physiology** **137**: 1174-1181.
- Yu J, Hu S, Wang J, Wong GK, Li S, Liu B, Deng Y, Dai L, Zhou Y, Zhang X, Cao M, Liu J, Sun J, Tang J, Chen Y, Huang X, Lin W, Ye C, Tong W, Cong L, Geng J, Han Y, Li L, Li W, Hu G, Huang X, Li W, Li J, Liu Z, Li L, Liu J, Qi Q, Liu J, Li L, Li T, Wang X, Lu H, Wu T, Zhu M, Ni P, Han H, Dong W, Ren X, Feng X, Cui P, Li X, Wang H, Xu X, Zhai W, Xu Z, Zhang J, He S, Zhang J, Xu J, Zhang K, Zheng X, Dong J, Zeng W, Tao L, Ye J, Tan J, Ren X, Chen X, He J, Liu D, Tian W, Tian C, Xia H, Bao Q, Li G, Gao H, Cao Y, Wang J, Zhao W, Li P, Chen W, Wang X, Zhang G, Xiong Y, Li Z, Mao L, Zhou C, Zhu Z, Chen R, Hao B, Zheng W, Chen S, Guo W, Li G, Liu S, Tao M, Wang J, Zhu L, Yuan L and Yang H (2002) A draft sequence of the rice genome (*Oryza sativa* L. ssp. *indica*). **Science** **296**: 79-92.
- Zhu H, Kim DJ, Back JM, Choi HK, Ellis LC, Küester H, McCombie WR, Peng HM and Cook DR (2003) Syntenic relationships between *Medicago truncatula* and *Arabidopsis* reveal extensive divergence of genome organization. **Plant Physiology** **131**: 1018-1026.