Modelling gene networks controlling transition to flowering in Arabidopsis

Stephen M. Welch[1], Zhanshan Dong[2], Judith L. Roe[3]

[1]Department of Agronomy, Kansas State University, Manhattan, KS USA 66506. Email: welchsm@ksu.edu
[2]Pioneer Hi-Bred International, Inc., 7300 NW 62nd Ave., Johnston, IA USA 50131. Email: zhanshan.dong@pioneer.com
[3]Division of Biology, Kansas State University, Manhattan, KS USA 66506. Email: jroe@ksu.edu

Abstract
Flowering is a critical stage in plant development that initiates grain production and is vulnerable to stress. The genes controlling flowering time in the model plant Arabidopsis thaliana are reviewed. Previously, interactions between these genes were described by qualitative network diagrams. We present a generalized mathematical formalism that relates environmentally dependent transcription, RNA processing, translation, and protein-protein interaction rates to resultant phenotypes. We have developed models (reported elsewhere) based on this concept that simulate flowering times for novel A. thaliana genotype-environment combinations and critical short day lengths (CSDL) in rice (Oryza sativa ssp. japonica cv. Nipponbare). Here we show how CSDL phenotypes emerge from gene expression dynamics. Functionally different but homologous photoperiod measurement genes in rice and A. thaliana nevertheless yield similar results. Other technologies for interrelating genotypes, phenotypes, and the environment are crop simulation models and the theory of quantitative genetics (QG). Some potential synergies between genetic networking (GN) and these older approaches are discussed. Twelve contrasts are drawn between QG and GN revealing that both have equal contributions to make to an ideal theory. Such a theory is initiated by discussing epistasis, dominance, and additivity (all QG basics) in GN terms. Three or less genes can account for the first two but additivity is a complex property dependent on the structure and function of entire subnets. Finally, the utility of simple models is evidenced by 80 years of quantitative genetics and mathematical ecology.

Media summary
Mathematical studies of plants whose genomes have been deciphered are leading to new insights regarding theories that have underpinned crop breeding efforts for many decades.

Keywords
Regulation, differential equations, photothermal, pathways

Introduction
An organism's genome is a functional control system. The surrounding environment provides external inputs to which that control system responds. The result is a time sequence of states whose observable attributes are phenotypes. According to Cooper *et al.* (2002), modelling the relationship between genotypes and the resulting phenotypes in particular environments is a major problem in computational biology. They refer to this as the "GP problem". Three technologies for predicting phenotypes are (in order of age) quantitative genetics, physiological crop simulation modelling, and genetic network theory. The first and last approaches, in particular, are exploiting the recent explosive advances in genomic science.

Multiple mathematical formalisms have been used to model genetic and, more generally, metabolic networks. Examples include (*i*) Boolean (ON/OFF) networks (Frank, 1998; Liang *et al.*, 1998; Mendoza and Alvarez-Buylla, 1998; Szallasi and Liang, 1998; Samsonova and Serov, 1999; Akutsu *et al.*, 2000; Mendoza and Alvarez-Buylla, 2000; Ideker *et al.*, 2000; Maki *et al.*, 2001), (*ii*) Petri (concurrent information flow) nets (Goss and Peccoud, 1999; Matsuno *et al.* 2000), (*iii*) S-systems (continuous time models motivated by chemical kinetics) (Liang *et al.*, 1998; Akutsu *et al.*, 1999; Tominaga *et al.*, 1999; Akutsu *et al.*, 2000; Maki *et al.*, 2001), (*iv*) differential equation models (Chen *et al.*, 1999; Wolf and Eeckman, 1998), (*v*) neural network models (Reinitz and Sharp, 1995; D'Haesseleer *et al.*, 1999; Weaver *et al.*, 1999; Marnellos *et al.,* 2000), and (*vi*) Bayesian networks (Friedman *et al.*, 2000; Barash and Friedman, 2001; Hartemink *et al.*, 2001). Despite this extensive effort, little attention has focused on predicting phenotypes of interest to crop scientists or on integrating the effects of multiple environmental factors.

In response, we have modeled the genetic network control of floral initiation in *Arabidopsis thaliana*. Our choice of this test system was motivated by two factors. Intensive research during the 1990's elucidated the structure of the control network. Also, many skillful empirical flowering time models exist, suggesting that this process might be "within reach". This rationale has proved felicitous. The resulting model not only reproduces its calibration data, but also simulates (from first principles) the inflorescence bud dates for many mutant and engineered genotypes that differ widely from those used for model calibration. This work is described in Dong (2003) and several papers in preparation. Here we (1) review genetic control of the *A. thaliana* floral transition, (2) present a general gene network model, (3) show that key physiological traits (exemplified by critical short day length) can arise as emergent properties of gene networks, and (4) discuss synergies between crop simulation, quantitative genetics, and gene network models, especially the latter two.

Gene regulatory network controlling flowering in *Arabidopsis*

The transition to flowering is influenced by both endogenous and exogenous signals. The underlying genetic regulatory network that integrates and transduces these signal has been elucidated in *Arabidopsis* using two complementary strategies; comparisons between naturally occurring ecotypes, and the genetic analysis of mutations that result in early or late flowering phenotypes (Levy and Dean, 1998). Genes have been positioned in several pathways that promote or repress flowering, depending on environmental or autonomous conditions, and how these pathways interact is an area of active study. A number of network models have been constructed and refined for flowering time control (Martinez-Zapater *et al.*, 1994; Haughn *et al.*, 1995; Blazquez, 2000) and an integrated current view is presented in Figure 1. Four major pathways (photoperiod, autonomous, vernalization, and gibberellin) converge on the meristem identity genes *LEAFY* (*LFY*) and *APETALA1* (*AP1*). A few key genes (boxed in Figure 1) integrate signals from the different converging pathways. Under appropriate conditions, then, their activities lead to increased *LFY* and *AP1* expression, which then mediate the switch to reproductive development in the shoot meristem.
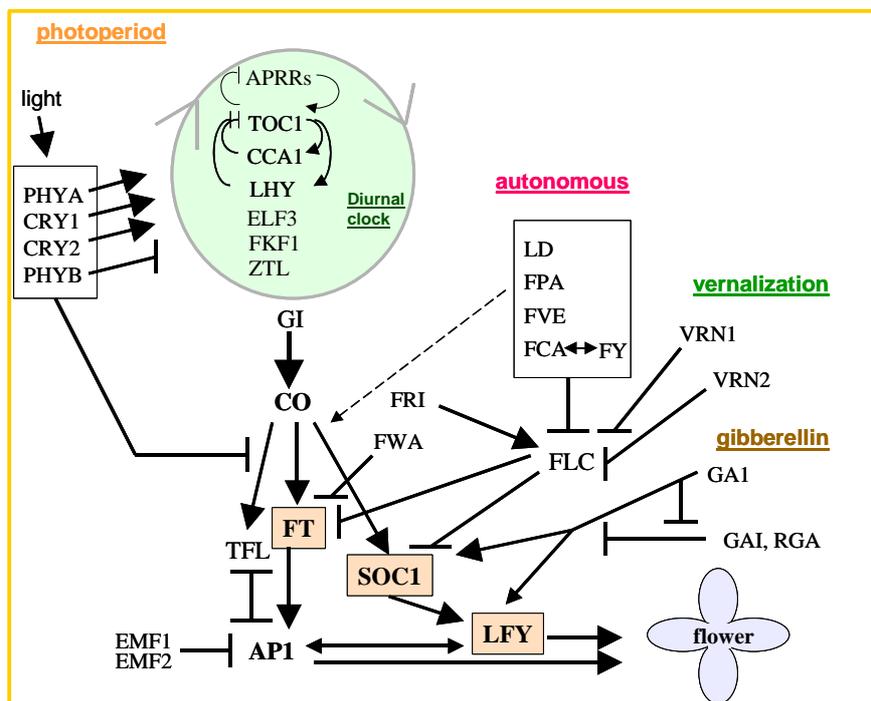

Figure 1. The genetic regulatory network controlling flowering time in *Arabidopsis*. The positive (arrows) and negative (bars) regulatory relationships are from reviews cited in the text. Genes in tinted boxes (*FT*, *LFY* and *SOC1*) integrate signals from multiple pathways. The dashed line depicts a possible connection between the autonomous pathway and the photoperiod pathway that is independent of *FLC*.

Photoperiod pathway
Arabidopsis is a facultative or quantitative long day (LD) plant that can flower, albeit much later, in short days (SD). Key regulatory genes appear to be conserved between *Arabidopsis* and rice, a SD plant (Blazquez *et al.*, 2001; Samach and Gover, 2001; Yano *et al.*, 2001; Goff et al., 2002; Shimamoto and Kyozuka, 2002; Mouradov *et al.*, 2002), suggesting that common pathways are utilized. Photoperiod is

perceived by the plant and transduced to a downstream signalling system by the interaction of photoperception mechanisms with the endogenous diurnal clock (reviewed in Hayama and Coupland, 2003). The light- and clock-regulated expression of the flowering time gene *CONSTANS* (*CO*) (*Hd1* in rice) is critical to the timing of flowering.

(a) Photoperception
The two main classes of photoreceptors in *Arabidopsis* are the phytochromes and the cryptochromes that are involved in sensing red/far red, and blue light and ultraviolet wavelengths, respectively (reviewed in Hudson, 2000; Smith, 2000; Lin, 2000; Devlin, 2002). The products of the five phytochrome genes, *PHYTOCHROME A (PHYA)* through *PHYTOCHROME E (PHYE)*, and the two cryptochrome genes, *CRYPTOCHROME 1 (CRY1)* and *CRY2* play critical roles in sensing light and entraining the circadian clock. Both *PHYA* and *CRY2* can promote flowering (Guo *et al.*, 1998; Mockler *et al.*, 1999), and mutations in these genes cause a late-flowering phenotype (Johnson *et al.*, 1994). On the other hand, *phyB* mutants flower earlier than wild type (Goto *et al.*, 1991; Reed *et al.*, 1993; Bagnall *et al.*, 1995), implying that *PHYB* represses flowering (Mockler *et al.*, 1999).

(b) Endogenous circadian clock
The details of the *Arabidopsis* circadian clock are reviewed elsewhere (Millar, 1999; Somers, 1999; Samach and Coupland, 2000; McClung, 2001; Johnson, 2001; Alabadi *et al.*, 2001; Devlin, 2002; Eriksson and Millar, 2003). The central oscillator houses a negative feedback loop whereby *TIMING OF CAB 1 (TOC1)* stimulates expression of *LATE ELONGATED HYPOCOTYL (LHY)* and *CIRCADIAN CLOCK ASSOCIATED 1 (CCA1)*, which then feed back and repress TOC1 expression (Alabadi *et al.*, 2001)(Fig. 1). *LHY* and *CCA1* encode similar single-MYB/SANT domain proteins (Wang *et al.*, 1997; Schaffer *et al.*, 1998; Wang and Tobin, 1998; Carre and Kim, 2002) and *TOC1/ARABIDOPSIS PSEUDO RESPONSE REGULATOR 1 (APRR1)* encodes a pseudo-response regulator (Makino *et al.*, 2000; Strayer *et al.*, 2000) that has a carboxyl-terminal CCT motif (<u>CO</u>, <u>CO-L</u>, <u>TOC1</u>) (Robson *et al.*, 2001). *TOC1* is a member of an *APRR* gene family (*TOC1/APPR1, APRR3, APRR5, APRR7*, and *APRR9*) whose staggered wave pattern of circadian expression may indicate a negative feedback loop and/or an oscillator (Matsushika *et al.*, 2000, 2002a, 2002b; Makino *et al.*, 2000, 2001, 2002; Murakami-Kojima *et al.*, 2002). Eriksson and Millar (2003) depict the circadian system in *Arabidopsis* as the two abovementioned loops linked by *TOC1*. Other genes have been described that also may function in clock control, such as *EARLY FLOWERING 3* (*ELF3*) (Hicks *et al.*, 2001; Liu *et al.*, 2001), and *4* (*ELF4*) (Doyle *et al.*, 2002), *ZEITLUPE* (*ZTL*) (Somers *et al.*, 2000; Jarillo *et al.*, 2001) and *FKF1* (flavin-binding, kelch repeat, F-box) (Nelson *et al.*, 2000).

(c) Photoperiod measurement, transduction, and integration of pathways
Photoperiod is measured by an interaction of light with the circadian clock and the output then traverses a signalling pathway leading directly to the meristem identity genes. Several important flowering time genes operate directly downstream of the circadian clock in the light-signaling pathway. *GIGANTEA* (*GI*) encodes a nuclear protein involved in phytochrome signaling whose expression is regulated by the circadian clock (Fowler *et al.*, 1999; Park *et al.*, 1999; Huq *et al.*, 2000). Its relationship with the circadian clock is complex, as it in turn can regulate some components of the circadian clock (Fowler *et al.*, 1999; Park *et al.*, 1999).

CO, a zinc-finger transcription factor (Putterill *et al.*, 1995), is tightly regulated by the circadian clock and is critical for photoperiod perception (Yanovsky and Kay, 2002). *CO* accelerates flowering in long days by promoting the expression of the downstream genes *SOC1* and *FT* (Suarez-Lopez *et al.*, 2001; Simpson and Dean, 2002). Roden *et al.* (2002) altered the timing of circadian rhythms of gene expression relative to dawn and dusk and found that a cycle was perceived as a long day condition when elevated *CO* expression coincided with daylight, consistent with the external coincidence model of photoperiodism (Samach and Coupland, 2000; Carre, 2001; Davis, 2002). This model says that only light that is coincident with a certain phase of the diurnal clock will promote flowering (Bunning, 1936; Pittendrigh, 1972).

SOC1 encodes a MADS-box transcription factor (Borner *et al.*, 2000; Lee *et al.*, 2000), while *FT* produces a small kinase inhibitor-like protein (Kardailsky *et al.*, 1999; Kobayashi *et al.*, 1999). These both integrate signals from the photoperiod pathway as described above (Kobayashi *et al.*, 1999;

Kardailsky *et al.*, 1999; Samach *et al.*, 2000; Suarez-Lopez *et al.*, 2001; Hepworth *et al.*, 2002; Blazquez *et al.*, 2002), and the autonomous pathway described below (Simpson *et al.*, 1999; Sheldon *et al.*, 2000; Rouse *et al.*, 2002), along with physiological age (Samach *et al.*, 2000) and gibberellin (Blazquez and Weigel, 1999), in the case of *SOC1* (see below). *SOC1* and *FT* promote expression of the meristem identity genes *LFY* and *AP1*, respectively, which then cause the transition to inflorescence development in the shoot meristem (reviewed in Reeves and Coupland, 2000; Mouradov *et al.*, 2002; Sung *et al.*, 2003).

Autonomous pathway
Genes in the autonomous pathway include *FCA*, *FY*, *FVE*, *FPA*, and *LUMINIDEPENDENS (LD)* (Koornneef *et al.*, 1991). Mutations in these genes cause a late flowering phenotype in long or short photperiods, but this is suppressed by a vernalization treatment (Martinez-Zapater and Somerville, 1990; Koornneef *et al.*, 1991). The genes independently promote flowering by downregulating expression of *FLC* , a floral repressor (Sheldon *et al.*, 1999, 2000), as does vernalization (below).

FCA encodes a protein with two RNA recognition motifs (RRM domains) (Burd and Dreyfuss, 1994), a WW protein interaction domain (Macknight *et al.*, 1997), and a domain interacting with AtSWI3B (Sarnowski *et al.*, 2002). Processing of *FCA* transcripts is complex (Macknight *et al.*, 1997; Macknight *et al.*, 2002) and shows negative autoregulation (Quesada *et al.*, 2003). *FY* is a conserved mRNA 3' end processing factor that functions with *FCA* (Simpson *et al*., 2003). *FPA* also encodes an RNA-binding protein (Schomburg *et al.*, 2001), whereas LD encodes a homeodomain protein (Lee *et al.*, 1994)

FVE encodes a WD-40 repeat protein similar to human RbAp48, AtMSI4, possibly involved in chromatin regulation (Kenzior and Folk, 1998; Morel *et al.*, 2002). *FVE* protein may not be simply involved in flowering time control but also during all stages of plant development (Martínez-Zapater *et al.*, 1995). *fve* and *fca* mutants are less sensitive to growth temperature than either wild-type or other mutants in the range of 24°C to 16°C (Blazquez *et al* , 2003; Welch *at al*,. 2003; Dong, 2003), suggesting that *FCA* and *FVE* function are especially responsive to temperature which may contribute to differential temperature-dependent growth rates, leaf development rates, and flowering time.

Vernalization pathway
FLOWERING LOCUS C (*FLC*) is a floral repressor gene encoding a MADS-box transcription factor (Michaels and Amasino, 1999; Sheldon *et al.*, 1999) that is regulated by both the autonomous and the vernalization pathways. Elevated expression of *FLC* leads to suppression of *FT* and *SOC1* and inhibition of flowering (Hepworth *et al.*, 2002). Cold treatment leads to down-regulation of *FLC* in genotypes that require vernalization (1 to 3 months under 4°C), protecting the plant from flowering before spring (Reeves and Coupland, 2000). Mutations decreasing *FLC* function have created naturally occurring vernalization-independent ecotypes (summer annuals) (Michaels *et al.* 2003). *FRIGIDA (FRI)*, a second major determinant of natural variation in *Arabidopsis* flowering time, enhances *FLC* function (Michaels and Amasino, 1999; Johanson *et al.*, 2000).

The vernalization pathway senses low temperatures and downregulates *FLC* mRNA levels by an unknown mechanism. A set of vernalization genes (*VRN*) has been identified in screens for mutants unable to respond to vernalization in a late-flowering mutant background (Chandler *et al.*, 1996). Two of these, *VRN1* and *VRN2*, are involved in maintenance of *FLC* repression; *VRN1* encodes a DNA-binding protein and *VRN2* encodes a nuclear-localized zinc-finger protein (Gendall *et al.*, 2001; Levy *et al.*, 2002).

Gibberellin pathway
The plant hormone, gibberellin (GA), influences flowering time, and is the main promoting pathway for flowering in short days in Arabidopsis (Reeves and Coupland, 2001). It plays only a minor role during LD, when other pathways are dominant. Several major genes that have been identified in the pathway are repressors of the GA response, including *GIBBERELLIN INSENSITIVE (GAI)* and *REPRESSOR OF GA (RGA),* which presumably must be inhibited during GA signalling (Peng *et al*., 1997; Silverstone *et al.* 1997, 1998). They are both members of the GRAS family of plant transcription factors (Lee *et al*., 2002). The gibberellin pathway stimulates flowering by causing up-regulation of *SOC1* and *LFY*, but not *FT* (Blazquez *et al*., 1998; Moon *et al*., 2003).

Meristem identity genes
Promotion of flowering from the flowering time genes finally stimulates expression of the floral meristem identity genes *LFY* and *AP1*, as the terminal output from the converging signal pathways (Hempel *et al.*, 1997). The floral meristem identity genes then control floral organ identity genes that pattern development of the floral organs (Simpson *et al.*, 1999). *LFY* encodes a nuclear DNA-binding protein that activates floral homeotic gene expression (Weigel *et al.*, 1992; Busch *et al.*, 1999). Overexpression of *LFY* causes early flowering (Weigel and Nilsson, 1995), whereas mutations in the *LFY* gene result in the conversion of early flowers into shoot-like structures (Weigel *et al.*, 1992). *AP1* encodes a MADS-box transcription factor that also regulates floral homeotic genes (Mandel *et al.*, 1992; Ng and Yanofsky, 2001; Lamb *et al.* 2002). In addition, another repressive factor that blocks meristem conversion during vegetative development by repressing *AP1* and *LFY*, *TERMINAL FLOWER 1* (*TFL1*), is mutually inhibited by both of these factors once they are activated (Liljegren *et al.*, 1999).

LFY is activated by *SOC1* (Mouradov *et al.*, 2002), possibly via an intermediary factor *AGAMOUS-LIKE 24 (AGL24)* (Yu *et al.*, 2002) and is up-regulated by gibberellin either directly or via *SOC1* as the activator (Blazquez and Weigel, 1999; Moon *et al.*, 2003), whereas *AP1* is activated by *FT* (Ruiz-Garcia *et al.*, 1997). *AP1* and *LFY* have distinct but overlapping functions and positively regulated each other (Bowman *et al.*, 1993; Wagner *et al.*, 1999; Liljegren *et al.*, 1999). While the primary activation of *LFY* and *AP1* occurs by parallel pathways, subsequent reciprocal activation amplifies floral meristem identify gene expression (Liljegren *et al.*, 1999; Mouradov *et al.*, 2002) suggesting a bi-stable switch function (Welch *et al.*, 2005).

Integration of Pathways
As mentioned above, signals from the four pathways converge at downstream points, creating a network of both environmental inputs and growth signals to influence flowering time quantitatively. The plant must integrate all of the signals and flower under appropriate conditions and developmental stage. The genes *FT*, *SOC1*, and *LFY* are the main integration points; all three respond to signals from the autonomous and vernalization pathways, *SOC1* and *LFY* also integrate GA signals, and *SOC1* and *FT* are regulated by the photoperiod pathway (Mouradov *et al.*, 2002; Halliday *et al.*, 2003; Moon *et al.*, 2003).

Other Floral Repressors and regulation of flowering time by microRNAs
In addition to the floral repressor *FLC*, other genes have been identified that appear to negatively regulate flowering, including the *EMBRYONIC FLOWERING* (*EMF*) genes, *EARLY BOLTING IN SHORT DAYS* (*EBS*) and *FWA* (Reeves and Coupland, 2000; Chou *et al.*, 2001; Sung *et al.*, 2003). Recently reported gene expression profiles of induced shoot apices reveal not only induced genes, but additional down-regulated genes that may function as floral repressors (Schmid *et al.*, 2003). In addition to traditional scenarios of gene repression, microRNAs are the products of a new class of genes that can inhibit the function of other genes at a post-transcriptional step (Carrington and Ambros, 2003). Recently, the *miR172* gene has been described that can target a family of AP2-domain containing genes, including some that may function as floral repressors (Park *et al.*, 2002; Auckerman and Sakai, 2003; Schmid *et al.*, 2003). So, the plant must coordinate the repression and activation of a diverse set of genes to achieve the switch to flowering.

General theory
Despite the complexities just discussed, a generalized system of differential equations can model molecular and environmental interactions within a single organism. State variables are RNA's and proteins although other metabolites could easily be added. Let **s** be a ($m + p$) column vector partitioned into **m** and **p**, which contain levels of RNA expression and protein, respectively. The rate of change of any s_i is the environmentally dependent difference between its production and degradation rates

$$\frac{ds_i}{dt} = R_i\, g_i\left(\mathbf{s},\mathbf{e},\mathbf{a_i}\right) - \lambda_i\, h_i\left(\mathbf{s},\mathbf{e},\mathbf{b_i}\right) s_i \tag{1}$$

where $\mathbf{a_i}$ and $\mathbf{b_i}$ are parameter vectors, $\mathbf{e}$ is a vector of environmental inputs, R_i, and λ_i are scalars, $s_i(0)=0$, and $0 \leq g, h \leq 1$. The $\mathbf{a_i}$, $\mathbf{b_i}$, R_i, λ_i parameters are all positive. The s_i are dimensionless due to normalization against standards during measurement (*e.g.*, Northern or Western blotting, microarray and rt-PCR technology). The parameter values offset the use of different standards for different i's. Upper bounds on g and h (not necessarily reached) are justified because all biological processes have finite

rates. The initial condition reflects the negligible size of a plant at $t=0$. Clearly, for some i, $g_i(s(0),\mathbf{e}(0),\mathbf{a_i}) \neq 0$.

Equation (1) allows RNA and protein to influence each other in all combinations as in Table 1.

Table 1. Processes involving RNA's and proteins.

Influences (row to col.)	RNA	Protein
RNA	Alternative splicing	Translation
Protein	Transcription control	Protein-protein interactions

Equation (1) can be modified to incorporate diploidy. Assume that the total RNA produced at each locus is the sum of that produced by each allele. We represent a genotype by the $1 \times (m+p)$ column vector

$$\mathbf{X} = \begin{bmatrix} \alpha_{1j_11} + \alpha_{1j_12} \\ M \\ \alpha_{mj_{m1}} + \alpha_{mj_{m2}2} \\ 1 \\ M \\ 1 \end{bmatrix} \tag{2}$$

where the last p entries are 1. α_{ijk} is the *effective* relative RNA production rate of allele j at locus i, where i runs sequentially across the haploid genome. The $k = 1,2$ subscript indexes the chromosomes in each homologous pair. The α_{ijk} are measured relative to single copy wild type alleles. If a mutation changes a wild type allele at locus i to a null allele then α_{ijk} changes from one to zero. A mutation that increases the copy number, increases α_{ijk} proportionately. A partial-loss-of-function mutant allele has an intermediate α_{ijk}. This model assumes that mutation alters regulation, product activity, or both proportionately across all situations. Letting $*$ denote component-wise multiplication, the model becomes

$$\mathbf{\dot{s}} = \mathbf{X} * \mathbf{R} * \mathbf{g} - \Lambda * \mathbf{h} * \mathbf{s} \tag{3}$$
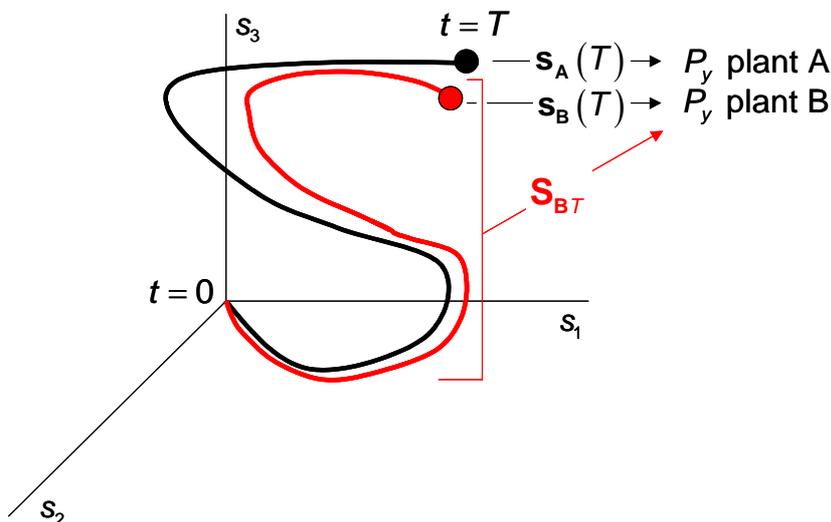

Figure 2. Biochemical trajectories for two plants (A and B). Because a plant's phenotype (P_y) depends on it's developmental history, phenotype can be related to the entire curve ($S_{.T}$) as shown in red for B.

Equation (3) provides a complete biochemical model of a plant but says nothing about phenotypic traits like height, flowering date, *etc.* A plausible assumption is that two plants are never biochemically identical except (trivially) at $t=0$. Therefore, at observation time T, the biochemical state, $s_A(T)$, of plant is unique. It would be simple to assume that the phenotype is a function of $s_A(T)$ but many traits (*e.g.*, yield at harvest) depend on earlier events (*e.g.*, water availability at silking). It is therefore convenient to relate the value of a trait to the entire time series of $\mathbf{s}$ from $t=0$ to T. This is illustrated in Figure 2 for $(m+p)=3$ with the time series of $\mathbf{s}$ symbolized by $\mathbf{S}$. This notation exhibits massive data compression.

For example, let equation (3) be the "rough network model" in Chen *et al*. (1999) and Baldi and Hatfield (2002, p. 151). If repeated expression level snapshots are taken with whole-genome chips, the single symbol "$\mathbf{S}_T$" is the entire time series from 0 to T (dropping the plant subscript).

$\mathbf{S}_T$ is a function of time. A *functional* is a mathematical operation that pairs a function with a real value *y* (*e.g.*, for some quantitative trait). The notation is

$$y = P_y\left(\mathbf{S}_T\right) \tag{4}$$

where P_y is the phenotype functional. The combination of equations (3) and (4) constitutes a genetic network model for a single individual and trait. Although P_y can take any form, a common functional calculates the cumulative effect of an ongoing process such as grainfill. In this case

$$P_y\left(\mathbf{S}_T\right) = \int_0^T p\left(\mathbf{S}_\tau\right) d\tau \tag{5}$$

Dong (2003) developed a model of *Arabidopsis* flowering time control that specialized the above theory. The h_i were either zero or one and most g_i operated below an upper bound. The model describes the dynamics of mRNA for nine genes and four protein levels as influenced by temperature and photoperiod. Equation (4) was a rule that triggered budding when *LFY* expression (Fig. 1) exceeded a threshold. Parameter estimates were obtained by nonlinear least squares fits to two data sets. The first contained bud dates from replicated growth chamber experiments on seven mutants (*fca-6*, *fpa-2*, *fve-2*, *co-6*, *fha-1*, *gi-6*, and *phyB-1)* in the Ler background under constant photothermal regimes. Additional data were gene expression time series culled from the literature. The model accounts for 85% of the variation in observed bolting time for the seven mutants and the wild type. It also mimics rhythmic expression features of most photoperiod pathway genes. The model was validated against 114 independent observations, all but eight of which were from literature. Of the 114, four *fve-1* double mutants and two genotypes reared at 6ºC were statistical outliers that were discarded. The retained data represented a far wider range of genotypes (heterozygotes, double mutants, and over-expression lines) and environments (variable temperatures and alternative photoperiods, including continuous light) than were present in the calibration data. As is common with literature data, bud dates were reported as total leaf numbers (TLN) rather than days after planting. Typical correlations between these two measures exceed 90 %. TLN observations ranged from 4.3 to 74.4 and were quite uniformly distributed on this interval. The model accounts for 74% in the variation in TLN. For 41 data points bud dates were estimated from TLN via ancillary data. The model R^2 for these estimated dates was 76%.

Critical short day length – an emergent property
Existing crop models integrate a wealth of physiological concepts, many of which derive from physical and/or chemical first principles while others reflect empirical observations not yet mechanistically explained. An example of the latter is critical short day length (CSDL), a photoperiod above which developmental rates begin to decline (increase) in SD (LD) plants. Because accurate flowering time simulation is important, CSDL is a key parameter in some crop models (Tsuji, *et al.*, 1994) and methods have been developed for its efficient estimation (Irmak *et al.*, 2000; Welch *et al.*, 2002). But how is it determined biologically?

CO (Fig. 1) is strongly tied to day length measurement in *A. thaliana* and its expression profiles have been studied under LD and SD conditions (citations above). However, the resulting qualitative inferences cannot suggest what patterns might occur at transitional photoperiods. In contrast, Figure 3 compares *CO* loss of function to an autonomous pathway mutant as simulated by the Dong (2003) model. The latter clearly shows a photoperiod response transition with a mildly temperature-dependent CSDL. The pattern is absent in *co-6*, which has lost the ability to measure day length. Unfortunately, molecular geneticists, unlike plant physiologists, seldom investigate intermediate day lengths so the model cannot be evaluated in that range.
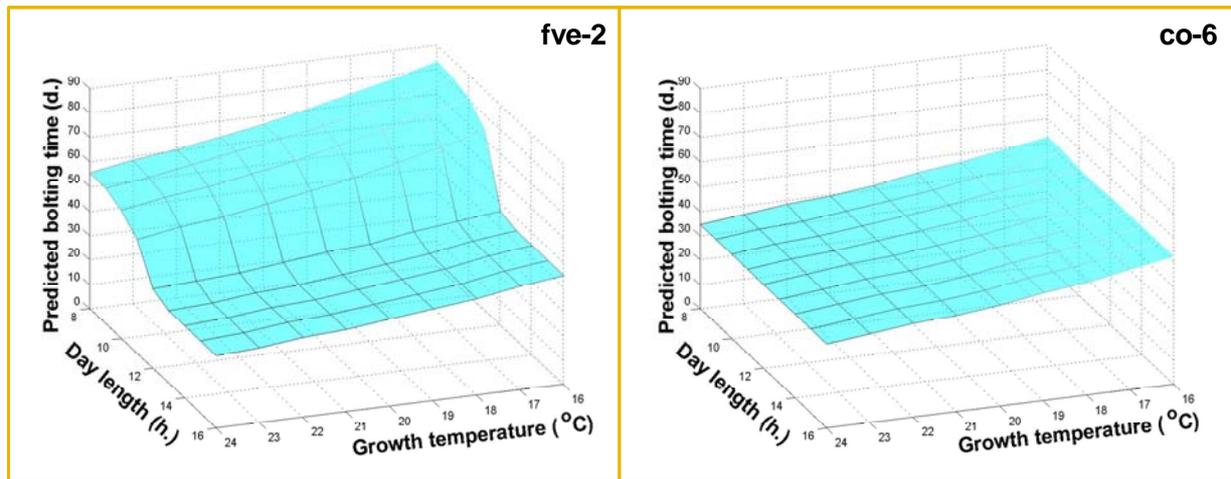
Figure 3. The autonomous path mutant (*left*) shows a CSDL below which development responds only slightly to photoperiod. A loss of function mutation removing the ability to measure day length (*right*) cancels the response.

Welch *et al.* (2005) model genes as Hopfield neurons, a special case of equation (1), and give examples of feasible signal processing functions. One example is *HEADING DATE 1* (*Hd1*), the rice (*Oryza sativa*) homolog of *CO*. The form of the model was suggested by *Hd1* time series data for Nipponbare, a *japonica* cultivar, collected under SD (9 h.) and LD (15 h.) by Kojima *et al.* (2002). Specifically,

$$\frac{d}{dt}(Hd1) = \left.\begin{matrix} R_L \\ R_D \end{matrix}\right\} g_{NN}\big(C(t)\big) - (Hd1)\left\{\begin{matrix} \lambda_L \\ \lambda_D \end{matrix}\right. \tag{6}$$

The model assumes (Suárez-López *et al.*, 2001) that rates of maximum production (R_L, R_D) or specific degradation (λ_L, λ_D) may differ under light and dark conditions (*L* and *D*, respectively). Production is governed by $g_{NN}(u) = \big[1 + \exp(-u)\big]^{-1}$, a sigmoid function used in neural networks, and is driven by a sine wave clock input, $C(t)$. Parameters minimize the sum of absolute errors against the Kojima *et al.* (2002) data.

Welch *et al.* (2005) assume that day length is encoded as the time-averaged *Hd1* level and that higher values slow development (Kojima *et al.*, 2002), making rice a SD plant. While per observation errors average 22% due to model simplicity and the variability of gene expression data, errors in the time averages are 10.3% and 1.3% for SD and LD, respectively. They next compare time averages, plotted as a function of photoperiod, with developmental rates from the rice phenology model of Yin *et al.* (1997) with Nipponbare parameters. The photoperiod above which development rates start to decline differ by *ca.* 15 min between the two models, representing the first time that a CSDL has been estimated from gene expression data alone.

The results in Figure 3 and Welch *et al.* (2005) were obtained by simulation. Both demonstrate the emergent property of a CSDL but neither explains the origin of the effect. We do so now. No assumptions are made about the clock waveform, $C(t)$, beyond periodicity, or about g, which controls transcription and translation.

Because Welch *et al.* (2005) use equation (6) only as an example, they do not report parameter values, which are R_L=1.71, R_D=1.03, λ_L=0.090, and λ_L=0.084. The 66% difference in R values *vs.* 7% for λ suggests using a single λ for both light and dark photophases. Refitting the data gives λ=0.086 and a slight increase in the goodness-of-fit. Thus, equation (6) can be rewritten as

$$\frac{dy}{dt} - ay = G(t) \tag{7}$$

where $a = -\lambda$ and $G(t) = L(t)g\big(C(t)\big)$ with $L(t) = R_L$ when the lights are on and R_D otherwise. Define *t* to be Zeitgeber (ZT) time (dawn at *t*=0) and assume that the lights are on for a fixed fraction, *f*, of the *p*=24 h. diurnal cycle. G is periodic because *C* and *L* are. The Fourier series expansion for G is

---

$$G(t) = \sum_{n=0} a_n(f)\cos(s_n t) + b_n(f)\sin(s_n t) \qquad (8)$$

where $s_n = 2n\pi / p$ and

$$a_n(f) = \left(\frac{2}{p}\right)\int_0^p G(\tau)\cos(s_n\tau)\,d\tau \qquad b_n(f) = \left(\frac{2}{p}\right)\int_0^p G(\tau)\sin(s_n\tau)\,d\tau \qquad (9)$$

Ignoring short-term transients, the solution (CRC, 1996, p. 405) of equation (7) is

$$y(t) = -\sum_{n=0} a_n(f)\left[(a\cos(s_n t) - s_n\sin(s_n t)\right]/(a^2 + s_n^2) \qquad (10)$$
$$- \sum_{n=0} b_n(f)\left[(a\sin(s_n t) + s_n\cos(s_n t)\right]/(a^2 + s_n^2)$$

Next average $y(t)$ over one period. Because $\int_0^p \cos(s_n\tau)\,d\tau = \int_0^p \sin(s_n\tau)\,d\tau = 0$ for $n > 0$ and $b_0(f) = 0$,

$$\bar{y}(f) = p^{-1}\int_0^p y(\tau)\,d\tau = a_0(f)/\lambda = \left(\frac{2}{p}\right)\left[\frac{R_L}{\lambda}\int_0^{fp} g(C(\tau))\,d\tau + \frac{R_D}{\lambda}\int_{fp}^p g(C(\tau))\,d\tau\right] \qquad (11)$$

Call the two integrals in the right hand side of equation (11) the *light* and *dark* integrals, respectively. *Hd1* production begins in the afternoon. The light integral thus contributes little for earlier *fp*, so $\bar{y}(f)$ is constant at the average of $2R_D g(C(t))/\lambda$. Since $R_L > R_D$, $\bar{y}(f)$ rises with further increases in *fp* to plateau at the average of $2R_L g(C(t))/\lambda$. Thus, CSDL in rice relates to the ZT time of the clock-driven onset of *Hd1*.

Equation (11) is unaltered if, in addition to the clock, g also responds to *f*. This is relevant to *A. thaliana* in which *CO* regulation is more complex. Under SD, *CO* expression increases in the afternoon, exhibits a single peak during the night, and drops to a low level before dawn. Under LD, however, two peaks appear in the dark with the second one extending past dawn (Suárez-López *et al.*, 2001). Although this pattern differs from *Hd1*, the result is similar. Under short but increasing day lengths, CSDL occurs, as in rice, when sunset begins to fall later than the rise in *CO*. However, $\bar{y}(f)$ increases accelerate as the second peak develops and affects the light integral. While the second peak affects the dark integral, this is more than offset by a shrinking interval of integration. Thus, both species operate consistently with the external coincidence model in that daylight during the expression period determines progress toward flowering.

Relationships between phenotype prediction technologies

Prediction of crop plant phenotypes in differing environments is of critical importance to all aspects of agriculture including new variety development (breeding and marketing), crop production management (variety selection and cultural practices), and utilization (grain quality and quantity forecasting). The need for high quality feeds creates a demand for crop phenotype data even within animal agriculture. Genetic networking (GN) is the newest of three phenotype prediction technologies applicable to plant systems. The others are quantitative genetics (QG), the basis of scientific breeding program design (Walsh, 2001), and crop simulation (CS) modeling, which has been used in research (Hanks and Ritchie, 1991), policy analysis (Rosenzweig *et al.*, 1996; Tubiello *et al.*, 1999), and decision support (McCown *et al.*, 2002).

CS models lead in the amount of biological process information they embody and the degree of quantitative integration. Hanks and Ritchie (1991) listed more than 30 published models for 13 individual crops and several general models, usable for multiple crops. These models mimic plant physiological responses to the daily or hourly dynamics of temperature, solar radiation, soil water content, and soil nutrients. Although not genomically based, CS models are the nearest things to 'virtual plants' (*sensu* Salk Institute, 2000) currently in existence. Simulated crop responses encompass crop phenological development, dry matter production, and biomass partitioning among plant tissues including economic yield (usually grain). Of these, yield has been the hardest to simulate correctly and phenology the easiest.

A natural approach to the GP problem is to try to exploit the realism of CS models (Weiss, 2003). White and Hoogenboom (1996) regressed genetic coefficients in a dry bean (*Phaseolus vulgaris*) model on the presence of dominant alleles at seven loci affecting phenology, growth habit, and seed size. Chapman *et al.* (2003) simulated sorghum breeding program alternatives by merging QU-GENE (Podlich and Cooper, 1998) with the APSIM crop model (McCown *et al.*, 1996). In this study, four traits (*i.e.*, flowering time, staygreen, canopy transpiration efficiency, and osmotic adjustment) involving 15 additive genes determined the sorghum grain yield. The complex gene × gene and gene × environment interactions on grain yield were generated indirectly through APSIM by associating the values of four adaptive traits with the number of positive alleles (Chapman *et al.*, 2003). In an interesting reversal, (Yin *et al.*, 1999a, 1999b, 2000) associated QTL's with specific parameters in a barley model as did Reymond *et al.* (2003) for maize. The implication is that genes in those intervals are responsible for the traits that the coefficients quantify, a relationship that degrades when parameters are not constant across environments (Tardieu, 2003). Work melding genomics and CS models is discussed elsewhere in these proceedings; we will attempt to interrelate GN and QG.

At least 12 contrasts can be made between current GN and QG formulations:
1. QG is algebraic; GN is dynamic (differential equations or discrete time models);
2. QG variables have phenotypic units; GN variables are dimensionless, normalized biochemical levels;
3. As a consequence, QG needs no mathematical mechanism to convert "genotype" to "phenotype" values; GN does, in the form of equation (4);
4. QG is a Mendelian theory, explicitly depicting alleles, loci, chromosomes, mutation, crossover, reproduction, selection, *etc*; GN focuses on the processes of transcription, translation, their controls, and the interactions of the resulting chemical species although mechanisms like the **X** vector in equation (4) do allow for alternative alleles;
5. Thus, QG well describes the physical structure of the genetic mechanism (chromosomes, alleles, markers *etc*.); GN better represents biochemical and information processing functions;
6. QG models are linear with interaction terms whose number can rise rapidly to impractical levels; GN models are nonlinear with *far* fewer direct interactions that are explicitly defined by directed graphs;
7. QG is a population-level theory that addresses relationships between means and (co)variances; GN operates at the individual level or, most often, lower;
8. Relatedly, QG relies heavily on basic concepts of probability; GN models may include stochastic elements (*e.g.*, due to molecular randomness) but most often do not;
9. QG yields useful results even when the genetic basis of a trait is unknown; GN models explicitly reference particular genes;
10. QG is a mature theory with an accepted set of core axioms and an expanding set of applications, some of great commercial value; GN is, as yet, none of these;
11. QG views phenotypes in a sequence of increasingly complex contexts: single allele effects at one locus (additivity), multiple allele effects at one locus (dominance), multiple locus effects (epistasis), environmental effects (general / specific / GxE); GN sees all phenotypes as emerging from the interaction of environmentally influenced network elements;
12. QG defines the allele as the unit of parental contribution, leading to the conclusion that only additive genetic variation is heritable; GN seldom considers inheritance but it should be remembered that gametophytes are living organisms and thus have functional, albeit haploid, genetic networks.

It is worth considering what traits one might want in an ideal hybrid of both theories. Plausibly, a hybrid ought to be QG-like for 4, 8, 9, and 10; GN-like for 1, 6, 11, and 12; and a blend of both for 2, 3, 5, and 7. In short, both approaches seem to have equal contributions to make. But how can such a theory be constructed?

Our flowering time research was motivated by hope that a GN model could reproduce, explain, and extend empirical relationships already built into successful crop phenology simulators. By analogy, a starting point for a hybrid GN/QG theory might be to examine the basics of QG in GN terms. Such basics would certainly include epistasis, dominance, and additivity. Of these, epistasis is trivial – the structure of the GN is precisely that of the epistatic relationships between the genes involved.
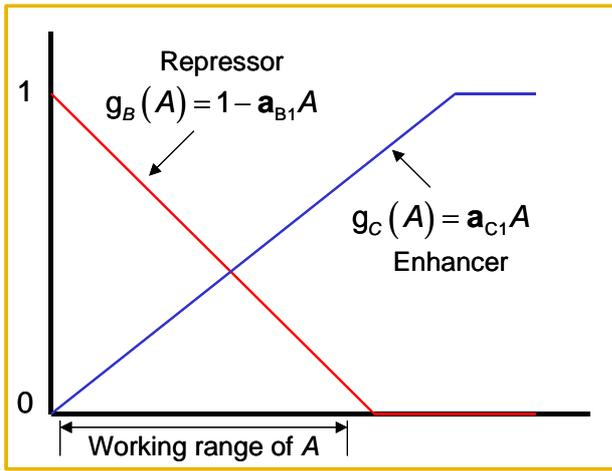
Figure 4. Gene product production functions for a repressor (g_B) and an enhancer (g_C).

Dominance is more involved. In garden peas (*Pisum sativum*) Mendel observed that the presence of certain alleles (*e.g.*, Round, Yellow) disproportionately influenced phenotype. It was a short mental step to define a property called "dominance" and attribute it to these special alleles. However, no satisfactory explanation for dominance emerged for many years. Indeed, the decades-long, acrimonious debate between Sewell Wright and R.A. Fisher on this topic ruined their early friendship (Lynch and Walsh, 1998). In 1981, Kacser and Burns gave an account of dominance for genes encoding enzymes, the only portion of the genome that had been heavily studied at that time. Fundamental to their reasoning was the shape of enzyme reaction rate curves that rise from zero to a plateau. Since the g and h functions presented here have upper bounds, similar arguments can be applied to processes involving RNA.

However, defining dominance as an allelic property may undervalue the role of other genes in manifesting dominant phenotypes. The genetic background was central in Fisher's (1928a,b) theory of dominance (now out of favor). The background is also present, albeit much reduced, in Kacser and Burns (1981) model. They depict a structureless background that influences the gene of interest via an unspecified distribution of interaction strengths. This ignores the apparent existence of fine-scale modularity in gene networks (Ravasz *et al.*, 2002) and the extensive computational abilities of small genetic circuits (Welch *et al.*, 2005).

As an example, consider a gene A that regulates B and C but in opposite directions ($C \leftarrow A$ " B). Let the dynamics be modeled with a simplified form of equation (3), $\dot{s} = X_s R_s g_s - \lambda_s s$, where s is either A, B, or C. If A is constitutively expressed, then, from the definition of X_A in equation (2) and by setting $\dot{A} = 0$, the possible equilibrium levels are $\bar{A} = 2\alpha_A R_A / \lambda_A$, $(\alpha_A + \alpha_a)R_A / \lambda_A$, and $2\alpha_a R_A / \lambda_a$ where the subscripts on α denote two different alleles. Let g_B and g_C be the simple piecewise linear functions shown in Figure 4. Finally, suppose that phenotype is determined by the heterodimer $D \leftarrow B + C$, whose formation is governed by the Law of Mass Action. Then the steady state levels of B, C, and D are

$$\bar{B} = X_B R_B (1 - \mathbf{a_{B1}}\bar{A}) / \lambda_B \tag{12}$$

$$\bar{C} = X_C R_C (\mathbf{a_{C1}}\bar{A}) / \lambda_C$$

$$\bar{D} = K_{eq}\bar{B}\bar{C} = K_{eq}\mathbf{a_{C1}}(X_B R_B)(X_C R_C)(\bar{A} - \mathbf{a_{B1}}\bar{A}^2)/(\lambda_B \lambda_C)$$

It is obvious that $\bar{D}$ is maximized when $\bar{A} = 1/(2\mathbf{a_{B1}})$. This would happen at the steady state level of the A heterozygote if the system parameters were such that $(\alpha_a + \alpha_A) = \lambda_A /(2\mathbf{a_{B1}}R_A)$. Such exactitude is unlikely. However, there will be finite parameter ranges for which the condition is well enough satisfied that $\bar{D}$ for the A heterozygote exceeds that of either homozygote. This demonstrates a simple model for over-dominance. Many other circuits would behave similarly.

Additivity is, surprisingly, a greater mystery than either dominance or epistasis, especially when present at high levels as in flowering time. The prerequisite for additivity is linearity, taken for granted by many to be the simplest mathematical form. In reality, linearity is merely the most *tractable* mathematical form.

Frequent observations of additivity and the ease of subsequent QG derivations may have obscured the depth of the additivity puzzle. Clearly, linearity is conspicuously absent in equation (3), given the kinetic complexities buried in **g** and **h**. CS models are also highly nonlinear, even in phenology (Yin *et al.*, 1997). Nevertheless, examples of additivity have been observed in CS outputs (Boote *et al.*, 2003) and in genetic coefficients, which can also be viewed as quantitative traits (White and Hoogenboom, 1996; Stewart *et al.*, 2003). Finally, while one to three genes are sufficient to create epistasis and dominance effects, plausibly a single gene can destroy linearity and, thus, additivity. So, what properties preserve additivity in genetic networks despite the seeming rarity and fragility of linearity?

Suppose that none of the downstream genes regulated by A (either directly or indirectly) in turn regulates A. That is, A is part of one or more feed-forward paths but is not in any feedback loop. Under these conditions g_A and h_A are not dependent on A and, without loss of generality, can be written as functions of time only. The A component of equation (3) is then

$$\dot{s}_A + \lambda h_A(t) s_A = X_A R_A g_A(t) \tag{13}$$

which is a linear differential equation. That is, if $s_1(t)$ is a solution to equation (13) when written without X_A, and if α and β are the X_A values for two different alleles of A, then $s_A(t) = 2\alpha s_1(t)$ for the α homozygote, $(\alpha + \beta) s_1(t)$ for the $\alpha\beta$ heterozygote, and $2\beta s_1(t)$ for the β homozygote.

For phenotypes to be additive, linearity must be preserved by downstream **g** and **h** functions and by the phenotype functional in equation (4). Relative to the former, we are studying simple thermodynamic models of promoter reactions based on concepts in (Kingston and Narlikar, 1999), which can produce curves like those in Figure 4 (Welch, unpub.). A manipulation of equation (13) shows that the "working range linearity" in Figure 4 maintains additivity even in feedback situations. Linearity in phenotype functionals is promoted when they are accumulative, as is equation (5), since integration is a linear operation. A specific example is the time averaging in equation (11). In the over-dominance illustration, however, the chain of linearity was broken by $\bar{D}$, which, as a protein, does not have an X-factor and whose levels are governed by the nonlinear Law of Mass Action. Note that the over-dominance of A would disappear if either B or C were to be present in excess since the formation of D would then have first order kinetics. This illustrates how dominance behaviour can depend on particular backgrounds or environments.

These arguments show that additivity, far from being an allelic property, depends on the structure and operation of entire gene subnets. Feed-forward mechanisms may be important as they introduce "new" linearity (via X-factors) at each successive pathway locus. In this vein, it is interesting that, aside from the diurnal oscillator and meristem identity switches, the flowering time control network seems to be primarily feed-forward (Figure 1). A second example (Davidson *et al.*, 2002) is the elaborate network controlling the embryonic differentiation of sea urchin (*Strongylocentrotus purpuratus*) endomesoderm. This network is also remarkably feed-forward (to visual inspection), except for switches that turn on major subsections.

Of course, two feed-forward examples barely qualify as anecdotal evidence, especially given the apparent incidence of feedback inhibition and other mechanisms of physiological homeostasis. Perhaps feed-forward characteristics are more closely associated with certain processes or hierarchical levels (*sensu* Csete and Doyle, 2002) of internal plant control than others. Both of the above examples are high-level developmental processes and perhaps different patterns are found elsewhere. Whatever the truth may be, the importance of heritability in breeding programs seemingly indicates a high priority on understanding its origin, additivity.

Synthesis
As research on the GP problem progresses, the three existing approaches to phenotype prediction (GN, QG, and CS models) will increasingly synergize. Genetic networks can be expected to contribute to each of the others. For example, equation (11) can be converted into an algebraic form directly usable in CS models. The light and dark integrals can be evaluated once for each of a series of fp that span the photoperiodic range and interpolating polynomials, $P_L(f)$ and $P_D(f)$, constructed from the results (Press *et al.*, 1992). Welch *et al.* (2005) incorporate thermal effects into genetic network models by replacing

$\left(R/\lambda \right)$ quotients with what they call form-invariant functions of temperature, T, such as the beta model of Yin *et al.* (1995). This gives a developmental rate of $1/\overline{y} = 1/\left[\eta_L \beta_L(T) P_L(f) + \eta_D \beta_D(T) P_D(f) \right]$ where the η's are constants to be fit.

Yin *et al.* (1999a, 1999b, 2000) and Reymond *et al.* (2003) mapped QTL's for crop model coefficients. It would be interesting to attempt the same for gene parameters in a network model. Lacking complete knowledge of the network, models can only include subsets of the genes actually present. Thus parameter estimates will be lumped values influenced by other genes closely associated in the network. Quite possibly, parameter QTL mapping will reveal the regions containing these other genes, thereby contributing to gene discovery and network expansion. If nothing else, mapping the parameters of any given gene should yield a QTL containing that gene. This would be a known-ground-truth validation of the Yin-Reymond approach.

According to QG, the rate of crop improvement through selection is proportional to heritability (*i.e.*, the additivity of controlling genes). Perhaps targeted substitutions at influential, non-additive loci might accelerate this process when heritability is low. GN could help plan such efforts by giving quantitative estimates of expected phenotypes, particularly in situations involving multiple, non-additive loci. More speculatively, it may sometimes be possible to increase trait additivity by focal disruption of non-additive circuitry. Low heritability also results when a trait is governed by multiple upstream processes that contribute individual increments of environmental variation. GN can help by identifing more heritable upstream selection targets including traditional, physiological observables and/or desirable patterns of upstream gene expression. Finally, efforts are currently underway to construct libraries of useful alleles by mining large varietal collections (McNally, 2004). GN methods can add value to such libraries through more accurate prediction of non-additive allelic effects at multiple loci in differing backgrounds and environments.

In 2001 an NSF reviewer told the authors that models like the above are simplistic and "unlikely to have relevance to real genetic systems", a view oblivious to the 80 years of utility demonstrated by even simpler QG models (Fisher, 1918; Wright, 1921a-d). Perhaps because the vast majority of differential equations lack closed-form solutions, little attention has been given to genetic network models where some analysis is feasible. Yet, in ecology, a domain whose complexity rivals genomics, simple mathematical models have generated insight and practical uses for nearly as long (Lotka, 1925; Kot 2001). It may be that, without computers, early ecologists had to develop insights by other means. These insights now guide ecologists even as they use computers, today of ubiquitous importance. Currently, bioinformaticists struggle to develop software that can extract meaning from large masses of genomic data. Perhaps by first learning from models that have some realism but great simplicity, we can find better ways to guide the machines in their search.

References
Akutsu, T., Miyano, S., and Kuhara, S. (1999). Identification of genetic networks from a small number of gene expression patterns under the boolean network model. In: Proceedings of the Pacific Symposium on Biocomputing, 4:17-28. World Publishing Co., Singapore.

Akutsu, T., Miyano, S., and Kuhara, S. (2000). Inferring qualitative relations in genetic networks and metabolic pathways. Bioinformatics, 16:727-34.

Alabadi, D., Oyama, T., Yanovsky, M., Harmon, F., Mas, P., and Kay, S. (2001). Reciprocal regulation between *TOC1* and *LHY/CCA1* within the *Arabidopsis* circadian clock. Science 293:880-883.

Auckerman, M., and Sakai, H. (2003). Regulation of Flowering Time and Floral Organ Identity by a MicroRNA and Its *APETALA2*-Like Target Genes. Plant Cell 16:2730-2741.

Bagnall, D., King, R., Whitelam, G., Boylan, M., Wagner, D., and Quail, P. (1995). Flowering responses to altered expression of phytochrome in mutants and transgenic lines of *Arabidopsis thaliana* (L.) Heynh. Plant Physiol 108:1495-1503.

Baldi, P., and Hatfield, G. (2002). DNA Microarrays and Gene Expression. (Cambridge, UK, Cambridge University Press).

Barash, Y., and Friedman, N. (2001). Context specific bayesian clustering for gene expression data. In: Proceedings of the Fifth Annual International Conference on Computational Molecular Biology RECOMB-2001:2-11. ACM_SIGACT, New York.

Blazquez, M. (2000). Flower development pathways. J Cell Science 113:3547-3548.

Blazquez, M., Green, T., Nilsson, O., Sussman, M., and Weigel, D. (1998). Gibberellins promote flowering of *Arabidopsis* by activating the *LEAFY* promoter. Plant Cell 10:791-800.

Blazquez, M., Koornneef, M., and Putterill, J. (2001). Flowering on time: Genes that regulate the floral transition. EMBO reports 2:1078-1082.

Blazquez, M., Trenor, M., and Weigel, D. (2002). Independent control of gibberellin biosynthesis and flowering time by the circadian clock in *Arabidopsis*. Plant Physiol 130:1770-1775.

Blazquez, M., and Weigel, D. (1999). Independent regulation of flowering by phytochrome B and gibberellins in *Arabidopsis*. Plant Physiol 120:1025-1032.

Blazquez, M. A., Ahn, J., and Weigel, D. (2003). A thermosensory pathway controlling flowering time in *Arabidopsis thaliana*. Nature Genetics 33:168-171.

Boote, K., Jones, J., Batchelor, W., Nafziger, E., and Myers, O. (2003). Genetic coefficients in the CROPGRO-Soybean model: Links to field performance and genomics. Agron J 95:32-51.

Borner, R., Kampmann, G., Chandler, J., Gleissner, R., Wisman, E., Apel, K., and Melzer, S. (2000). A *MADS* domain gene involved in the transition to flowering in *Arabidopsis*. Plant J 24:591-599.

Bowman, J., Alvarez, J., Weigel, D., Meyerowitz, E., and Smyth, D. (1993). Control of flower development in *Arabidopsis thaliana* by *APETALA1* and interacting genes. Dev 119:721-743.

Bunning, E. (1936). Die endonome tagesrhythmik als grundlage der photoperiodischen reaktion. Ber Deut Bot Ges 54:590-607.

Burd, C., and Dreyfuss, G. (1994). Conserved structures and diversity of functions of RNA-binding proteins. Science 265:615–621.

Busch, M., Bornblies, K., and Weigel, D. (1999). Activation of a floral homeotic gene in *Arabidopsis*. Science 285:585-587.

Carre, I. (2001). Daylength perception and the photoperiodic regulation of flowering in *Arabidopsis*. J Biol Rhythms 16:415-423.

Carre, I., and Kim, J. (2002). MYB transcription factors in the *Arabidopsis* circadian clock. J Exp Bot 53:1551-1557.

Carrington, J., and Ambros, V. (2003). Role of microRNAs in plant and animal development. Science 301:336-338.

Chandler, J., Wilson, A., and Dean, C. (1996). *Arabidopsis* mutants showing an altered response to vernalization. Plant J 10:637–644.

Chapman, S., Cooper, M., Podlich, D., and Hammer, G. (2003). Evaluating plant breeding strategies by simulating gene action and dryland environment effects. Agron J 95:99-113.

Chen, T., He, H., and Church, G. (1999). Modeling gene expressions with differential equations. In: Proceedings of Pacific Symposium on Biocomputing, 4:17-28. World Publishing Co., Singapore.

Chou, M., Haung, M., and Yang, C. (2001). *EMF* genes interact with late-flowering genes in regulating floral initiation genes during shoot development in *Arabidopsis thaliana*. Plant Cell Physiol 42:499-507.

Cooper, M., Chapman, S., Podlich, D., and Hammer, G. (2002). The GP problem: Quantifying gene-to-phenotype relationships. *In Silico* Biol 2:151-164.

CRC (1996). Standard mathematical tables and formula. (Boca Raton, FL, CRC Press).

Csete, M., and Doyle, J. (2002). Reverse engineering of biological complexity. Science 295:1664-1669.

D'Haesseleer, P., Wen, X., Fuhrman, S., and Somogyi, R. (1999). Linear modeling of mRNA expression levels during CNS development and injury. In: Proceedings of the Pacific Symposium on Biocomputing, 4: 41-52. World Publishing Co., Singapore.

Davidson, E., Rast, J., Oliveri, P., Ransick, A., Calestani, C., Yuh, C., Minokawa, T., Amore, G., Hinman, V., Arenas-Mena, C., *et al.* (2002). A genomic regulatory network for development. Science 295:1670-1678.

Davis, S. (2002). Photoperiodism: The coincidental perception of the season. Curr Biol 12:R841-R843.

Devlin, P. (2002). Signs of the time: environmental input to the circadian clock. J Exp Bot 53:1535-1550.

Dong, Z. (2003) Incorporation of genomic information into the simulation of flowering time in *Arabidopsis thaliana*. Ph.D. dissertation, Kansas State University.

Doyle, M., Davis, S., Bastow, R., McWatters, H., Kozma-Bognar, L., Nagy, F., Millar, A., and Amasino, R. (2002). The *ELF4* gene controls circadian rhythms and flowering time in *Arabidopsis thaliana*. Nature 419:74-77.

Eriksson, M., and Millar, A. (2003). The circadian clock. A plant's best friend in a spinning world. Plant Physiol 132:732-738.

Fisher, R. (1918). The correlation between relatives on the supposition of Mendelian inheritance. Trans Royal Soc Edinburgh 52:399-433.

Fisher, R. (1928a). The possible modification of the response of the wild type to recurrent mutations. Am Nat 62:115-126.

Fisher, R. (1928b). Two further notes on the origin of dominance. Am Nat 62:571-574.

Fowler, S., Lee, K., Onouchi, H., Samach, A., Richardson, K., Morris, B., Coupland, G., and Putterill, J. (1999). *GIGANTEA*: a circadian clock-controlled gene that regulated photoperiodic flowering in *Arabidopsis* and encodes a protein with several possible membrane-spanning domains. EMBO J 18:4679-4688.

Frank, S. (1998). Population and quantitative genetics of regulatory networks. J Theoretical Biol 197:281-294.

Friedman, N., Linial, M., Nachman, I., and Pe'er, D. (2000). Using Bayesian networks to analyze expression data. J. Comput. Biol. 7:601-620.

Gendall, A., Levy, Y., Wilson, A., and Dean, C. (2001). The *VERNALIZATION2* gene mediates the epigenetic regulation of vernalization in *Arabidopsis*. Cell 107:525–535.

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

Goss, P., and Peccoud, J. (1999). Analysis of the stabilizing effect of Rom on the genetic network controlling ColE1 plasmid replication. Pac Symp Biocomput. 4:65-76

Goto, N., Kumagai, T., and Koornneef, M. (1991). Flowering responses to lightbreaks in photomorphogenic mutants of *Arabidopsis thaliana*, a long-day plant. Physiol Plant 83:209-215.

Guo, H., Yang, H., Mocker, T., and Lin, C. (1998). Regulation of flowering time by *Arabidopsis* photoreceptors. Science 279:1360-1363.

Halliday, K., Salter, M., Thingnaes, E., and Whitelam, G. (2003). Phytochrome control of flowering is temperature sensative and correlates with expression of the floral integrator *FT*. Plant J 33:875-885.

Hanks, J., and Ritchie, J. (1991). Modeling plant and soil systems. Agron. Monogr. 31. ASA, CSSA, and SSSA, Madison, WI.

Hartemink, A., Gifford, D., Jaakkola, T., and Young, R. (2001). Using graphical models and genomic expression data to statistically validate models of genetic regulatory networks. In: Proceedings of the Pacific Symposium on Biocomputing, 6:422-433. World Publishing Co., Singapore.

Haughn, G., Schultz, E., and Martinez-Zapater, T. (1995). The regulation of flowering in *Arabidopsis thaliana*: meristems, morphogenesis, and mutants. Canadian J Botany 73:959-981.

Hayama, R., and Coupland, G. (2003). Shedding light on the circadian clock and the photoperiodic control of flowering. Curr Opin Plant Biol 6:13-19.

Hempel, F., Weigel, D., Mandel, M., Ditta, G., Zambryski, P., Feldman, L., and Yanofsky, M. (1997). Floral determination and expression of floral regulatory genes in *Arabidopsis*. Dev 124:3845-3853.

Hepworth, S., Valverde, F., Ravenscroft, D., Mouradov, A., and Coupland, G. (2002). Antagonistic regulation of flowering-time gene *SOC1* by *CONSTANS* and *FLC* via separate promoter motifs. EMBO J 21:4327-4337.

Hicks, K., Albertson, T., and Wagner, D. (2001). *EARLY FLOWERING3* encodes a novel protein that regulates circadian clock function and flowering in *Arabidopsis*. Plant Cell 13:1281-1292.

Hudson, M. (2000). The genetics of phytochrome signaling in *Arabidopsis*. Sem Cell Dev Biol 11:475-483.

Huq, E., Tepperman, J., and Quail, P. (2000). *GIGANTEA* is a nuclear protein involved in phytochrome signaling in *Arabidopsis*. Proc Natl Acad Sci 97:9789-9794.

Ideker, T., Thorsson, V., and Karp, R. (2000). Discovery of regulatory interactions through perturbation: inference and experimental design. In: Proceedings of the Pacific Symposium on Biocomputing, 5:302-313. World Publishing Co., Singapore.

Irmak, A., Jones, J., Mavromatis, T., Welch, S., Boote, K., and Wilkerson, G. (2000). Evaluating Methods for simulating soybean cultivar responses using cross validation. Numbered contribution of the Florida AES. Agron J 92:1140-1149.

Jarillo, J., Capel, J., Tang, R.-H., Yang, H.-Q., Alonso, J., Ecker, J., and Cashmore, A. (2001). An *Arabidopsis* circadian clock component interacts with both *CRY1* and *phyB*. Nature 410:487-490.

Johanson, U., West, J., Lister, C., Michaels, S., Amasino, R., and Dean, C. (2000). Molecular analysis of *FRIGIDA*, a major determinant of natural variation in *Arabidopsis* flowering time. Science 290:344-347.

Johnson, C. (2001). Endogenous timekeepers in photosynthetic organisms. Annu Rev Physiol 63:695-728.

Johnson, E., Bradley, M., Harberd, N., and Whitelam, G. (1994). Photoresponses of light-grown *phyA* mutants of *Arabidopsis*. Plant Physiol 105:141-149.

Kacser, H., and Burns, J. (1981). The molecular basis of dominance. Genetics 97:639-666.

Kardailsky, I., Shakla, V., Ahn, J., Dagenais, N., Christensen, S., Nguyen, J., Chory, J., Harrison, M., and Weigel, D. (1999). Activation tagging of the floral inducer *FT*. Science 286:1962-1965.

Kenzior, A., and Folk, W. (1998). AtMSI4 and RbAp48 WD-40 repeat proteins bind metal ions. FEBS Lett 440:425-429.

Kingston, R., and Narlikar, G. (1999). ATP-dependent remodeling and acetylation as regulators of chromatin fluidity. Genes Dev 13:2339-2352.

Kobayashi, Y., Kaya, H., Goto, K., Iwabuchi, M., and Araki, T. (1999). A pair of related genes with antagonistic roles in mediating flowering signals. Science 286:1960-1962.

Kojima, S., Takahashi, Y., Kobayashi, Y., Monna, L., Sasaki, T., Araaki, T., and Yano, M. (2002). *Hd3a*, a rice ortholog of the *Arabidopsis FT* gene, promotes transition to flowering downstream of *Hd1* under short-day conditions. Plant Cell Physiol 43:1096-1105.

Koornneef, M., Hanhart, C., and van der Veen, J. (1991). A genetic and physiological analysis of late flowering mutants in *Arabidopsis thaliana*. Mol Gen Genet 229:57–66.

Kot, M. (2001). Elements of Mathematical Ecology. (Cambridge, UK, Cambridge University Press).

Lamb, R., Hill, T., Tan, Q., and Irish, V. (2002). Regulation of *APETALA3* floral homeotic gene expression by meristem identity genes. Devel 129:2079-2086.

Lee, H., Suh, S., Park, E., Cho, E., Ahn, J., Kim, S., Lee, S., Kwon, Y., and Lee, I. (2000). The *AGAMOUS-LIKE 20 MADS* domain protein integrates floral inductive pathways in *Arabidopsis*. Genes Devel 14:2366-2376.

Lee, I., Aukerman, M., Gore, S., Lohman, K., Michaels, S., Weaver, L., John, M., Feldmann, K., and Amasino, R. (1994). Isolation of *LUMINIDEPENDENS*: a gene involved in the control of flowering time in *Arabidopsis*. Plant Cell 6:75-83.

Lee, S., Cheng, H., King, K., Wang, W., He, Y., Hussain, A., Lo, J., Harberd, N., and Peng, J. (2002). Gibberellin regulates *Arabidopsis* seed germination via *RGL2*, a *GAI/RGA*-like gene whose expression is up-regulated following imbibition. Genes Devel 16:646-658.

Levy, Y., and Dean, C. (1998). The transition to flowering. Plant Cell *10*, 1973-1989.

Levy, Y., Mesnage, S., Mylne, J., Gendall, R., and Dean, C. (2002). Multiple roles of *Arabidopsis VRN1* in vernalization and flowering time control. Science 297:243-247.

Liang, Fuhrman, and Somogyi, R. (1998). REVEAL: A general reverse engineering algorithm for inference of genetic network architecture. In: Proceedings of the Pacific Symposium on Biocomputing, 3:18-29. World Publishing CO., Singapore.

Liljegren, S. J., Gustafson-Brown, C., Pinyopich, A., Ditta, G., and Yanofsky, M. (1999). Interactions among *APETALA1, LEAFY*, and *TERMINAL FLOWER1* specify meristem fate. Plant Cell 11:1007–1018.

Lin, C. (2000). Photoreceptors and regulation of flowering time. Plant Physiol 123:39-50.

Liu, X., Covington, M., Fankhauser, C., Chory, J., and Wagner, D. (2001). *ELF3* encodes a circadian clock-regulated nuclear protein that functions in an *Arabidopsis PHYB* signal transduction pathway. Plant Cell 13:1293-1304.

Lotka, A. (1925). Elements of physical biology. (Baltimore, Williams & Wilkins Co.).

Lynch, M., and Walsh, B. (1998). Genetics and analysis of quantitative traits. (Sunderland, MA, Sinauer Associates, Inc.).

Macknight, R., Bancroft, I., Page, T., Lister, C., Schmidt, R., Love, K., Westphal, L., Murphy, G., Sherson, S., Cobbett, C., and Dean, C. (1997). *FCA*, a gene controlling flowering time in *Arabidopsis*, encodes a protein containing RNA-binding domains. Cell 89:737–745.

Macknight, R., Duroux, M., Laurie, R., Dijkwel, P., Simpson, G., and Dean, C. (2002). Functional significance of the alternative transcript processing of the *Arabidopsis* floral promoter *FCA*. Plant Cell 14:877-888.

Maki, Y., Tominaga, D., Okamoto, M., Watanabe, S., and Eguchi, Y. (2001). Development of a system for the inference of large scale genetic networks. In: Proceedings of the Pacific Symposium on Biocomputing, 6:446-458. World Publishing Co., Singapore.

Makino, S., Kiba, T., Imamura, A., Hanaki, N., Nakamura, A., Taniguchi, M., Ueguchi, C., Sugiyama, T., and Mizuno, T. (2000). Genes encoding pseudo-response regulators: insight into *His*-to-*Asp* phosphorelay and circadian rhythm in *Arabidopsis thaliana*. Plant Cell Physiol 41:791-803.

Makino, S., Matsushika, A., Kojima, M., Oda, Y., and Mizuno, T. (2001). Light response of the circadian waves of the *APRR1/TOC1* quintet: When does the quintet start singing rhythmically in *Arabidopsis*? Plant Cell Physiol 42:334-339.

Makino, S., Matsushika, A., Kojima, M., Yamashino, T., and Mizuno, T. (2002). The *APRR1/TOC1* quintet implicated in circadian rhythms of *Arabidopsis thaliana*: I. characterization with *APRR1*-overexpressing plants. Plant Cell Physiol 43:58-69.

Mandel, M., Gustafson-Brown, C., Savidge, B., and Yanofsky, M. (1992). Molecular characterization of the *Arabidopsis* floral homeotic gene *APETALA1*. Nature 360:273-277.

Marnellos, G., Deblandre, G., Mjolsness, E., and Kintner, C. (2000). Delta-notch lateral inhibitory patterning in the emergence of ciliated cells in *Xenopus*: Experimental observations and a gene network model. In: Proceedings of the Pacific Symposium on Biocomputing, 5:326-337. World Publishing Co., Singapore.

Martinez-Zapater, J., Coupland, G., Dean, C., and M. Koornneef, M. (1994). The transition to flowering in *Arabidopsis*. In *Arabidopsis*, E. Meyerowitz, and C. Somerville, eds. (Cold Spring Harbor, NY, Cold Spring Harbor Lab Press), pp. 403-433.

Martinez-Zapater, J., and Somerville, C. (1990). Effect of light quality and vernalization on late-flowering mutants of *Arabidopsis thaliana*. Plant Physiol 92:770–776.

Martínez-Zapater, J. M., Jarillo, J., Cruz-Alvarez, M., Roldán, M., and Salinas, J. (1995). *Arabidopsis* late-flowering fve mutants are affected in both vegetative and reproductive development. Plant J 7:543-551.

Matsuno, H., Doi, A., Nagasaki, M., and Miyano, S. (2000). Hybrid Petri net representation of gene regulatory network. In: Proceedings of the Pacific Symposium on Biocomputing, 5:338-349. World Publishing Co., Singapore.

Matsushika, A., Imamura, A., Yamashino, T., and Mizuno, T. (2002a). Aberrant expression of the light-inducible and circadian-regulated *APRR9* gene belonging to the circadian-associated *APRR1/TOC1*

quintet results in the phenotype of early flowering in *Arabidopsis thaliana*. Plant Cell Physiol 43:833-843.

Matsushika, A., Makino, S., Kojima, M., and Mizuno, T. (2000). Circadian waves of expression of the *APRR1/TOC1* family of pseudo-response regulators in *Arabidopsis*. Plant Cell Physiol 41:1001-1012.

Matsushika, A., Makino, s., Kojima, M., Yamashino, T., and Mizuno, T. (2002b). The *APRR1/TOC1* quintet implicated in circadian rhythms of *Arabidopsis thaliana*: II. characterization with *CCA1*-overexpressing plants. Plant Cell Physiol 43:118-122.

McClung, C. (2001). Circadian rhythms in plants. Annu Rev Plant Physiol Biol 52:139-162.

McCown, R. L., Hammer, G., Hargreaves, J., Holzworth, D., and Freebairn, D. (1996). APSIM: A novel software system for model development, model testing, and simulation in agricultural systems research. Ag Syst 50:255-271.

McCown, R. L., Hochman, Z., and Carberry, P. (2002). Probing the enigma of decision support systems for farmers: Learning from experience and from theory. Ag Sys 74:1-10.

McNally, K. (2004). Allele mining workshop. In: Plant and Animal Genome XII Conference. January 10-4. San Diego, California. http://www.intl-pag.org/12/12-allele.html

Mendoza, L., and Alvarez-Buylla, E. (1998). Dynamics of the Genetic Regulatory Network for *Arabidopsis thaliana* Flower Morphogenesis. J Theoretical Biol 193:307-319.

Mendoza, L., and Alvarez-Buylla, E. (2000). Genetic Regulation of Root Hair Development in *Arabidopsis thaliana*: A Network Model. J Theoretical Biol 204:311-326.

Michaels, S., and Amasino, R. (1999). *FLOWERING LOCUS C* encodes a novel *MADS* domain protein that acts a repressor of flowering. Plant Cell 11:949-956.

Michaels, S., He, Y., Scortecci, K., and Amasino, R. (2003). Attenuation of *FLOWERING LOCUS C* activity as a mechanism for the evolution of summer-annual flowering behavior in *Arabidopsis*. Proc Natl Acad Sci 100:10102-10107.

Millar, A. (1999). Biological clocks in *Arabidopsis thaliana*. New Phytol 141:175-197.

Mockler, T., Guo, H., Yang, H., Duong, H., and Lin, C. (1999). Antagonistic actions of *Arabidopsis* cryptochromes and phytochrome B in the regulation of floral induction. Dev 126:2073-2082.

Moon, J., Suh, S.-S., Lee, H., Choi, K.-R., Hong, C., Paek, N.-C., Kim, S.-G., and Lee, I. (2003). The *SOC1 MADS*-box gene integrates vernalization and gibberelling signals for flowering in *Arabidopsis*. Plant J 35:613-623.

Morel, P., Tréhin, C., Monéger, F., and Négrutiu, I. (2002). Control of floral induction and plant yield: role of the *MSI4* gene coding for a WD-repeat protein. In: XIII International Conference on Arabidopsis Research (Seville, Spain).

Mouradov, A., Cremer, F., and Coupland, G. (2002). Control of flowering time: integrating pathways as a basis for diversity. Plant Cell 14:S111-S130.

Murakami-Kojima, M., Nakamichi, N., Yamashino, T., and Mizuno, T. (2002). The *APRR3* component of the clock-associated *APRR1/TOC1* quintet is phosphorylated by a novel protein kinase belonging to the WNK Family, the gene for which is also transcribed rhythmically in *Arabidopsis thaliana*. Plant Cell Physiol 43:675-683.

Nelson, D., Lasswell, J., Rogg, L., Cohen, M., and Bartel, B. (2000). *FKF1*, a clock-controlled gene that regulates the transition to flowering time in *Arabidopsis*. Cell 101:331-340.

Ng, M., and Yanofsky, M. (2001). Activation of the *Arabidopsis* B class homeotic genes by *APETALA1*. Plant Cell 4:739-753.

Park, D., Somers, D., Kim, Y., Choy, Y., Lim, H., Soh, M., Kim, H., Kay, S., and Nam, H. (1999). Control of circadian rhythms and photoperiodic flowering by the *Arabidopsis GIGANTEA* gene. Science 285:1579-1582.

Park, W., Li, J., Song, R., Messing, J., and Chen, X. (2002). *CARPEL FACTORY*, a *Dicer* homolog, and *HEN1*, a novel protein, act in microRNA metabolism in *Arabidopsis thaliana*. Curr Biol 12:1484-1495.

Peng, J., Carol, P., Richards, D., King, K., Cowling, R., Murphy, G., and Harberd, N. (1997). The *Arabidopsis GAI* gene defines a signaling pathway that negatively regulates gibberellin responses. Gen Dev 113:194-207.

Pittendrigh, C. (1972). Circadian surfaces and the diversity of possible roles of circadian organization in photoperiodic induction. Proc Natl Acad Sci 69:2734-2737.

Podlich, D., and Cooper, M. (1998). QU-GENE: a simulation platform for quantitative analysis of genetic models. Bioinformatics 14:632-653.

Press, W. H., Teukolsky, S., Vetterling, W., and Flannery, B. (1992). Numerical recipes in C: the art of scientific computing. (Cambridge, Cambridge University Press).

Putterill, J., Robson, F., Lee, K., Simon, R., and Coupland, G. (1995). The *CONSTANS* gene of *Arabidopsis* promotes flowering and encodes a protein showing similarities to zinc finger transcription factors. Cell 80:847-857.

Quesada, V., Macknight, R., Dean, C., and Simpson, G. (2003). Autoregulation of *FCA* pre-mRNA processing controls *Arabidopsis* flowering time. EMBO J 22:3142-3152.

Ravasz, E., A. Somera, D. Mongru, Z. Oltvai, and A.L. Baraba´si. (2002). Hierarchical organization of modularity in metabolic networks. Science 297:1551-1555.

Reed, J., Nagpal, P., Poo1e, D., Furuya, M., and Chory, J. (1993). Mutations in the gene for the red/far-red light receptor phytochrome B alter cell elongation and physiological responses throughout *Arabidopsis* development. Plant Cell 5:147-157.

Reeves, P., and Coupland, G. (2000). Response of plant development to environment: control of flowering by daylength and temperature. Curr Op Plant Biol 3:37-42.

Reeves, P., and Coupland, G. (2001). Analysis of flowering time control in *Arabidopsis* by comparison of double and triple mutants. Plant Physiol 126:1085-1091.

Reinitz, J., and Sharp, D. (1995). Mechanism of eve stripe formation. Mechanisms of Development 49:133-158.

Reymond, M., Muller, B., Leonardi, A., Charcosset, A., and Tardieu, F. (2003). Combining quantitative trait loci analysis and an ecophysiological model to analyze the genetic variability of the responses of maize leaf growth to temperature and water deficit. Plant Physiol 131:664–675.

Robson, F., Costa, M., Hepworth, S., Vizir, I., Pineiro, M., Reeves, P., Putterill, J., and G. Coupland, G. (2001). Functional importance of conserved domains in the flowering-time gene *CONSTANS* demonstrated by analysis of mutant alleles and transgenic plants. Plant J 28:619-631.

Roden, L., Song, H., Jackson, S., Morris, K., and Carre, I. (2002). Floral responses to photoperiod are correlated with the timing of rhythmic expression relative to dawn and dusk in *Arabidopsis*. Proc Natl Acad Sci 99:13313-13318.

Rosenzweig, C., Philips, J., Goldberg, R., Carroll, J., and Hodges, T. (1996). Potential impacts of climate change on citrus and potato production in the US. Ag Sys 52:455-479.

Rouse, D. T., Sheldon, C., Bagnall, D., Peacock, W., and Dennis, E. (2002). *FLC*, a repressor of flowering, is regulated by genes in different inductive pathways. Plant J 29:183-191.

Ruiz-Garcia, L., Madueno, F., Wilkinson, M., Haughn, G., Salinas, J., and Martinez-Zapater, J. (1997). Different roles of flowering-time genes in the activation of floral initiation genes in *Arabidopsis*. Plant Cell 9:1921-1934.

Salk Institute. (2000). Functional genomics and the virtual plant: A blueprint for understanding how plants are built and how to improve them. (NSF Workshop Report, The Arabidopsis Information Resource (TAIR)).

Samach, A., and Coupland, G. (2000). Time measurement and the control of flowering in plants. BioEssays 22:38-47.

Samach, A., and Gover, A. (2001). Photoperiodism: The consistent use of *CONSTANS*. Curr Biol 11:R651-R654.

Samach, A., Onouchi, H., Gold, S., Ditta, S., Schwarz-Sommer, Z., Yanofsky, F., and Coupland, G. (2000). Distinct roles of *CONSTANS* target genes in reproductive development of *Arabidopsis*. Science 288:1613-1616.

Samsonova, M., and Serov, V. (1999). NetWork: An interactive interface to the tools for analysis of genetic network structure and dynamics. In: Proceedings of Pacific Symposium on Biocomputing,4:102-111. World Publishing Co., Singapore.

Sarnowski, T. J., Swiezewski, S., Pawlikowska, K., Kaczanowski, S., and Jerzmanowski, A. (2002). *AtSWI3B*, an *Arabidopsis* homolog of *SWI3*, a core subunit of yeast *Swi/Snf* chromatin remodeling complex, interacts with *FCA*, a regulator of flowering time. Nucleic Acids Research 30:3412-3421.

Schaffer, R., Ramsay, N., Samach, A., Corden, S., Putterill, J., Carre, I., and Coupland, G. (1998). The late elongated hypocotyls mutation of *Arabidopsis* disrupts circadian rhythms and the photoperiodic control of flowering. Cell 93:1219-1229.

Schmid, M., Uhlenhaut, N., Godard, F., Demar, M., Bressan, R., Weigel, D., and Lohmann, J. (2003). Dissection of floral induction pathways using global expression analysis. Devel 130:6001-6012.

Schomburg, F., Patton, D., Meinke, D., and Amasino, R. (2001). *FPA*, a gene involved in floral induction in *Arabidopsis*, encodes a protein containing RNA-recognition motifs. Plant Cell 13:1427–1436.

Sheldon, C., Burn, J., Perez, P., Metzger, J., Edwards, J., Peacock, J., and Dennis, E. (1999). The *FLF MADS* box gene: A repressor of flowering in *Arabidopsis* regulated by vernalization and methylation. Plant Cell 11:445–458.

Sheldon, C., Rouse, D., Finnegan, E., Peacock, W., and Dennis, E. (2000). The molecular basis of vernalization: The central role of *FLOWERING LOCUS C (FLC)*. Proc Natl Acad Sci 97:3753-3758.

Shimamoto, K., and Kyozuka, J. (2002). Rice as a model for comparative genomics of plants. Annu Rev Plant Biol 53:399-419.

Silverstone, A., Ciampaglio, C., and Sun, T. (1998). The Arabidopsis *RGA* gene encodes a transcriptional regulator repressing the gibberellin signal transduction pathway. Plant Cell 10:155-169.

Silverstone, A., Mak, P., Martinez, E., and Sun, T. (1997). The *RGA* locus encodes a negative regulator of gibberellin response in *Arabidopsis thaliana*. Genetics 146:1087-1099.

Simpson, G., and Dean, C. (2002). *Arabidopsis*, the Rosetta stone of flowering time? Science 296:285-289.

Simpson, G., Dijkwel, P., Quesada, V., Henderson, I., and Dean, C. (2003). *FY* is an RNA 3' end-processing factor that interacts with *FCA* to control the Arabidopsis floral transition. Cell 113:777-787.

Simpson, G., Gendall, A., and Dean, C. (1999). When to switch to flowering. Annu Rev Cell Dev Biol 99:519-550.