Phytochemicals in plants: genomics-assisted plant improvement for nutritional and health benefits
Michael A Grusak

Plants are an important source of essential nutrients and health-beneficial components that are crucial for human life. Because the intake of these phytochemicals is not always adequate, the resources of plant biotechnology are being used to enhance the nutritional quality of our plant-based food supply. Various improvement strategies are feasible, depending on whether the phytochemical target is a major or minor constituent. Recent efforts in gene discovery and functional genomics are providing the necessary understanding to develop and evaluate different approaches to manipulate phytochemical composition.

Addresses
USDA/ARS Children's Nutrition Research Center, Department of Pediatrics, Baylor College of Medicine, 1100 Bates Street, Houston, Texas 77030, USA; e-mail: mgrusak@bcm.tmc.edu

Introduction
Our plant-based food supply offers a diverse mixture of chemical constituents that are essential for human life and which can contribute to the promotion of good health. Because the density of nutritive and health-beneficial phytochemicals can vary dramatically amongst plant sources, with the levels of many compounds being quite low or non-existent in some plant foods, efforts are under way to develop and execute strategies that would improve the phytochemical composition of certain food crops [1]. Although conventional breeding is one means to achieve this goal, the genetic diversity available within existing germplasm collections for any given crop will limit the extent of phytochemical improvement that can be achieved through conventional techniques alone. Thus, transgenic approaches will be necessary to enable effective and significant increases in phytochemical content, especially in those cases where the compound of interest is introduced into a plant tissue in a novel way.

Efforts to genetically engineer improvements in complex, quality traits, such as phytochemical content, require a sound understanding of the biochemistry and nutritional physiology of both the phytochemical and the target plant in question [2]. This understanding is critical if one is to identify the key gene or genes whose products are responsible for the synthesis or acquisition of various phytochemicals, but is also necessary to assist in the functional optimization of any transgenic improvement strategy. Fortunately, the plant kingdom is a vast repository of genes for diverse biosynthetic and transport machinery, and certain plants are known for their ability to synthesize/accumulate high levels of a given phytochemical. Thus, researchers can draw upon the genomic resources of any of several agronomic, model or non-model species to identify the molecular components and regulatory elements that are pertinent to these important plant constituents [3].

The role of biotechnology in these improvement efforts is both essential and critical to the future of human society. Clearly, the inexcusably high prevalence of protein, energy, and/or micronutrient malnutritions that exist in various population groups throughout the world [4•] could be reduced by a more nutritious food supply, especially if efforts are focused on staple crops such as rice, maize, wheat, bean, and cassava [5,6]. However, even for presumably well-nourished populations it appears that not all essential nutrients are being obtained at recommended levels of intake [7] and strong evidence exists to support a health-beneficial role for various non-nutritive phytochemicals [8]. Additionally, many researchers have noted that transgenic efforts to improve the quality of our food supply might help to show that this technology can be used for consumer-oriented output traits (not just producer-oriented input traits). Such observations might serve to counteract some of the anti-biotechnological sentiment that exists in today's society [9]. In this article, we review some of the strategies, the constraints, and the genomic tools that can be used to effectively manipulate phytochemical levels in crop plants.

Phytochemical targets and strategies
Phytochemicals can be categorized broadly as major or minor constituents, depending on their general abundance in plants. Major constituents include carbohydrates, lipids, and proteins and are present at gram levels in each 100 g portion of food [10]. Minor constituents include vitamins, health-beneficial secondary metabolites, and minerals; when present, these constituents generally are found in the microgram to milligram range (per 100 g of food) [10]. Improvement strategies can be developed for both major and minor constituents, but the success and impact of the improvement will depend on whether a quantitative or qualitative engineering approach is taken. Significant quantitative changes are most feasible for minor constituents, where increases in the microgram to milligram range would require minimal diversion of precursors and only limited modifications in the plant's ability to store or sequester the target phytochemical. Examples of this include the engineering of β-carotene in rice grains (1.6 µg/g dry weight) [11], elevated iron content in rice (twofold increase) [12], enhanced ascorbic acid in lettuce

(sevenfold increase) [13], and increased flavonols in tomato (~60-fold increase) [14]. Although these changes are quantitatively small, they are nonetheless large changes in the functionality of the foods, because these phytochemicals provide nutritional and health benefits at low doses.

The engineering of qualitative changes is possible for all phytochemicals; however, strategies designed to divert metabolites from one end product to another are of particular interest for major constituents like lipids or carbohydrates. This type of strategy enables significant compositional changes in target phytochemicals without the need for concomitant modifications in the global productivity of the plant. Expression of a Δ6 desaturase from *Pythium irregulare* in *Brassica juncea*, for instance, led to the accumulation of γ-linolenic acid in seeds, at the expense of linoleic and α-linolenic acids [15]. Similarly, the expression of a fructan synthetic gene (1-sucrose:sucrose fructosyl transferase from *Helianthus tuberosus*) in sugar beet resulted in the nearly complete conversion of sucrose to fructan in the taproot with no change in total sugars [16].

Molecular issues and constraints

Complete and near-complete genomic sequences are now available for *Arabidopsis* and rice [17–19], and several sequencing efforts are generating expressed sequence tags and limited genomic sequences for various crop and model plant species. Nevertheless, we still know little with respect to the function of the thousands of gene products made by plants [20]. Thus, gene discovery and functional analysis of putative enzymes and transporters will continue to be critical issues in the development of transgenic improvement strategies. This is especially true for secondary metabolites, for which gene identification is still needed for numerous rate-limiting enzymes, or for critical branch points within complex pathways [21••]. Thankfully, the commonality of many metabolic pathways across species allows us to use sequence comparisons with known genes (or families of genes) from other organisms, to narrow our focus on small sets of candidate sequences in a species of interest. Functional genomic approaches can then be used to verify the putative metabolic pathway genes. Recent successes have utilized complementation of yeast mutants with *Arabidopsis* genes to identify critical enzymes in folate metabolism [22], heterologous expression of soybean genes in *Arabidopsis* to identify isoflavone synthase, a key enzyme in the synthesis of the phytoestrogens genistein and daidzein [23], and positional cloning in *Arabidopsis* and tomato mutants to identify a carotenoid isomerase enzyme that is essential for β-carotene biosynthesis [24,25].

For certain phytochemicals, the complexity of their biochemical pathway (e.g. secondary metabolites), or the complexity of their trafficking through the plant (e.g. minerals), may require the identification of several genes for use in a transgenic improvement strategy. This can be facilitated by the global analysis of expressed genes using DNA microarrays or gene chips [26] developed either for specific plants or for model species. The success of these analyses will depend on the nature of the genes that are on the array or chip (i.e. the extent of that species' genome) and the ingenuity of the researcher to select appropriate target tissues for mRNA extraction. Microarrays have been used to study developing *Arabidopsis* seeds [27,28] in an effort to identify seed-specific genes that are involved in the regulation of carbohydrate conversion into seed oil. The technology also has been used to investigate aspects of iron nutritional physiology, both in dicots and monocots [29•,30], as a means to identify genes whose coordinate expression has relevance to the ability of roots to acquire iron.

With multiple genes in hand, the problem still remains that multigene transfer is a difficult proposition [31], although some successes have been noted. Ye *et al.* [11] inserted three genes necessary for β-carotene synthesis into rice, in a single transformation effort. A vector containing a daffodil (*Narcissus pseudonarcissus*) phytoene synthase along with a bacterial (*Erwinia uredovora*) phytoene desaturase co-transformed with a vector containing a daffodil lycopene β-cyclase. As mentioned earlier, these investigators generated co-transformants that yielded 1.6 μg/g β-carotene in rice grains; untransformed rice would normally contain no β-carotene. Multigene combinations have also been attempted in an effort to improve iron content and bioavailability in rice [12]. Independent transformants were created using either a ferritin gene from bean (*Phaseolus vulgaris*) to increase iron storage capacity, a thermostable phytase gene from *Aspergillus fumigatus* or an endogenous rice gene for a cysteine-rich metallothionein-like protein; the latter two gene products had the intended purpose of enhancing iron bioavailability. The investigators used standard breeding approaches to cross the transgenic lines in an attempt to combine the various quality traits. The nutritional efficacy of these 'improved' iron lines, or of the β-carotene-enriched rice, have yet to be established.

In some cases, the difficulty of multiple gene insertions could be handled through an alternative approach. The successful overexpression of a transcription factor that coordinately regulates several genes in a pathway could simplify the transformation effort [21]. The transcription factor would need to be somewhat selective for the metabolic pathway of interest and, of course, the target plant would need to have the pathway genes available for activation in its genome. Such a strategy has proven successful in maize cell cultures, where induction of different components of the flavonoid pathway was accomplished by the ectopic expression of the transcription factors R/C1 or P [32], thereby causing cells to accumulate cyanidin derivatives or 3-deoxy flavonoids, respectively. For other pathways, the identification of appropriate regulatory genes should make it feasible to enhance the levels of a specific phytochemical or a group of phytochemicals [33].

Whole plant issues and constraints

Plants function within the framework of a limited pool of organic and inorganic resources that are partitioned spatially and temporally throughout the organism. Because various metabolic pathways have numerous points of intersection, via a shared pool of metabolites, any transgenic effort to influence the level of a specific phytochemical must take into account the consequences of non-targeted modifications to other compounds. For instance, one should consider whether an indirect loss of certain phytochemicals or an unintended increase in other phytochemicals would have a detrimental impact on the plant's growth and development. Alternatively, it is possible that an unanticipated loss or gain of certain non-target phytochemicals might lead to a reduction in the overall nutritional quality of the food. Efforts to increase Fe or Zn content in edible tissues, which might involve an induction of metal transporters (directly or indirectly), must be wary of simultaneously enhancing the levels of phytotoxic and/or health-detrimental heavy metals like cadmium or lead. Functional characterization of most membrane-localized metal transporters shows that they are not uniquely selective [34•,35], such that when other heavy metals are available in the soil environment they also can be accumulated in the plant. It seems prudent, therefore, that any nutritionally enhanced line should be screened for possible non-target (organic or inorganic) alterations, when grown under varying environmental conditions, before it is considered for commercialization.

For minerals, a sound understanding of the whole plant transport and partitioning dynamics of a target mineral is central to the design of any improvement strategy [36]. Unlike organic phytochemicals, which can be synthesized in discrete compartments or tissues when given the proper precursors and enzymatic machinery, there are of course no gene products that facilitate the synthesis of a mineral ion. Minerals must be absorbed through the root system, transported from roots to shoots via the xylem pathway, and in some cases remobilized to various sink tissues via the phloem system [1]. Mineral improvement strategies that are ultimately directed towards seeds may therefore require multiple transgenes to achieve a sizable enhancement in mineral concentration. Not only may source-sink partitioning need to be modified (e.g. phloem loading for leaf to seed delivery), but enhanced root uptake may also need to be engineered to increase the available pool of that mineral within the plant. Additionally, it should be noted that as most minerals are handled as ionized species, an increase in the synthesis of storage proteins or organic counter-ions may also need to be introduced. Information is becoming available on phloem-specific genes and promoter sequences [37] and new insights into various metal transporters continue to be found [38,39,40•], such that sensible improvement strategies for certain minerals can now be formulated.

Conclusions

There clearly is strong and concerted interest in the nutritional enhancement of our plant-based food supply, as evidenced by the breadth of research being conducted with all classes of phytochemicals. Continued efforts in gene discovery, functional genomics, secondary metabolism, and nutritional physiology must be pursued, as these investigations will provide the necessary molecular insights and whole plant understanding to help us formulate sensible, high-impact improvement approaches. The use of model systems to assess the efficacy of various transgenic strategies will remain an important proof-of-concept tool, but researchers must continue to look towards the transformation of agronomically important crops as soon as an effective strategy is identified. We must push forward with biotechnological efforts now, because improved, stable lines in different genetic backgrounds will not be accomplished quickly, and the growth of the human population will go on (for better or worse) whether we address these agricultural goals or not.

References and recommended reading
Papers of particular interest, published within the annual period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Grusak MA, DellaPenna D: **Improving the nutrient composition of plants to enhance human nutrition and health.** *Annu Rev Plant Physiol Plant Mol Biol* 1999, **50**:133-161.

2. DellaPenna D: **Plant metabolic engineering.** *Plant Physiol* 2001, **125**:160-163.

3. DellaPenna D: **Nutritional genomics: manipulating plant micronutrients to improve human health.** *Science* 1999, **285**:375-379.

4. Food and Agriculture Organization of the United Nations: *The State of*
•   *Food Insecurity in the World 2000.* Rome:FAO; 2000; p 31.
A forceful and comprehensive document that outlines the severity of the world's food situation and the need for dramatic change in our food resources.

5. Conway G, Toenniessen G: **Feeding the world in the twenty-first century.** *Nature* 1999, **402**:C55-C58.

6. Bouis HE, Graham RD, Welch RM: **The consultative group on international agricultural research (CGIAR) micronutrients project: justification and objectives.** *Food Nutr Bull* 2000, **21**:374-381.

7. Krebs-Smith SM, Cleveland LE, Ballard-Barbash R, Cook DA, Kahle LL: **Characterizing food intake patterns of American adults.** *Am J Clin Nutr* 1997, **65**:1264S-1268S.

8. World Cancer Research Fund/American Institute for Cancer Research: **Food, nutrition and the prevention of cancer: a global perspective.** Washington DC: American Institute for Cancer Research; 1997.

9. Kleter GA, van der Krieken WM, Kok EJ, Bosch D, Jordi W, Gilissen LJWJ: **Regulation and exploitation of genetically modified crops.** *Nat Biotechnol* 2001, **19**:1105-1110.

10. US Department of Agriculture: Agricultural Research Service: **USDA Nutrient Database for Standard Reference, Release 14.** Nutrient Data Laboratory Home Page; 2001. URL: http//www.nal.usda.gov/fnic/foodcomp

11. Ye X, Al-Babili S, Klöti A, Zhang J, Lucca P, Beyer P, Potrykus I: **Engineering the provitamin A (β-carotene) biosynthetic pathway into (carotenoid-free) rice endosperm.** *Science* 2000, **287**:303-305.

12. Lucca P, Hurrell R, Potrykus I: **Genetic engineering approaches to improve the bioavailability and the level of iron in rice grains.** *Theor Appl Genet* 2001, **102**:392-397.

13. Jain AK, Nessler CL: **Metabolic engineering of an alternative pathway for ascorbic acid biosynthesis in plants**. *Mol Breed* 2000, **6**:73-78.

14. Muir SR, Collins GJ, Robinson S, Hughes S, Bovy A, De Vos RCH, van Tunen AJ, Verhoeyen ME: **Overexpression of petunia chalcone isomerase in tomato results in fruit containing increased levels of flavonols**. *Nat Biotechnol* 2001, **19**:470-474.

15. Hong H, Datla N, Reed DW, Covello PS, MacKenzie SL, Qiu X: **High-level production of γ-linolenic acid in *Brassica juncea* using a Δ6 desaturase from *Pythium irregulare***. *Plant Physiol* 2002, **129**:354-362.

16. Sévenier R, Hall RD, van der Meer IM, Hakkert HJC, van Tunen AJ, Koops AJ: **High level fructan accumulation in a transgenic sugar beet**. *Nat Biotechnol* 1998, **16**:843-846.

17. The *Arabidopsis* Genome Initiative: **Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana***. *Nature* 2000, **408**:796-815.

18. Yu J, Hu S, Wang J, Wong GK-S, Li S, Liu B, Deng Y, Dai L, Zhou Y, Zhang X *et al.*: **A draft sequence of the rice genome (*Oryza sativa* L. ssp. *indica*)**. *Science* 2002, **296**:79-92.

19. Goff SA, Ricke D, Lan T-H, Presting G, Wang R, Dunn M, Glazebrook J, Sessions A, Oeller P, Varma H *et al.*: **A draft sequence of the rice genome (*Oryza sativa* L. ssp. *japonica*)**. *Science* 2002, **296**:92-100.

20. Meagher RB: **Plant biotechnology: post-genomics networking of biotechnology for interpreting gene function**. *Curr Opin Plant Biol* 2002, **5**:135-140.

21. Verpoorte R, van der Heijden R, Memelink J: **Engineering the plant
••   cell factory for secondary metabolite production**. *Transgenic Res* 2000, **9**:323-343.
An excellent review of the approaches and tools available for the manipulation of secondary metabolites in plants.

22. Ravanel S, Cherest H, Jabrin S, Grunwald D, Surdin-Kerjan Y, Douce R, Rébeillé F: **Tetrahydrofolate biosynthesis in plants: molecular and functional characterization of dihydrofolate synthetase and three isoforms of folylpolyglutamate synthetase in *Arabidopsis thaliana***. *Proc Natl Acad Sci USA* 2001, **98**:15360-15365.

23. Jung W, Yu O, Lau S-MC, O'Keefe DP, Odell J, Fader G, McGonigle B: **Identification and expression of isoflavone synthase, the key enzyme for biosynthesis of isoflavones in legumes**. *Nat Biotechnol* 2000, **18**:208-212.

24. Park H, Kreunen SS, Cuttriss AJ, DellaPenna D, Pogson BJ: **Identification of the carotenoid isomerase provides insight into carotenoid biosynthesis, prolamellar body formation, and photomorphogenesis**. *Plant Cell* 2002, **14**:321-332.

25. Isaacson T, Ronen G, Zamir D, Hirschberg J: **Cloning of *tangerine* from tomato reveals a carotenoid isomerase essential for the production of β-carotene and xanthophylls in plants**. *Plant Cell* 2002, **14**:333-342.

26. Wisman E, Ohlrogge J: ***Arabidopsis* microarray service facilities**. *Plant Physiol* 2000, **124**:1468-1471.

27. Girke T, Todd J, Ruuska S, White J, Benning C, Ohlrogge J: **Microarray analysis of developing *Arabidopsis* seeds**. *Plant Physiol* 2000, **124**:1570-1581.

28. White JA, Todd J, Newman T, Focks N, Girke T, de Ilárduya OM, Jaworski JG, Ohlrogge JB, Benning C: **A new set of *Arabidopsis* expressed sequence tags from developing seeds. The metabolic pathway from carbohydrates to seed oil**. *Plant Physiol* 2000, **124**:1582-1594.

29. Thimm O, Essigmann B, Kloska S, Altmann T, Buckhout TJ: **Response
•   of *Arabidopsis* to iron deficiency stress as revealed by microarray analysis**. *Plant Physiol* 2001, **127**:1030-1043.
A good example of the use of microarray technology to characterize and dissect patterns of gene expression in multiple biochemical pathways.

30. Negishi T, Nakanishi H, Yazaki J, Kishimoto N, Fujii F, Shimbo K, Yamamoto K, Sakata K, Sasaki T, Kikuchi S *et al.*: **cDNA microarray analysis of gene expression during Fe-deficiency stress in barley suggests that polar transport of vesicles is implicated in phytosiderophore secretion in Fe-deficient barley roots**. *Plant J* 2002, **30**:83-94.

31. Daniell H, Dhingra A: **Multigene engineering: dawn of an exciting new era in biotechnology**. *Curr Opin Biotechnol* 2002, **13**:136-141.

32. Grotewold E, Chamberlin M, Snook M, Siame B, Butler L, Swenson J, Maddock S, St. Clair G, Bowen B: **Engineering secondary metabolism in maize cells by ectopic expression of transcription factors**. *Plant Cell* 1998, **10**:721-740.

33. Verpoorte R, Memelink J: **Engineering secondary metabolite production in plants**. *Curr Opin Biotechnol* 2002, **13**:181-187.

34. Williams LE, Pittman JK, Hall JL: **Emerging mechanisms for heavy
•   metal transport in plants**. *Biochim Biophys Acta* 2000, **1465**:104-126.
A good review of structure/function aspects of the various types of metal transporters found in plants.

35. Rogers EE, Eide DJ, Guerinot ML: **Altered activity in an *Arabidopsis* metal transporter**. *Proc Natl Acad Sci USA* 2000, **97**:12356-12360.

36. Grusak MA: **Enhancing mineral content in plant food products**. *J Am Coll Nutr* 2002, **21**, in press.

37. Nishitani C, Demura T, Fukuda H: **Primary phloem-specific expression of a *Zinnia elegans* homeobox gene**. *Plant Cell Physiol* 2001, **42**:1210-1218.

38. Shigaki T, Cheng N-H, Pittman JK, Hirschi K: **Structural determinants of Ca^{2+} transport in the *Arabidopsis* H^+/Ca^{2+} antiporter *CAX1***. *J Biol Chem* 2001, **276**:43152-43159.

39. Li L, Tutone AF, Drummond RSM, Gardner RC, Luan S: **A novel family of magnesium transport genes in *Arabidopsis***. *Plant Cell* 2001, **13**:2761-2775.

40. Hamburger D, Rezzonico E, Petétot JM-C, Somerville C, Poirier Y:
•   **Identification and characterization of the *Arabidopsis* *PHO1* gene involved in phosphate loading to the xylem**. *Plant Cell* 2002, **14**:889-902.
A good example of the type of research that is needed for other minerals to identify individual components in the whole plant mineral puzzle.