

Review

Update in bioinformatics Toward a digital database of plant cell signalling networks: advantages, limitations and predictive aspects of the digital model

Marcela Beatriz Treviño Santa Cruz¹, Dominique Genoud, Jean-Pierre Métraux,
Thierry Genoud^{*}

Department of Biology, University of Fribourg, Rte Albert Gockel 3, 1700 Fribourg, Switzerland

Received 5 November 2004; received in revised form 24 November 2004

Abstract

The process of signal integration, which contributes to the regulation of multiple cellular activities, can be described in a digital language by a set of connected digital operations. In this article we delineate the basic concepts of cell signalling in the context of a logical description of information processing. Newly described instances of signal integration in plants are given as examples. The different advantages, limitations and predictive aspects of the digital modeling of signal transduction networks, as well as the minimal architecture of a computer database for plant signalling networks are discussed.

© 2004 Elsevier Ltd. All rights reserved.

Keywords: Cell signalling database; Signal transduction networks; Digital networks; Digital simulation; Cross-talks; Phytochrome

Contents

1. Introduction	268
2. Separating signalling events from metabolic events	268
3. Signal integration: basic concepts	269
4. Signal integration machinery: structure and sites.	271
5. The time factor	272
6. The predictive aspects in logical models.	273

^{*} Corresponding author. Present address: Center for Integrative Genomics (CIG), BEP, CH-1015 Lausanne-Dorigny, Switzerland. Tel.: +41 26 300 8810; fax: +41 26 300 9740.

E-mail address: thierry.genoud@unifr.ch (T. Genoud).

¹ Present address: School of Ocean and Earth Science, Southampton Oceanography Centre, University of Southampton, Room 346/39, European Way, Southampton SO14 3ZH, UK.

7. Conclusion	274
References	274

1. Introduction

During the processing of signals in a cell, multiple intrinsic and extrinsic information sets are combined by cellular processors to elicit the most logical/adapted response for any given conjunction of stimuli. The characterization of signal integration processes emerges as an important issue in biology, since they control gene expression, the activity of proteins and motor functions, the course of development, as well as adjustments of primary and secondary metabolism following the fluctuations of environmental and internal parameters (see for instance [Jordan et al., 2000](#); [Gasch et al., 2000](#); [Cases et al., 2003](#)). Models for the representation of the cellular control systems have been proposed since the early 1960s ([Jacob and Monod, 1961a,b](#); [Kauffman, 1974](#)); they mostly have considered the regulation of genes as the output of “genetic networks”, i.e. as the result of the activation of a causal network of genes responsible for the induction/repression of other genes in a cascade of events. This connectionist view has been applied on processes such as the progression of cell differentiation in animals or the induction of a specific stage during the course of development ([Yuh et al., 1998](#); [Davidson et al., 2002](#)). In this approach one starts from a mathematical standpoint and uses complex data from the high-throughput technologies such as microarrays and large-scale 2-hybrid systems in yeast to infer the minimal structure of a genetic network ([Somogyi and Sniegoski, 1996](#); [D’haeseleer et al., 1998, 2000](#); [Thieffry, 1999](#); [Guelzim et al., 2002](#); [Shen-Orr et al., 2002](#); [Alon, 2003](#); [Bray, 2003](#); [Wuchty et al., 2003](#)). On the other hand, the creation of dynamic models allowing computer representation and analysis has been the object of multiple theoretical studies (reviewed in [De Jong, 2002](#)). Differential equations and models such as Petri networks and Bayesian networks have been applied to complex biological networks of interacting cell components, with some exciting achievements (e.g. the lambda-phage activation cycle, [McAdams and Shapiro, 1995](#); the MAP kinase cascade, [Bhalla et al., 2002](#)). Globally, the modeling methods can be divided into two main groups; one providing a qualitative description, the other capitalizing on statistical variations and/or random factors embedded in cell signalling operators. These operators are described in the first case as simple processors adding or associating unitary signals to

create a new signal (connected in logical/Boolean networks; [Thomas, 1973](#)); and in the second case, as processors enclosing continuous or probabilistic values, yet equally rooted on the same logical rules of association (connected in statistical Boolean networks, Bayesian networks, or sets of differential equations; reviewed in [De Jong, 2002](#); [Bolouri and Davidson, 2002](#); discussed in [Shmulevich et al., 2002](#)).

A graphic digital model has been proposed for the description of signal transduction networks in plants ([Genoud and Métraux, 1999](#); [Zhang and Shapiro, 2002](#); [Heck et al., 2003](#); [Thum et al., 2003](#)). This model can be used to create a common formalism for a plant signalling database to be included in a database ontology, [such as the EcoCyc ontology ([Karp et al., 2002](#)), Rzhetsky’s ontology ([Rzhetsky et al., 2000](#)), the Biological Process Ontology ([The Gene Ontology Consortium, 2001](#)), Bhalla’s ontology ([Bhalla, 2002](#)), the BioPax ontology (<http://www.biopax.org>)], and expressed in a format for data exchange (e.g. SBML, Systems Biology Markup Language). It is also very suitable for the qualitative simulation of plant responses using computers. In this article we describe some basic features of biological signal processing that may serve for the identification of signalling proteins based on their sequence, structure, and biochemical properties. We also discuss the limitations of a digital model versus more quantitative models to represent and simulate the basic properties of plant signal processing, and propose the creation of a signalling database for plant biologists using a digital paradigm. Such a tool will be very helpful in plant signalling research, where results obtained from studies in different plant species and under a wide range of developmental and experimental conditions often make data interpretation challenging. By using digital models and including all relevant information in a comprehensive database, the use of available information could be optimized, so that a clearer and more complete picture of cell signalling events may emerge.

2. Separating signalling events from metabolic events

The idea of a ‘signal’ is central in the science of information, and therefore in any model of cell information processing. This notion contains a priori meanings that deserve to be shortly recapitulated here, since they bias

the definition of signal transduction activity, and therefore the representation of such an activity in a specific database. First, the notion of a signal tacitly separates the activity of perception from the activity of cyclic and constant modifications of the basic metabolism (the metabolome). The notion of a signal also contains the idea of a unit of induction; e.g. a defined quantity of stimulus generating a detectable physiological and/or molecular response, the significance of which is statistically determined. In the frame of a digital model of signalling, one statistically significant signal leads to a response with an assigned value of one (1). For a complete description of cell signalling, this basic definition of signal could be inadequate in several cases. For instance, some metabolic events also possess signalling properties (e.g. the function of hexokinase in sugar sensing, Jang et al., 1997; the role of aconitase as an iron sensor, Navarre et al., 2000; the acetyl- γ -glutamyl phosphate reductase Arg5,6, acting as a transcription factor; Hall et al., 2004); and in a systemic view any event connected to a regulatory loop might be regarded as a potential signal emitter. In addition, any change in the threshold of significance assigned to a biological response affects the level of the corresponding input signals. This may produce qualitative changes in the digital model and will require modifications in the details of such a representation. The construction of a large database of cellular signalling and responses must therefore include a precise description of the thresholds of significance (statistical values) and of all experimental conditions (see Fig. 1). Moreover, since the notion of signal implies observable biological changes, any exchange of information among processes controlling the cellular homeostasis and leading to no apparent changes in metabolite concentration or gene expression, would escape the definition of signalling and thus would be improperly neglected. Therefore, access to metabolic databases and to the related genetic data obtained from specific mutants in the metabolic pathways should be included in order to compensate for the missing information (Fig. 1).

3. Signal integration: basic concepts

Signal integration can be understood as a process by which two or more signals interact with a biochemical component and thereby create a new “output” signal (Monod and Jacob, 1961; Sugita, 1961, 1975; Walter et al., 1967; Kauffman, 1974; Prehoda and Lim, 2002). The operation of such a process requires a minimal set of stable and characteristic constraints. The concept of a molecular operator integrating several signals is not trivial, as it presupposes complex molecular structures capable of carrying out sophisticated biochemical functions (see for instance the regulation of genes during sea

urchin development: Arnone and Davidson, 1997; Yuh et al., 1998, 2001; the N-WASP function: Prehoda and Lim, 2002; Dueber et al., 2003). It is useful to shortly re-examine the basic necessary properties for the integration of cellular signals by a biochemical component in the light of a graphic digital representation and of recent biochemical discoveries.

The biochemical machinery of signal integration is experimentally robust (i.e. statistically reproducible), and must include a minimal number of stable rules of signal association. The way in which such combination rules may be embodied in a molecular signalling element has been the object of recent experimental investigations (Russo et al., 1996; Prehoda et al., 2000; Tarricone et al., 2001; Dueber et al., 2003). From the point of view of a basic theoretical approach, rules of signal transduction and/or integration might be compared to the concept of a translation system. Such a molecular system contains two main features: a recognition and memory ability of the input signal(s) to be translated (located in the input domain), and a mechanism for generating a response, or output signal, (which is the result of translation), e.g. a connection toward secondary components, such as the next element in a signalling network. For instance, the key-lock recognition of a substrate by an enzyme corresponds to the input perception step that implies a stereochemical memory of the substrate. In contrast, the catalytic activity of the enzyme generates the output of translation. Both aspects constitute the operation of this molecular translator. In general, the different features of a translation rule can be described in a qualitative manner. For example, the biochemical phenomenon of molecular recognition, which implies a molecular memory, relates to the specific and distinctive qualitative characteristics of the interacting partners. Thus a molecular interaction can be represented within a qualitative model, and the quantitative features of the interaction may be described discontinuously by several qualitative steps. The mechanism, as well as the output of a molecular operation, can also be digitalized. In fact, similarly to the arithmetical rules of addition and multiplication, it is possible to represent in a qualitative manner any mechanism associating two or more qualities to produce one or more new qualities (as already shown by Boole, 1854; Von Neumann, 1951). In a dynamic simulation program, this implies however that the activation of an operator located in a cascade of integration occurs, for instance, through variation of the input frequency, a feature that can be easily implemented in current digital simulation programs (Genoud et al., 2001, 2003).

Biochemical functions can be qualitatively described by the statistically most relevant steps involved in biochemical reactions such as hydrolysis, addition of residues, reduction, oxidation, allosteric modification of structures and activity, etc. Indeed, such transformations

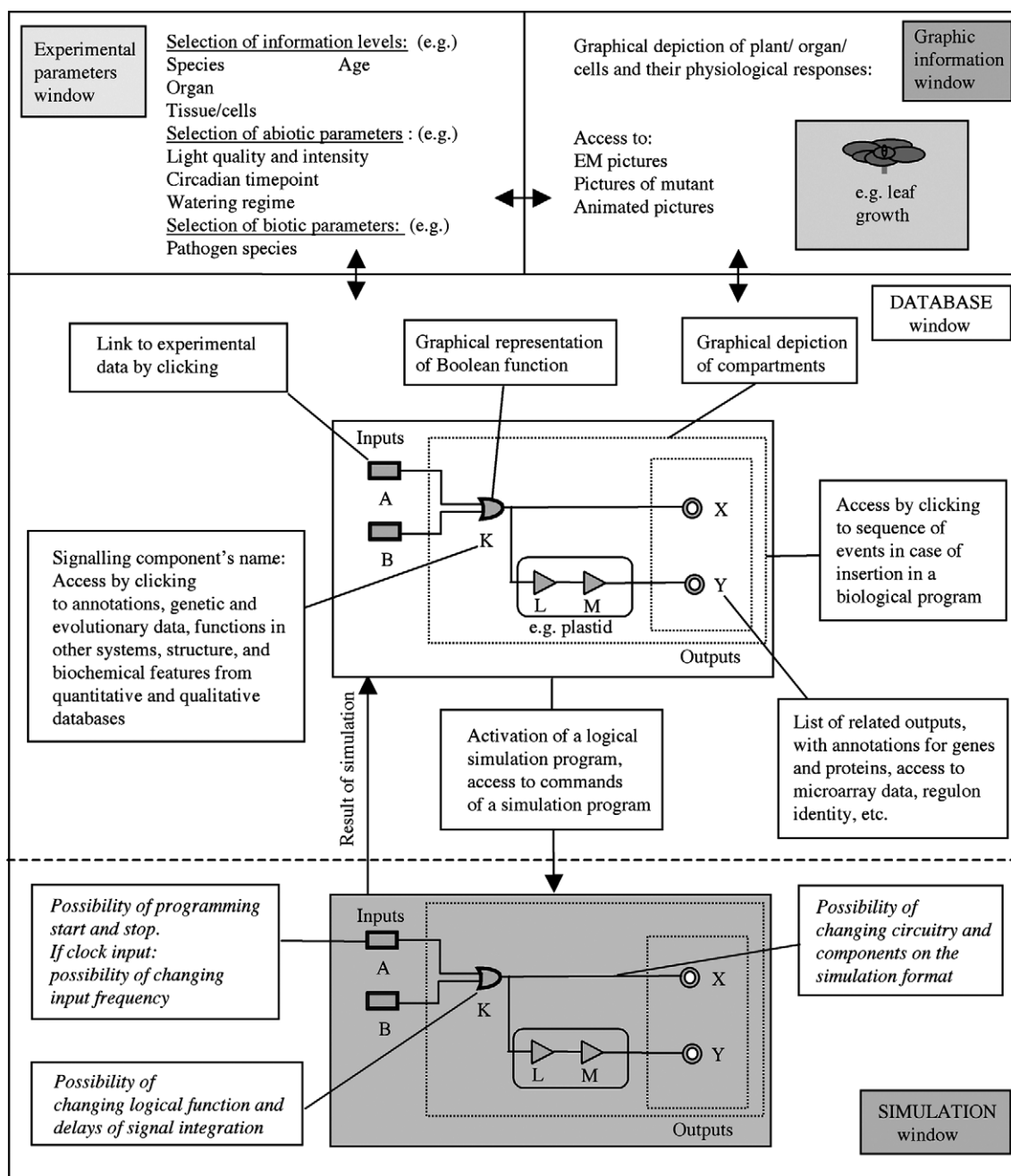


Fig. 1. Schematic representation of the architecture of a cell signalling database that includes a digital simulation program. The information that would be accessible by clicking on the different objects, and the links to other databases are described in text boxes. Arrows represent transitions between windows and computer applications. Text in *italics* corresponds to the options included in a digital simulation program such as DigSim. A,B: input elements (e.g. receptors); K,L,M: signalling elements (e.g. kinases).

imply either alterations in spatial qualities (tri-dimensional structure) or qualitative changes (variation of chemical properties). Qualitative models are therefore amenable for the representation of large networks of biochemical information processing. Digital models have been shown to be suitable for the description of cell signalling processes (Genoud and Métraux, 1999; Genoud et al., 2001; Zhang and Shapiro, 2002; Thum et al., 2003). Such models use defined categories for

the input and output signals and unitary association rules that cover any possible logical combination, and that can be directly applied in a logical simulation program (for instance inside the Matlab's Simulink application). However, the fine details of enzyme kinetics and protein stability, or those of the relative affinity for specific interacting factors, would need to be simplified into their statistically most significant components. Hence, a qualitative digital model would be appropriately

completed by quantitative parameters measured for the individual components of signal transduction networks, e.g. the *in vitro* measurements of enzyme kinetics. This additional information should be directly accessible from the digital model to give researchers the possibility of designing further experiments to target particular aspects of the cell signalling component (see Fig. 1).

For the digital representation, gene expression can be described as an all or none response and thus translated into an ON/OFF state. However, if needed, a description through categories of expression levels can also be used. This is a simple way to transcribe quantitative effects into a digital language. While the definition of distinct levels of expression as separate qualities in the output domain of a signalling network (i.e. categories such as 'low', 'medium', 'high', and 'very high' levels of gene expression) will increase the accuracy of a qualitative model, it could still not provide the level of precision of a quantitative description that uses, for instance, a continuous model of differential equations (Lee et al., 2003).

An additional difficulty may arise with the interpretation of gene expression levels, since feedback of gene products on their own expression (or autocatalytic effects) might blur the interpretation of signalling circuits. This might be critical for the inference of genetic networks from microarray data, since different levels of a gene's expression may correspond to the output of distinct circuits in a signalling network.

It is likely that quantitative data characterizing the behavior of each component controlling gene expression in a higher organism will not be available before several decades (Endy and Brent, 2001), and a fully quantitative model of signalling networks might not be realizable without a huge computer facility (Kitano, 2002). Until these data and technologies are available, a qualitative model seems appropriate.

4. Signal integration machinery: structure and sites

Theoretically, there is a stoichiometric proportion linking the input signals to the activity of a molecular operator. In fact, the first characterized signal integrators in biological systems associate two signals to a response event with a one-to-one stoichiometry (see for instance Prehoda and Lim, 2002; Mangan and Alon, 2003). Biochemical experiments have shown that the molecular machinery required to perform such a simple association is already highly complex. Hence, the simplest and perhaps energetically most economical way for a living organism to integrate biochemical information, may be through discrete biochemical associations. However, since most experimental results are derived from the observation of numerous signalling molecules, the data always appear as a statistical sum of unsyn-

chronized signalling activity, and therefore as continuous quantitative data. The digital model is not able to represent such a statistical property and therefore does not reflect the nonsynchronization of, for instance, an equal response in different cells of a tissue, or the cellular noise that produces randomness in gene expression levels (Thattai and van Oudenaarden, 2001; Elowitz et al., 2002). However, transcriptional noise can be evaluated experimentally (Blake et al., 2003), and simulated using a generator of variable random values coupled to the data inputs (Fig. 2(a)).

What are the molecular characteristics of the cell information processors? It seems obvious that the simplest molecular machines are composed of one or several complex molecules such as proteins, nucleic acids, or a combination of both. Fundamentally there are two main alternatives to operate signal integration:

- 1) The integration is simple: one logical operator per molecule, or per complex of molecules. This is the case of proteins such as N-WASP and Cdk2 (Prehoda and Lim, 2002), the CDK inhibitor (Nash et al., 2001), and the neural coincidence detectors (Bourne and Nicoll, 1993).
- 2) The integration is multiple: several connected operators per molecule or complex of molecules. This situation is typically found in the case of promoters containing several *cis*-regulatory domains (Arnone and Davidson, 1997; Yuh et al., 2001).

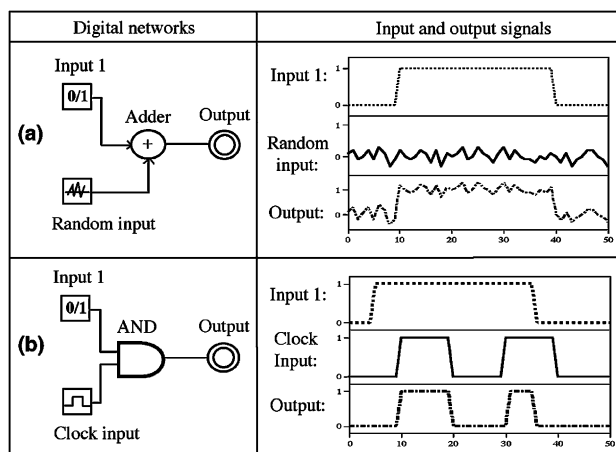


Fig. 2. Implementation of noise and clock signals on a digital simulator. The right panels contain the time signatures of the input and output signals represented on the left panels, using an arbitrary time scale (x-axis) for the signal value (y-axis). (a) In this example, a digital signal has been associated to a source of random signal through an adder in order to simulate biological noise; the amplitude of the random signal can be adjusted to simulate measured signalling background. (b) The activation of an oscillating clock-like signal has been placed under the control of a simple ON/OFF switch (input 1) through an AND operator. (Similarly, several signals may be connected to a central clock in order to simulate circadian rhythms).

The site of signal integration can be found at diverse locations throughout a signalling network. In fact two pathways may interact several times at different levels, such as at the level of hormone production, of histone modifications that mediate changes in chromatin configuration, of promoter activity, etc. To illustrate this fact it is useful to consider examples of signal integration recently described in plants.

For instance, evidence suggests that multiple connections exist between the phytohormone auxin and the phytochrome-mediated regulatory networks. O'Grady et al. (2001) identified in soybean plants a novel member of the GT-2 family of transcription factors, namely GmGT-2, which unlike other family members, was shown to be transcriptionally down-regulated by light in a phytochrome-dependent manner. As with GmGT-2, RNA-level analysis indicated that transcription of *Aux28*, a member of the *Aux/IAA* family of auxin-responsive genes encoding short-lived transcription factors, is also negatively regulated by phytochrome. Using electrophoretic mobility shift assays, GmGT-2 was demonstrated to bind to well-mapped protein binding sites in the promoter region of *Aux28*, suggesting that the transcriptional control of this *Aux/IAA* gene by phytochrome may be mediated by transcription factors such as GmGT-2. On the other hand, the promoter region of *Aux28* also contains elements similar to those responsible for the auxin-induced expression of the *GH3* gene, mediated by another class of transcription factors involved in auxin responsiveness, known as ARFs. Both *Aux/IAA* proteins and ARFs contain conserved domains that enable their intra- and interfamily dimerization (Kim et al., 1997; Ulmasov et al., 1999). As indicated by O'Grady et al. (2001), it remains to be determined whether the observed phytochrome-mediated regulation of the *Aux28* gene occurs through a reduction in auxin levels, through the down-regulation of positive activators such as GmGT-2 and ARFs, or both. At least for some *Arabidopsis* and pea *Aux/IAA* recombinant proteins (e.g. SHY2/IAA3 and Ps-IAA4), their interaction with and phosphorylation by recombinant phytochrome A (phyA) from oat has been shown in vitro (Colon-Carmona et al., 2000). This group of researchers also demonstrated increased in vivo steady-state levels of mutant IAA3 in *shy2-2* gain-of-function mutant plants, and in vivo phosphorylation of the SHY2-2 protein, leading them to propose the phytochrome-dependent phosphorylation of *Aux/IAA* proteins as a molecular mechanism for integration of light and auxin signalling.

A different type of cross-talk juncture between these two signalling pathways has been described in *Arabidopsis* by Hsieh et al. (2000), involving FIN219, a novel phyA signalling component. *fin219* was identified as a suppressor mutation of *cop1-6* during a screen for genes involved in the light-induced inactivation of COP1.

COP1 is a key repressor of photomorphogenic development known to be primarily controlled by phyA, phyB, and cryptochromes, and it is thought to negatively regulate transcription factors involved in light-regulated gene expression, presumably by targeting them for degradation by the 26S proteasome (Osterlund et al., 2000). FIN219 shows homology to the GH3 protein family, and as is the case for this protein family of yet unknown function, its expression is rapidly induced by auxin. Since the auxin-regulated processes in the *fin219* mutant are not altered, it is unlikely for FIN219 to be an auxin-signalling component. Therefore, this may be an example of cross-talk in which auxin regulates the expression level of a light signalling component. Taken together, these and other lines of evidence suggest the existence of complex and plastic interactions between the light- and auxin-mediated regulatory networks.

Yet a different type of regulation exists where one operator modulates two distinct pathways. For example, a recent study in *Arabidopsis* reveals the regulation of ABA- and phyA-mediated germination responses to be affected by a common modulator (Duque and Chua, 2003). This seed imbibition-inducible modulator, IMB1, is a member of the BET subgroup of bromodomain-containing proteins, a class of putative transcriptional regulators whose mode of action is thought to involve their association with acetylated histones (Florence and Faller, 2001). IMB1 appears to be specific to the promotion of seed germination, and it negatively and positively regulates the ABA and phyA transduction pathways, respectively. This is an interesting case from the emerging field of chromatin remodeling.

5. The time factor

The concentration of a protein/nucleic acid(s) processor can fluctuate in time and this in turn might alter the characteristics of the corresponding logical element in the model. Moreover, if such variation in concentration induces the emergence of new responses, an additional digital element must be defined and associated with that particular protein/nucleic acid(s) processor. These are inconveniences of a dynamic digital representation that need to be solved before the construction of a database.

In fact, continuous systems of oscillation are found at organismic and cellular levels (Goldbeter, 2002; Eriksson and Millar, 2003). They are composed of series of biochemical events subjected to transient negative feedback loops that may interrupt the expression of a particular gene responsible for cycling processes. A digital description of complex organized cellular events must consider the input of the cycling processes as a clock-like input. The cellular equivalent of a digital clock, which can define the tempo (or phase) of the diverse signal integration cascades, must be a chemical oscillation sys-

tem consisting of rapid fluctuations in the level of a small diffusable cellular component. Calcium ions and protons are possibly involved in such pacemaker functions, as they are known to modulate countless biochemical processes, including the regulation of enzymatic activities. For example, the concentration of calcium ions could influence the activity of several signalling proteins, either directly by steric modification (of, for instance, calcium-dependent protein kinases) or indirectly through the mediation of calcium-binding proteins such as calmodulins, calcitonins, annexins, phosphatases, etc., which in turn modulate other proteins involved in signalling (Johnson et al., 1995; Trewavas and Malho, 1998). Rapid diffusion waves of a chemical species such as Ca^{2+} may accompany the transduction of signals, or may be constitutively pulsing in the cytoplasm to coordinate the phase aspect of signal transduction (Guo et al., 2002; Nishida et al., 2003). The formation of additional and/or new organizing waves at the time of signal perception may reinforce or reduce the basic pacemaker process in order to emphasize important incoming information. Therefore, very accurate modeling of cell signals would require the addition of one or several clock oscillators for the synchronization of the signalling events. This is easily implemented in digital simulators that contain clock-like input sources (see Fig. 2(b)). These elements can be connected to any signalling operations in order to generate an oscillating output upon activation by an input signal.

In fact, the digital network's language allows dissection of cycling loops such as the circadian oscillation of gene expression. Dynamic representation and adaptation of input settings from a specific computer window will permit the construction of accurate models and a useful analysis of cycling cellular events (such as for example mitosis).

6. The predictive aspects in logical models

The representation of signalling networks using digital operators has the power of a precise language. However, digital networks will always contain some simplifications and assumptions, since this knowledge is rarely/never complete. In fact, such gaps provide the predictive quality of the qualitative model: any knowledge in the absence of the complete details represents a series of postulates or expectations. The apparent determinism of the sets of logical operations represented in digital models has in fact a probabilistic nature, and includes uncertainty at every level. Moreover, inherent variations in experimental measurements produce noise, which add uncertainty to the model. However, because the predictive power of a digital model can be a matter of controversy, and because prediction is central to science, this point deserves the following comment.

The simplest case of a postulate is the statement that a certain result, for instance an experimental observation, is reproducible in a given context, i.e. that it will occur in the same given way as long the prescribed conditions are met. This is the property of any model or knowledge database. Two different cases of postulate arise when: (1) two sets of data containing the same output response are superimposed and connected, and (2) when a single data set is obtained by the effect of 2 stimuli acting on a system simultaneously, but not when acting in an alternate manner. In these last two cases a prediction may result in fact from the insertion of putative logical operators into a synthetic model. This is illustrated in Fig. 3 with an example of the first case (superimposition of data), where two stimuli (a and b) have been applied independently to the same system, and both give the response X. Joining inputs a and b to output X in a model (for instance using an OR operator), even though the effects of a and b when applied in conjunction are not known, represents a prediction. This can subsequently be tested by an experiment, where stimuli a and b are applied together to possibly produce a second response different from

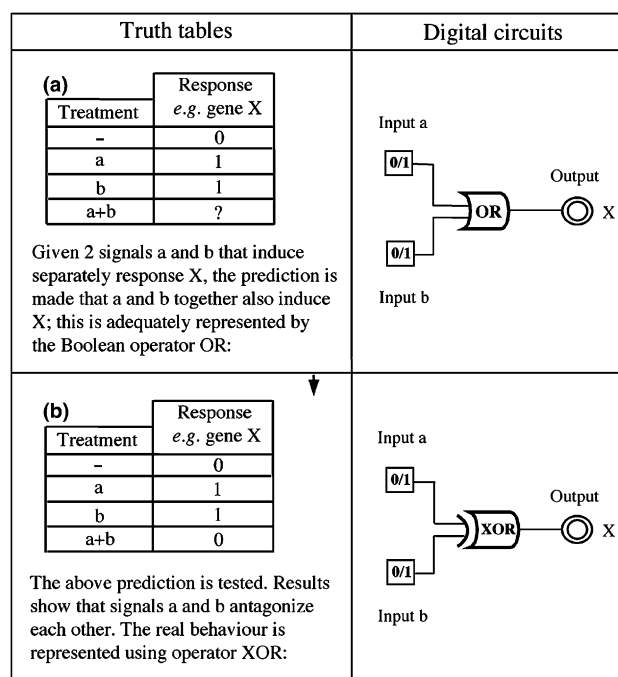


Fig. 3. The predictive aspect in a digital circuit is a function of the available data. In panel (a), the qualitative information (1 = ON, 0 = OFF) corresponding to the effect of two different input signals (a,b) is incomplete, the result of a simultaneous treatment with both a and b being unknown. (b) A combinatorial set of data provides the complete qualitative information to design digital circuits. The result of the effect of a + b is known, the postulated Boolean element can be replaced by a different operator (XOR). The circuit has now a higher predictive power for the effect of the two possible inputs: it can fully simulate the observed output in function of any combination of inputs a and b.

the one postulated. In the case represented in Fig. 3, the OR operator linking a and b, must be changed into an XOR operator. Thus, an interesting feature of digital logic dwells in its flexibility and versatility: it provides the possibility of substituting basic operators such as NOT, AND, OR by a combination of NOR, and/or NAND gates, or of replacing any operator with the same given number of inputs by an operator of different logic without changing the complete network. Usually, the predictive/speculative aspects of a digital system correspond to specific confined structures in a network (the results of additional experiments will only modify part of the network), and providing the possibility of confirmation and improvement by further targeted experiments (Ideker et al., 2000; Bolouri and Davidson, 2002).

7. Conclusion

In this article we discuss some basic concepts of signal integration in the context of cell biology, showing that a very precise description of experimental results can be provided by a digital computing framework. This digital model appears presently as a conveniently accessible ontological formalism for the representation of signalling networks. It offers an easily readable overview, it can be readily improved through alterations made to fit new knowledge, and it can also take into account kinetic and quantitative aspects of signalling networks. This model could therefore be amenable for the general framing of a signalling database. Such a database should include easy access to all available quantitative information, such as genetic data (e.g. Mendelian characters of a mutation), experimental parameters, topological data (e.g. protein localization), biochemical characteristics (in vitro measurement of kinetics, and in vivo concentration of signalling components), etc., in order to provide plant biologists with the best tools for signal integration analyses.

References

- Alon, U., 2003. Biological networks: the thinkerer as an engineer. *Science* 301, 1866–1867.
- Arnone, M.I., Davidson, E.H., 1997. The hardwiring of development: organization and function of genomic regulatory systems. *Development* 124, 1851–1864.
- Bhalla, U.S., 2002. The chemical organization of signalling interactions. *Bioinformatics* 17, 829–837.
- Bhalla, U.S., Ram, P.T., Iyengar, R., 2002. MAP kinase phosphatase as a locus of flexibility in a mitogen-activated protein kinase signalling network. *Science* 297, 1018–1023.
- Blake, W.J., Kaern, M., Cantor, C.R., Collins, J.J., 2003. Noise in eucaryotic gene expression. *Nature* 422, 633–637.
- Bolouri, H., Davidson, E.H., 2002. Modeling transcriptional regulatory networks. *BioEssays* 24, 1118–1129.
- Boole, G., 1854. An Investigation of the Laws of Thought on which are Founded the Mathematical Theories of Logic and Probabilities. MacMillan and Co., Cambridge.
- Bourne, H.R., Nicoll, R., 1993. Molecular machines integrate coincident synaptic signals. *Cell* 72, 65–75.
- Bray, D., 2003. Molecular networks: the top-down view. *Nature* 301, 1864–1865.
- Cases, I., de Lorenzo, V., Ouzounis, C.A., 2003. Transcription regulation and environmental adaptation in bacteria. *Trends Microbiol.* 11, 248–253.
- Colon-Carmona, A., Chen, D.L., Yeh, K.C., Abel, S., 2000. Aux/IAA proteins are phosphorylated by phytochrome in vitro. *Plant Physiol.* 124, 1728–1738.
- Davidson, E.H., Rast, J.P., Oliveri, P., Ransick, A., Caestani, C., Yuh, C.-H., Minokawa, T., Amore, G., Veronica, H., Arenas-Mena, C., Otim, O., Brown, C.T., Livi, C.B., Lee, P.X., Revilla, R., Rust, A.G., Pan, Z.J., Schilstra, M.J., Clarke, P.J.C., Arnone, M.I., Rower, L., Cameron, R.A., McClay, D.R., Hood, L., Bolouri, H., 2002. A genomic regulatory network for development. *Science* 295, 1669–1678.
- De Jong, H., 2002. Modeling and simulation of genetic regulatory systems: a literature review. *J. Comput. Biol.* 9, 67–103.
- D'haeseleer, P., Liang, S., Somogyi, R., 2000. Genetic network inference: from co-expression clustering to reverse engineering. *Bioinformatics* 16, 707–726.
- D'haeseleer, P., Wen, X., Fuhrman, S., Somogyi, R., 1998. Mining the gene expression matrix: inferring gene relationships from large-scale gene expression data. In: Paton, R.C., Holcombe, M. (Eds.), *Information Processing in Cells and Tissues*. Plenum Press, New York, pp. 203–212.
- Dueber, J.E., Yeh, B.J., Chak, K., Lim, W.A., 2003. Reprogramming control of an allosteric signalling switch through modular recombination. *Science* 301, 1904–1908.
- Duque, P., Chua, N.-H., 2003. IMB1, a bromodomain protein induced during seed imbibition, regulates ABA- and phyA-mediated responses of germination in *Arabidopsis*. *Plant J.* 35, 787–799.
- Elowitz, M.B., Levine, A.J., Siggia, E.D., Swain, P.S., 2002. Stochastic gene expression in a single cell. *Science* 297, 1183–1186.
- Endy, D., Brent, R., 2001. Modelling cellular behaviour. *Nature* 409, 391–395.
- Eriksson, M.E., Millar, A.J., 2003. The circadian clock. A plant's best friend in a spinning world. *Plant Physiol.* 132, 732–738.
- Florence, B., Faller, D.V., 2001. You BET-cha: a novel family of transcriptional regulators. *Front. Biosci.* 6, d1008–d1018.
- Gasch, A.P., Spellman, P.T., Kao, C.M., Carmel-Harel, O., Eisen, M.B., Storz, G., Botstein, D., Brown, P.O., 2000. Genomic expression programs in the response of yeast cells to environmental changes. *Mol. Biol. Cell* 11, 4241–4257.
- Genoud, T., Métraux, J.-P., 1999. Crosstalks in plant cell signalling: structure and function of the genetic network. *Trends Plant Sci.* 4, 503–507.
- Genoud, T., Treviño Santa Cruz, M.B., Métraux, J.-P., 2001. Numeric simulation of plant signalling networks. *Plant Physiol.* 126, 1430–1437.
- Genoud, T., Treviño Santa Cruz, M.B., Métraux, J.-P., 2003. Networks of cellular information processing: digital description and simulation. *Curr. Genomics* 4, 27–36.
- Goldbeter, A., 2002. Computational approaches to cellular rhythms. *Nature* 420, 238–245.
- Guelzim, N., Bottani, S., Bourguin, P., Képès, F., 2002. Topological and causal structure of the yeast transcriptional regulatory network. *Nat. Genet.* 31, 60–63.
- Guo, Y., Xiong, L., Song, C.P., Gong, D., Halfter, U., Zhu, J.K., 2002. A calcium sensor and its interacting protein kinase are global regulators of abscisic acid signalling in *Arabidopsis*. *Dev. Cell.* 3, 233–244.

- Hall, D.A., Zhu, H., Zhu, X., Royce, T., Gerstein, M., Snyder, M., 2004. Regulation of gene expression by a metabolic enzyme. *Science* 306, 482–484.
- Heck, S., Grau, T., Buchala, A., Métraux, J.-P., Nawrath, C., 2003. Genetic evidence that expression of NahG modifies defence pathways independent of salicylic acid biosynthesis in the *Arabidopsis-Pseudomonas syringae* pv. *tomato* interaction. *Plant J.* 36, 342–352.
- Hsieh, H.L., Okamoto, H., Wang, M., Ang, L.H., Matsui, M., Goodman, H., Deng, X.W., 2000. FIN219, an auxin-regulated gene, defines a link between phytochrome A and the downstream regulator COP1 in light control of *Arabidopsis* development. *Genes Dev.* 14, 1958–1970.
- Ideker, T.E., Thorsson, V., Karp, R.M., 2000. Discovery of regulatory interactions through perturbation: inference and experimental design. In: Altman, R.B., Lauderdale, K., Dunker, A.K., Hunter, L., Klein, T.E. (Eds.), *Proc. Pac. Symp. Biocomput. (PSB 2000)*, vol. 5. World Scientific Publishing, Singapore, pp. 302–313.
- Jacob, F., Monod, J., 1961a. Genetic regulatory mechanisms in the synthesis of proteins. *J. Mol. Biol.* 3, 318–356.
- Jacob, F., Monod, J., 1961b. On the regulation of gene activity. *Cold Spring Harb. Symp. Quant. Biol.* 26, 193–211.
- Jang, J.C., Leon, P., Zhou, L., Sheen, J., 1997. Hexokinase as a sugar sensor in higher plants. *Plant Cell* 9, 5–19.
- Johnson, C.H., Knight, M.R., Kondo, T., Masson, P., Sedbrook, J., Haley, A., Trewavas, A., 1995. Circadian oscillations of cytosolic and chloroplastic free calcium in plants. *Science* 269, 1863–1865.
- Jordan, J.D., Landau, E.M., Iyengar, R., 2000. Signalling networks: the origins of cellular multitasking. *Cell* 103, 193–200.
- Karp, P.D., Riley, M., Saier, M., Paulser, I.T., Collado-Vides, J., Paley, S.M., Pellegrini-Toole, A., Bonavides, C., Gama-Castro, S., 2002. The EcoCyc Database. *Nucleic Acids Res.* 30, 56–58.
- Kauffman, S.A., 1974. The large-scale structure and dynamics of gene control circuits: an ensemble approach. *J. Theor. Biol.* 44, 167–190.
- Kim, J., Harter, K., Theologis, A., 1997. Protein–protein interactions among the Aux/IAA proteins. *Proc. Natl. Acad. Sci. USA* 94, 11786–11791.
- Kitano, H., 2002. Computational systems biology. *Nature* 420, 206–210.
- Lee, E., Salic, A., Krüger, R., Heinrich, R., Kirschner, M.W., 2003. The roles of APC and axin derived from experimental and theoretical analysis of the Wnt pathway. *PLoS Biol.* 1, 116–132.
- Mangan, S., Alon, U., 2003. Structure and function of the feed-forward loop network motif. *Proc. Natl. Acad. Sci. USA* 100, 11980–11985.
- McAdams, H.H., Shapiro, L., 1995. Circuit simulation of genetic networks. *Science* 269, 650–656.
- Monod, J., Jacob, F., 1961. General conclusions: teleonomic mechanisms in cellular metabolism, growth, and differentiation. *Cold Spring Harbor Symp. Quant. Biol.* 26, 389–401.
- Nash, P., Tang, X., Orlicky, S., Chen, Q., Gertler, F.B., Mendenhall, M.D., Sicheri, F., Pawson, T., Tyers, M., 2001. Multisite phosphorylation of a CDK inhibitor sets a threshold for the onset of DNA replication. *Nature* 414, 514–521.
- Navarre, D.A., Wendehenne, D., Durner, J., Noad, R., Klessig, D.F., 2000. Nitric oxide modulates the activity of tobacco aconitase. *Plant Physiol.* 122, 573–582.
- Nishida, M., Sugimoto, K., Hara, Y., Mori, E., Morii, T., Kurosaki, T., Mori, Y., 2003. Amplification of receptor signalling by Ca²⁺ entry-mediated translocation and activation of PLCgamma2 in B lymphocytes. *EMBO J.* 22, 4677–4688.
- O'Grady, K., Goekjian, V.H., Naim, C.J., Nagao, R.T., Key, J.L., 2001. The transcript abundance of GmGT-2, a new member of the GT-2 family of transcription factors from soybean, is down-regulated by light in a phytochrome-dependent manner. *Plant Mol. Biol.* 47, 367–378.
- Osterlund, M.T., Hardtke, C., Wei, N., Deng, X.W., 2000. Targeted destabilization of HY5 during light regulated development of *Arabidopsis*. *Nature* 405, 462–466.
- Prehoda, K.E., Scott, J.A., Dyché Mullins, R., Lim, W.A., 2000. Integration of multiple signals through cooperative regulation of the N-WASP-Arp2/3 complex. *Science* 290, 801–806.
- Prehoda, K.E., Lim, W.A., 2002. How signalling proteins integrate multiple inputs: a comparison of N-WASP and Cdk2. *Curr. Opin. Cell Biol.* 14, 149–154.
- Rzhetsky, A., Koike, T., Kalachikov, S., Gomez, S.M., Krauthammer, M., Kaplan, S.H., Kra, P., Russo, J.J., Friedman, C., 2000. A knowledge model for analysis and simulation of regulatory networks. *Bioinformatics* 16, 1120–1128.
- Russo, A.A., Jeffrey, P.D., Pavletich, N.P., 1996. Structural basis of cyclin dependent kinase activation by phosphorylation. *Nat. Struct. Biol.* 3, 696–700.
- Shen-Orr, S., Milo, R., Mangan, S., Alon, U., 2002. Network motifs in the transcriptional regulation network of *Escherichia coli*. *Nat. Genet.* 31, 64–68.
- Shmulevich, I., Daugherty, E.R., Zhang, W., 2002. From Boolean to probabilistic Boolean networks as models of genetic regulatory networks. *Proc. IEEE* 90, 1778–1792.
- Somogyi, R., Sniegoski, C.A., 1996. Modeling the complexity of genetic networks: understanding multigenic and pleiotropic regulation. *Complexity* 1, 45–63.
- Sugita, M., 1961. Functional analysis of chemical systems in vivo using a logical circuit equivalent. *J. Theor. Biol.* 1, 415–430.
- Sugita, M., 1975. Functional analysis of chemical systems in vivo using a logical circuit equivalent. V. Molecular biological interpretation of the self-reproducing automata theory and chemico-physical interpretation of information in biological systems. *J. Theor. Biol.* 53, 223–237.
- Tarricone, C., Xiao, B., Justin, N., Walker, P.A., Rittinger, K., Gamblin, S.J., Smerdon, S.J., 2001. The structural basis of Arfap1in-mediated cross-talk between Rac and Arf signalling pathways. *Nature* 411, 215–219.
- Thattai, M., van Oudenaarden, A., 2001. Intrinsic noise in gene regulatory networks. *Proc. Natl. Acad. Sci. USA* 98, 8614–8619.
- The Gene Ontology Consortium, 2001. Creating the gene ontology resource: design and implementation. *Genome Res.* 11, 1425–1433.
- Thieffry, D., 1999. From global expression data to gene networks. *BioEssays* 21, 895–899.
- Thomas, R., 1973. Boolean formalization of genetic control circuits. *J. Theor. Biol.* 42, 563–585.
- Thum, K.E., Shasha, D.E., Lejay, L.V., Coruzzi, G.M., 2003. Light- and carbon-signalling pathways. Modeling circuits of interactions. *Plant Physiol.* 132, 440–452.
- Trewavas, A.J., Malho, R., 1998. Ca²⁺ signalling in plant cells: the big network!. *Curr. Opin. Plant Biol.* 1, 428–433.
- Ulmasov, T., Hagan, G., Guilfoyle, T.J., 1999. Activation and repression of transcription by auxin-response factors. *Proc. Natl. Acad. Sci. USA* 96, 5844–5849.
- Von Neumann, J., 1951. The general and logical theory of automata. In: Jeffries, L.A. (Ed.), *Cerebral Mechanisms in Behavior*. John Wiley and Sons, New York.
- Walter, C., Parker, R., Ycas, M., 1967. A model for binary logic in biochemical systems. *J. Theor. Biol.* 15, 208–217.
- Wuchty, S., Oltai, Z.N., Barabasi, A.L., 2003. Evolutionary conservation of motif constituents in the yeast protein interaction network. *Nat. Genet.* 35, 176–179.
- Yuh, C.-H., Bolouri, H., Davidson, E.H., 1998. Genomic *cis*-regulatory logic: experimental and computational analysis of a sea urchin gene. *Science* 279, 1896–1902.
- Yuh, C.-H., Bolouri, H., Davidson, E.H., 2001. *Cis*-regulatory logic in the *endo16* gene: switching from a specification to a differentiation mode of control. *Development* 128, 617–628.
- Zhang, C., Shapiro, A.D., 2002. Two pathways act in an additive rather than obligatorily synergistic fashion to induce systemic acquired resistance and PR gene expression. *BMC Plant Biol.* 2, 9

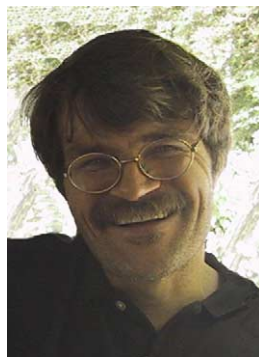


Marcela B. Treviño Santa Cruz has a background in microbiology, biochemistry and molecular biology. She has participated in diverse research projects in the areas of food technology and plant molecular genetics. Marcela is currently conducting marine biology research, developing molecular probes for phytoplankton population ecology studies. She aims to get involved in marine bioremediation and marine conservation work.

"Notre avenir dépend non de la science politique, mais d'une politique de la science." (Marc Gendron).



Dominique A. Genoud holds a PhD in computer communication systems, and has more than 15 years of experience in digital signal processing and digital signal modeling. During his stay at the University of Berkeley he contributed, together with the authors of the present review, to the new models of cellular signalling inspired by signal processing knowledge. His present interests converge on a linear and a nonlinear approach of cell signalling using both data driven information and Boolean signal abstraction.



Jean-Pierre Métraux holds a chair in Biology. He is trained in plant biology and has been working on resistance of plants to pathogens since 1983. He led a basic research team in the Agricultural Department of CIBA from 1983 to 1991. During this period, his group discovered the first synthetic inducer of resistance that was eventually commercialized. In 1991, he moved to the University of Fribourg where he has continued to study various aspects of induced systemic resistance with graduate and postdoctoral students. He has published over 100 articles, is a highly cited researcher (isihighly-cited.com) and coauthor of one patent.



Thierry Genoud holds a PhD in biochemistry of plant cell wall biosynthesis. During his postdoctoral stay at the Rockefeller University in New York, he studied the perception of light in *Arabidopsis* using classical genetics and new concepts of cell signal processing derived from computer knowledge systems. He is currently developing bioinformatic concepts and tools for the representation and analysis of cell signalling networks, notably networks controlling plant responses to pathogens.