

**DRAFT WORKING PAPER. NOT FOR CITATION.**

**Brazilian genomics and bioinformatics: instituting innovation processes in a global context.<sup>1</sup>**

**Mark Harvey and Andy McMeekin.**

**ESRC Centre for Research in Innovation and Competition.**

**Frontiers and Trends  
Frontiers of Innovation Research and Policy  
25-26 September 2002  
Rio de Janeiro**

**ABSTRACT**

Genomics and bioinformatics are transforming wide areas of scientific, technological and economic activity. New areas of innovation are opening up across the life science industries, agriculture and food provision. But at the same time, there are high levels of uncertainty of how, when, and where new markets, products, and services will emerge. Early expectations from genomics as the ‘key to life’ of many organisms have subsequently been seen to be too simplistic, with intervening complexity at every level from genome to organism (transcriptome, proteome, metabolome). In pharmaceuticals, the promised systemic shortening of innovation pipelines remains a promise. In agrigenomics, developments have largely remained at the primitive stage of enabling specific chemicals to be sprayed on specific crops. An informatic explosion of data and problems of interoperability within and between these different biological levels, present new challenges. In the context of this uncertainty, there are many alternative firm strategies, shifting boundaries and interchanges between public and private sectors, and pre-competitive co-operation. New classes of economic agent appear and disappear, or in the case of an Incyte or a Celera, change economic function.

Moreover, the development of these alternative institutional innovation trajectories have a geopolitical significance and specificity. The different major centres of gravity in the US, Europe and Japan also reveal forms of competition and diversity. The significance of Brazilian genomics and bioinformatics is that distinctive processes of innovation are being instituted that have achieved global leadership in specific domains and applications, linked to the needs and socio-economic interests of the country. The paper reports ongoing research on the sugar transcriptome, bacterial genomics and models of pathogenesis, and structural genomics of leishmaniasis, as examples of these processes.

---

<sup>1</sup> This paper is a working paper, based on continuing comparative research in Europe, the USA and Brazil. The research is at an early stage, and, for Brazilian genomics and bioinformatics, is part of a longitudinal study in co-operation with GEOPI at Unicamp, Campinas, Brazil, to whom we express gratitude for all the help and support. The European and USA research is a project ‘Bioinformatics and Economies of Knowledge in Europe and the USA’ funded by a grant from the ESRC, in addition to the funding for CRIC.

The paper concludes by arguing that distinctively instituted and distributed innovation processes lead to processes of variation and comparative advantage in an area of rapid institutional change and increasing complexity of organisational interactions.

## Introduction

The revolutionary transformation encapsulated in the terms ‘genomics and bioinformatics’ is a combination of the new science and technology and the new organisational forms of public science and private firms, including boundaries between, and interactions across, them. Scientifically, from the moment when the term bioinformatics was first publicly coined (Ryback, 1968<sup>2</sup>, 1978) to the present day, it has been a matter of contention as to what scientific epistemic activity or technologies it embraces (Harvey and McMeekin, 2002, McMeekin and Harvey, 2001). One way of describing this changing universe is in terms of a confluence between two distinct but complementary developments, each of which is now witnessing a major, often fragmented, expansion. On the one hand, as a continuation of the discovery of the double helix through to the completion of the first genomes, bioinformatics was defined partly by the nature of the digitalised data that was being generated. It was soon clear that bioinformatics could not be circumscribed by a particular type of DNA sequence data exemplified in the genome. Rapidly expanding new areas of digitalised content, each almost a domain in its own right, were emerging: the transcriptome, proteome, and metabolome, not to mention virtual organisms<sup>3</sup>. Each level of content adds new complexity between ‘dry’ informatic and ‘wet’ laboratory science, and achieving interoperability between heterogeneous data

---

<sup>2</sup> Ryback envisaged bioinformatics as an extension of computational biology, with the increasing digitisation of all types of biological data, long before genomic sequence data was even dreamt of (Ryback, 1968, 1978).

<sup>3</sup> The term ‘genomics’ in the title is thus not intended to restrict the field to the genome strictly speaking, but to include all those areas opening out in post-genomic domains.

domains presents increasing challenges to bioinformatics. On the other hand, from its earliest formation, bioinformatics was not so much to do with content as with a fundamental revolution in the nature of biological sciences and their methodology. What matters here is less the digitised content than the modelling of biological systems (Wolkenhauer, 2001; Grindrod, 2001). Mathematical modelling increasingly leads rather than follows the process of wet-science experimentation<sup>4</sup>, so requiring a radical ‘redisciplining’ of biology (Bornberg-Bauer, et al.). But here too, models imported from other sciences and the expansion of the domains of informatics to include chemo-, neuro-, demo-, and medico- informatics has entailed more proliferation of models and their data domains than integration between them. The confluence between these two expanding universes, therefore, has produced a multitude of new possibilities, alternative trajectories, and competing priorities, with as yet little overall coherence.

Organisationally, there has been a similar proliferation but of the different forms, scales, and interrelations of organisations engaged in genomic, post-genomic and bioinformatic activities in the public and private sphere. Just as there are different possible scientific and technological pathways, so are there also different possible organisational pathways, either differentiating or integrating bioinformatic and genomic activity from other activities, such as small molecule chemistry, or gene therapy and engineering. In terms of a ‘distributed innovation process’ (DIPs) approach (Coombs et al. 2001; McMeekin and Harvey, 2002), new forms of products and services (algorithms, integrated software packages, cryo-electron microscopy, interoperability mid-ware, 3D imaging of protein folding, biological sequence data, etc.) have emerged, as well as new economic agents, dedicated bioinformatics firms

---

<sup>4</sup> Interviews at GSK, European Molecular Biology Laboratory (EMBL), European Bioinformatics

or tool-providers joining the ranks of dedicated biotechnology companies. There have been new forms of interrelation between them, and as others have noted, a ‘fuzziness’ of boundaries between public and private sectors (Joly and Mangematin, 1997b; Fransman, 2001; Dasgupta and David, 1993; David et al. 2000). New public-private hybrids have appeared, and new forms of knowledge appropriation, within both the public and the private domains, have been developed (firewalls, conventions over publication and release of source codes or data into the public domain). It is for this reason that we have developed a concept of ‘economies of knowledge’ that combines different forms of resource flow (public and private) with different forms and directions of knowledge flow between public and private spheres.

A DIPs approach thus suggests that new combinations of new actors cutting across existing ‘instituted’ organisational boundaries of sector or technology are typical of the organisational forms embodying genomic, post-genomic and bioinformatic activity. Ink-jet printers, silicon chip manufacturers, microarray companies, major computer manufacturers are combining in new ways with drug discovery or food provisioning activities. But equally, a national systems or even regional systems of innovation approach becomes less applicable to an area of organisational change marked by regional clusters *combining* with local SMEs and global resources (Genbank, EBI, DGJB), or by firm-to-firm alliances that transect in multiple ways previous patterns of spatial organisation (McMeekin, Harvey, and Gee, 2002). So in parallel to the proliferation of alternative scientific and technological pathways there has developed a proliferation of alternative organisational pathways. This organisational fluidity is well exemplified by the switch in business focus of two major bioinformatic firms, Celera and Incyte, providing genome sequence data

‘content’<sup>5</sup>, to direct engagement in the drug discovery process. But their place in this new universe is less certain in relation to competing models of drug discovery and development, either within the large pharmaceuticals, or by smaller drug discovery firms whose economic existence is built around a particular integrative pathway through genome, transcriptome, proteome and metabolome<sup>6</sup>.

It is in this context of alternative organisational pathways that the geopolitics of innovation assumes significance. Both the scientific director of EMBL<sup>7</sup> and the director of the EBI<sup>8</sup>, the premier public science molecular biology and bioinformatics institutions in Europe, were clear that a multiplicity of competing poles were as much essential for the development of global science as for industrial competitiveness of Europe as a region. Thus, although there is 24-hour open and complete exchange of data between the three major genomic databases in Europe, the USA and Japan, - hence global integration – the existence of three main conduits from different regional scientific communities results in greater scientific productivity than a scenario of a single global and hegemonic centre based in the USA. It is striking that during the early development of genomics, the decision to sequence the yeast genome was seen as a European strategy to beat the Americans.

‘Europe needed to begin the project as soon as possible, perhaps on chromosome III; otherwise it would lose a unique opportunity for a competitive advantage over the USA and Japan. Goffeau<sup>9</sup> reduced the apparent dichotomy between cooperation and

---

<sup>5</sup> The term ‘content’ here refers to proprietary sequence data bases, which, following the Celera-HGMP race to be first to the rough draft of the human genome, lost their promise as a reliable source of business revenue.

<sup>6</sup> The interweaving of scientific, technological and organisational pathways is evidenced in the development of the firm organisation of companies such as Oxford Glycoscience, that has elected for a particular pathway to give it competitive distinctiveness. The business model is presented as part and parcel of the model for the organisation of scientific knowledge from identification of protein targets through to pre-clinical trials, on the basis of a distinctive experimental methodology (Parekh and Lyall, 2000; Lyall, 1997)

<sup>7</sup> Iain Mattaj, interview, Heidelberg, June 2002.

<sup>8</sup> Janet Thornton, interview, EBI, Cambridge, August 2002.

<sup>9</sup> See below.

competition with the USA with his vision of close association. For him, Europe should first reach a strong position from which to negotiate, after which it could dictate the conditions of cooperation.’ (Goujon, 2001, 385)

As we shall see, the organisational strategy behind the yeast genome was the model closely followed in Brazil in the ONSA network, so the geopolitics of this vision of scientific activity is significant for the argument to be developed here. It is clear that the additional significance in terms of comparative advantage for local and regional industries in biotechnology and informatics is geopolitically strategic. Throughout the development of genomics and bioinformatics capability within Europe there have been competing concerns between national, local and regional levels, some intensely national as around the genome of *Bacillus subtilis* (Danchin, 2000), others with a more regional perspective, where a European scale is seen as the only viable level at which to compete with the US or Japan (European Commission, 2001; Senker and van Zwanenberg, 2001). Yet, a DIPs approach suggests strongly that the new units of competition might not all reside equally easily within national, local or regional boundaries. Innovation processes may not fit into the same geopolitical boxes as those of policy-makers.

It is at this point that we introduce the notion of instituting innovation processes in a global context. Given the alternative possible scientific and organisational pathways, and given that these alternatives are inherently geopolitical in character, we will attempt here to present a preliminary analysis of some key developments in Brazilian genomics and bioinformatics. Distinctive pathways, organisational, scientific and technological, are instituted with specific resource flows from public and private sectors resulting in particular strengths at the global level. It will be shown that there is a variable geometry of alliances and spatial arrangements,

local and global, involving different aspects of genomics and bioinformatics in the recent development of Brazilian innovation capability. Of course, these are early days, both in the emergence of this capability and its development of new markets – and in our research of it. But in geopolitical terms, the development of distinctive pathways outside dominant centres (where markets and organisations may equally be suffering from crises in resource flows or locked in organisational or market deadlock), may be of great importance in the long term for the world balance of scientific and technological development.

No attempt is being made to present a comprehensive view of Brazilian genomics and bioinformatics. Rather, four particular cases have been selected to demonstrate the diversity of organisational pathways, the first being analysed in some depth and detail, the remaining three used to highlight contrasting aspects.

- The transposition of the European network organisation of the yeast genome into the São Paulo region and the creation of the Organisation for Nucleotide Sequencing and Analysis (ONSA) network for the genomics of *Xylella fastidiosa* and *Xanthomonas citri*.<sup>10</sup>
- The collaboration between Unicamp bioinformatics, Washington University and Pioneer Hi-Bred Inc. in the achievement of the draft of the genome and comparative genomics for *Agrobacterium tumefaciens* – ‘the natural genetic engineer’ (Wood, Setubal, et al. 2001)<sup>11</sup>.

---

<sup>10</sup> We thank Jose Fernando Perez, Scientific Director, FAPESP; Professors Paulo Arruda, Joao Meidanis, Joao Setubal, and Dr Joao Kitajima at Unicamp; Dr Anaclaudia Rosera, DNA coordinator for *Xanthomonas* at USP; and Professor Steve Oliver, University of Manchester for interviews in relation to the ONSA network. We thank Dr Fernando Ribeiro, Director, FINEP, and Dr Sergio Salles Filho, Operational Planning Superintendent, FINEP and Director of GEOPI, Unicamp, for interviews related to FINEP and innovation strategy in biotechnology.

<sup>11</sup> We particularly thank Joao Setubal and Joao Kitajima for information and interviews on this project

- The strategic creation of a sugar cane transcriptome data bank (SUCEST), the spin-off of a biotechnology company, and collaboration with Belgian firm dedicated to developing crop improvement traits (CropDesign).<sup>12</sup>
- The formation of an international network, with Brazil as initiator and collaborator, of the *Leishmania* Genome Network, targeting a disease that typically afflicts the poor in tropical climates, with a high incidence in Brazil (WHO, 2000)<sup>13</sup>.

Each of these four cases exemplify different aspects of how Brazilian genomic and bioinformatic innovation processes are being instituted in a global context, in terms of geography of collaboration, global significance of the innovatory activity, and the particular Brazilian orientation behind the process of institution.

## Four Brazilian innovation pathways

### **1. Model organisation and model organism**

There can be little doubt that the establishment of the ONSA network in São Paulo state in May 1997 was a platform for a breakthrough of Brazilian genomics onto the global scene, culminating in the publication of the world's first genome of a plant pathogen (*Xylella fastidiosa*), celebrated on the front page of *Nature* (Nature, 2000). A signature list of 107 names, beginning with Simpson and Reinach, did not reach the record of 147 authors achieved by chromosome III of the yeast genome

---

<sup>12</sup> We particularly thank Paulo Arruda for extensive interviews on this subject.

<sup>13</sup> The WHO reports that 12 million people are infected worldwide, with 1.5 million new cases each year. Co-infection with HIV is increasing. We particularly thank Dr Angela Cruz at USP and Dr Fernando Ribeiro, Director, FINEP, for interviews in relation to the LGN programme.

(*Nature*, Oliver et al., 1992), but testified to a common network organisational model that stretched even scientific norms of authorship.

The features shared by the European yeast genome project and the ONSA *xylella* project are significant to the argument of this paper, but as much because of the novelties of the Brazilian redesign of the model as because of its origins in the first European template.<sup>14</sup> The instituting of a distinctive innovation pathway can be analysed by comparing the yeast genome project from the ONSA genoma programme across the three dimensions, the nature and resources of the network organisation, the scientific and technological trajectory, and geopolitical positioning. This comparison is designed to demonstrate the inadequacy of notions of technology transfer or dependency models of development to account for the innovation processes in question. The intention is to show both the similarities and differences between the European template and the Brazilian model.

*The network organisation and resourcing.* The historical link between the yeast genome network and the Brazilian ONSA network is most strongly reflected in the Advisory Committee established to launch ONSA. André Goffeau, the main inspiration and overall coordinator of the European BAP, BRIDGE and EUROFAN programmes through which the yeast genome was principally financed, was on the Committee together with Steve Oliver<sup>15</sup>, DNA coordinator of the chromosome III sequence, which had piloted the whole network organisation. The Advisory Committee was instrumental in designing the functioning network, and responsible

---

<sup>14</sup> In 1997, Joly and Mangematin asked the question as to whether the European yeast network model was exportable, and were highly sceptical, as well as critical of aspects of the way *public* proprietary rights over delayed publication were assigned to each laboratory of the network, prior to release to public databases through the Martinsried Institute of Protein Sequencing (MIPS). (Joly and Mangematin, 1997a).

<sup>15</sup> A further interesting ramification of the linkage between European yeast and Brazilian genomics was that Steve Oliver had involvement with alcohol production, and alcohol production was to be important to the sugarcane SUCEST project, granted Brazilian use of fuel alcohol.

for soliciting proposals from laboratories, selecting the 34 laboratories taking part in ONSA and in choosing the organism for complete genome sequencing.

Interestingly, for the original European organisational innovation, there were many similarities to the subsequent Brazilian innovation pathway. In establishing a yeast genome sequencing as a network organisation, Goffeau was pioneering a model that contrasted strongly with the US and Japan programmes that were highly centralised and technologically based on the use of a supercomputer. The US project model involved the installation of a Cray supercomputer at a cost of US\$15 million with \$3million annual servicing costs.

‘If the project was decentralised, the central computer for the yeast genome could be a superminicomputer of the Convex type.’ (Goujon, 380.)

Capital costs of the network system were a twentieth of a Cray, and maintenance a tenth. But the advantages of the network organisation were by no means solely technological or financial. In addition to the decentralised workstation-based intranet network between 35 laboratories – effectively a composite virtual laboratory – the organisation was seen to have a major objective in building up and distributing genome sequencing skills and capabilities across Europe, with common peer reviewed standards.

‘It allowed the pleasure of small science to be maintained, i.e., the interactions, collaborations, excitement of confirming discoveries, avoidance of conflicts and problems involved in hierarchy and the maintenance of efficiency in work carried out at reasonable cost.’ (Goujon, 2001, 441)

If European genome sequencing capability was a central positive gain from the networking organisation, the manner in which the network functioned was also highly innovatory in a number of ways. Although developed in the course of the

project, the key elements were dual coordination by a DNA coordinator and a bioinformatics coordinator. The DNA coordinator distributed clones for sequencing to each of the laboratories, whilst the bioinformatics coordinator ensured that the results were quality controlled in terms of data standards between all 35 laboratories. In itself, this networked process of quality standardisation significantly developed the bioinformatics involved in sequencing.

To ensure that all the laboratories pulled together, a unique incentive structure was developed, whereby funds were released for each unit sequenced, a system that has been described as ‘payment by the piece’ (Joly and Mangematin, 1997a, 11; Joly and Mangematin, 1997b, 4). Given the current state of technology, an initial rate was set at 5 ecu per base pair. Eventually a system of rules were codified, so this payment system was further incentivised by a ‘first come, first served’ arrangement (Vassarotti et al. 1995), whereby units for sequencing were distributed on receipt of already completed sequenced units.<sup>16</sup> The more a laboratory did, competitively within the network, the more it received in funding. The rules on intellectual property rights were a final key element of this European organisational template. Unlike US or UK publication pattern for genomic sequences at the time, where small units of 110-200kb were quickly released to the public domain, each laboratory in the European network submitted results to the central coordinator, and results were published only when complete chromosomes were achieved, in effect a temporary appropriation of scientific results by each laboratory for up to two years (Vassarotti, 1995, 136). Further, each laboratory was given rights to patent its results. In the view of Joly and Mangematin both temporary appropriation and arrangements for patenting amounted to forms of private appropriation of public science. The consequence could result in

delaying scientific progress by restricting access to knowledge (Joly and Mangematin, 1997a, 1997b).

Chromosome III sequencing, undertaken between January 1989 and May 1991, was a world first for a whole eucaryote chromosome (Goujon, 2001, 434). Following the publication of chromosome III in May 1992, this same network model was then extended to the other 15 yeast chromosomes under the BRIDGE T programme from 1993 to 1997, and to other organisms, such as bacillus subtilis. A central feature of the yeast genome programmes was the nature of its resourcing. With regard to organisational models and the instituting of innovation processes, the nature and regulation of resources sustaining them are critical (Senker and van Zwanenberg, 2001, 35-37). The yeast genome was financed on a project-by-project basis, within research funding programmes, successively the Biotechnology Action Programme (BAP), the Biotechnology Research for Innovation, Development and Growth (BRIDGE), and the European Functional Analysis Network (EUROFAN). During the processes of acquiring funds there were continuing inter-directorate and inter-organisational conflicts, for example over whether yeast genomics was an agricultural or industrial project, or a fundamental or applied science project. Indeed the Industrial Research and Development Advisory Committee (IRDAC)

‘remained opposed to the yeast project in particular, and more generally, to systematic sequencing proposals, considering them of no use to industry.’ (Goujon, 2001, 413).

Yet from the earliest days, Goffeau recruited industries using yeast to support the genome project, and especially the food and brewing industries. In the event, some financing for the project came from business subscriptions through the Yeast Industry

---

<sup>16</sup> ‘The Perfect Gentleman Sequencer: The development of Deontological Rules.’ June 1992, developed

Platform (YIP) that included 13 leading companies.<sup>17</sup> However, the bulk of the remaining finance was from national governments supporting the national laboratories involved in the network. Nonetheless, despite subsequent involvement by non-European laboratories in sequencing later chromosomes, the large number newly discovered genes were dubbed ‘EEC’ genes by one of the coordinators (Piotr Slonminski). In true European acronymophiliac style, it had to be explained that this stood for ‘Esoteric, Elusive and Conspicuous’ genes that were companions to FUN genes (Function unknown). Thus, overall there was a complex mixture of public and private resource flows, with funds for discrete projects on a national and European basis, as well as from Transnational Corporations. In this sense, one of Goffeau’s major achievements as an European Commission official, was in the organisational diplomacy required to negotiate continuity of funding for yeast genomics across different programmes, in the context of considerable organisational tensions and conflicting interests. But in the end, the ‘official history’ of the project recognised the triumph of this type of organisation.

‘Yeast will become the navigation instrument for the voyages being undertaken across the world into large genomes of higher complexity.

The fruit of world collaborations and a unique network of laboratories, conceived and established by André Goffeau and sustained by the European Community through its various programmes, has ensured the success of the yeast program and allowed the modest creature used for centuries by bakers and brewers to enter the pantheon of sequenced organisms.(Goujon, 2001, 501)

The organisational ‘navigation instrument’ was taken up by the ONSA network in many ways, although significantly redesigned. Of course, at one level, it is

---

by Bernard Dujon.

<sup>17</sup> Notably, Interbrew, Unilever, Royal Gist Brocades, and Rhone Poulenc).

impossible to disentangle model organisation from model organism. The yeast genome has 16 chromosomes with 13 billion base pairs, as against the *Xylella fastidiosa* genome of one chromosome and 2.6 million base pairs. The more complex the organism, the more complex are the organisational requirements for a networked system. Indeed, this was one reason why Joly and Mangematin (1997b) considered that the decentralised network organisation was not exportable to organisms significantly more complex than yeast.

ONSA replicated the European template by having an intranet ‘virtual laboratory’ structure with 34 laboratories; a strong coordination by a DNA coordinator distributing clones for sequencing; a centralising function for bioinformatics at Unicamp; and an incentive system of payment by results.<sup>18</sup> A similar, if not greater, emphasis was placed on skill formation and distribution within the network, as a Brazilian resource. But, there is much to suggest a much higher level of inter-laboratory co-operation in the Brazilian model. Technical improvements were immediately shared, as were results of sequencing. With one chromosome, there was no question of division of property rights between laboratories, and there was a strong view that the ‘genome was done in a community of laboratories’ (Rosera, interview). There is also little doubt that the ONSA network achieved a synergy between previously dispersed and institution-bound researchers, as well as attracting and training much new talent. One key was the enrolment of two computational biologists, Meidanis and Setubal, to undertake the bioinformatics of sequencing in which they had no previous involvement.

But, in organisational terms, perhaps the sharpest contrast with the European network was in its resourcing, and in the consequent ability to take risks and play for

high stakes within an environment of assured continuous funding underwritten by the Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP). There were none of the European diplomatic complexities that beset the yeast genome sequencing. Steve Oliver was in no doubt that the funding system for ONSA was critical to its rapid success in global genomics, not least because of the commitment and courage of FAPESP's Scientific Director, José Fernando Perez in initiating and supporting ONSA. Since 1962, the system of financing research in São Paulo state was through hypothecation of 1% of the state's total revenue for research, freeing it from normal political pressures. Half of FAPESP's Board of Trustees, consisting of twelve members elected for a six year term, were from nominees within the scientific community. This ensured a strong science priority to the allocation of funds. Although also supported financially by the National Committee for Scientific Research (CNPQ, intermediated by FINEP), the essential decision that enabled the organisational breakthrough to establish the network came through FAPESP. Finally, further resources were provided by Fundicitrus, which, for reasons that will become clear when discussing the model organism chosen, had an interest in supporting the sequencing of a plant pathogen attacking citrus production.

*The model organism.*

The scientific and technological trajectory of the Brazilian ONSA network, as with the European network, was decisively conditioned by the choice of the organism to be sequenced. Both *Xylella fastidiosa* and *Saccharomyces cerevisiae* were strategic organisms, and in this section we analyse this in scientific and technological terms, to deal with their geopolitical significance in the next. Again there are significant

---

<sup>18</sup> The going rate at the time of the *Xylella f.* genome was \$R4 for each complete Open Reading Frame of 400 bp.s with a thread value of 20). League tables of productivity were published on the intranet.

similarities, but yet stronger contrasts between the European template and the Brazilian model.

The case made for the first European genome project (Vassarotti et al. 1995) has uncanny similarities with subsequent Brazilian reasoning. Yeast was a small genome, and therefore could be tackled by the skill and technological resources available in Europe at the time. It also had socio-economic relevance important to industrial support and collaboration as yeast is central to brewing and baking. Likewise, *Xylella fastidiosa* was a small genome that was estimated to be within the potential of the developing resources of the ONSA network. It also had considerable socio-economic significance, especially to São Paulo state. The citrus industry is a major global crop on which 400,000 jobs depended, and citrus fruit is attacked by this bacterial pathogen<sup>19</sup>, reducing the overall yield of the crop by 25%. It was also known that other strains of *xylella* were responsible for socio-economically significant diseases in grapes (Pierce's disease) alfalfa, and plum. One alternative candidate genome was a bacterium that could metabolise metal with potential for bio-mining of gold in the Amazon. The choice for São Paulo state was to go for gold in the Amazon, or orange juice in the A, B, C region. It was also a choice between understanding a pathogen in order to aid agriculture in the long term, and going straight for riches.

The differences between the two model organisms were significant, and to become more so as the programmes developed. The yeast genome was already well-known, and with an established yeast scientific community,<sup>20</sup> there was already a fully developed physical genetic map of yeast. *Xylella* was virtually unknown genetically, and Steve Oliver considered it high-risk to choose as a model for that reason. There

---

<sup>19</sup> The pathogen is transmitted to citrus trees by sharpshooter leafhoppers, and causes Citrus Variegated Chlorosis (CVC).

<sup>20</sup> Given the strong influence of francophones in the network, these are commonly referred to as *levuristes*.

was a continuing tension as to whether the yeast genome was to be considered as a generic model organism; a platform for drug discovery as a consequence of sharing homologous genes with humans; or whether it was primarily an organism with traits of industrial interest (Vassarotti et al. 1995); or all three. This tension persists to this day and is reflected in a European leading proteomics company, Cellzome. Having achieved the world first yeast proteome to complement the genome (Gavin et al. 2002), it is as yet unclear how this contributes to integrated drug discovery, its declared business goal.<sup>21</sup> By contrast, *Xylella f.* is a known pathogen, and this has led to functional genomics enabling ONSA to now proceed towards ‘lead-pesticide’ discovery. The prospect is for chemical treatments where traditional pesticides have failed, leading to the total eradication of the disease. This is yet quite far from being achieved, but the innovation pathway is quite clearly defined, and as an innovation strategy strikingly parallels lead-drug discovery for human pharmacological therapeutics. Furthermore, following the initial success in genome sequencing, ONSA was invited by the USDA to undertake the annotation work on strains of *xylella* affecting Californian grape crops, in preference to the American Joint Genomics Institute that had proved relatively uncompetitive in this task.

So a major difference between yeast and *Xylella* is that the Brazilian trajectory is directed toward agricultural genomics and technologies for crop protection, whereas yeast genomics has an as yet uncertain possible outcome, in food processing and pharmaceuticals. Yet, it is in the nature of post-genomics that even this divide between agricultural and pharmaceutical outcomes has become much less clear. As a consequence of the functional genomics arising from the sequencing of the *Xylella F.* genome, it has become evident that plants share many of the characteristics of

---

<sup>21</sup> Interviews with Georg Casari, Walter Blackstock, Cellzome, 2002.

pathogenicity with animals, including humans. In this way, the pathway opened by this strategic organism has delivered a far greater potential than originally anticipated, one that has been reinforced by the subsequent inclusion of other pathogenic microbial organisms, notably *Xanthomonas citri*, and *Agrobacterium tumefaciens*. This plant-route to the understanding of pathogenicity again has the advantage of starting with relatively simple models of biological systems compared with direct study with human pathogenicity. Experimentation on phytopathogenic processes is much more tractable, presents fewer ethical dilemmas, and is yet capable of producing generic understanding. In the view of Steve Oliver, the Brazilian pathway has thus found itself well placed, in global terms, as a consequence of pursuing its apparently more modest objectives.

*The geopolitics of organisations and organisms.* For Europe, the yeast genome project was of geopolitical significance in building European genome sequencing skills and capability. It created an alternative model to centralised big science, fostered an integrated European science base, and developed a skills base spread across the region. There is one important way in which what the yeast genome did for Europe, *Xylella* did for Brazil. Choosing the yeast genome meant that Europe was able to achieve advances over the Americans with their precipitous concentration on large genomes. This gave Europe an advantage when it came to tackle the human genome, one that the Americans only latterly recognised.

‘The Europeans did not make human genome sequencing an absolute priority just to later see that the study of model organisms was a step that could be bypassed. They moved directly to that step by sequencing small genomes. Before the Americans had even started sequencing the human genome, the Europeans had already sequenced large chunks of the yeast and *Bacillus subtilis* genomes. For Europe, this head start was very important.’ (Goujon, 2001, 420.)

ONSA's achievement of the front page of *Nature* within two years of its launch with a single chromosome bacterium proved to be of equivalent geopolitical significance, later to be consolidated by other projects. Within a global context of considerable flux where there is a multiplicity of possible pathways, the grounding of the research programme in local socio-economic interests, and the targeting of a specific pathogen, managed to achieve globally significant results with relatively restricted resources. That the USDA contracted with ONSA in preference to the Joint Genomics Institute for completing the sequencing of genomes relevant to US agriculture both recognised and reinforced this global status. The fact that the organisation and the organism combined to then open out into much broader potential and importance may have been partly fortuitous, but nonetheless demonstrates how innovation leverage can be gained on the basis of small-scale but critically significant scientific and technological research programmes. The geopolitics of ONSA and *Xylella* thus illustrate one particular way in which Brazilian genomics and bioinformatics operating on a *local* state basis create a distinctive *global* presence. Moreover it has now become the Brazilian template for the national genome project, also funded by forms of tax hypothecation, through the CNVPq and FINEP (Salles and Ribeiro, interviews). The networked innovation organisation has thus expanded from state level to federal level, so becoming an organisational innovation platform for Brazilian biotechnology.

## **2. 'The natural genetic engineer'.**

The leverage gained by sequencing the world's first phytopathogen formed the basis of opening next Brazilian pathway to be considered. Very different in inter-organisational links and global connections, it too is a strategic organism with great geopolitical charisma: *Agrobacterium tumefaciens*, the 'natural genetic engineer'

(Wood, Setubal, et. al. 2001; Goodner, Hinkle, et al. 2001). *A. tumefaciens* has earned this sobriquet because of its generalised use in biotechnology for genetic engineering of plants, as a consequence of its natural characteristics as a plant pathogen that induces cancerous growth – genetic modification – in the host, the ‘crown gall’ seen on many trees. By extracting the pathogenic genes from the natural genetic engineer, and replacing them with genes for traits desired for agriculture or nutrient dense food, the bacterium becomes an ideal tool for the human genetic engineer.

However, as with yeast, *A. tumefaciens* presents this dual characteristic, as an organism of fundamental scientific interest and immediate technological significance. Understanding the nature of its pathogenicity is at the same time understanding a technology of gene transfer.

Brazilian presence in the publication of the genome (through Joao Setubal and Nalvo Almeida) was achieved through the development of automated annotation software for *Xylella*, the same software that had given ONSA the edge over the Joint Genomics Institute. But in this project, Setubal and Almeida collaborated with the University of Washington and Pioneer Hi-Bred, on a pre-competitive basis in which participation was seen to be of mutual benefit. Clearly, those institutional linkages underpinning the achievement demonstrate a global genomic activity, in which Brazil was able to take a leading role in the key domain of bioinformatics. In that respect, although resting on the ONSA network and the bioinformatics at Unicamp, organisationally it presents a different pathway to that of the São Paulo virtual laboratory.

The technological and scientific pathway is open to a number of alternatives, although the functional genomics already has demonstrated ways of enhancing

genetic transfer through an understanding of the promotion of different levels of virulence of the pathogen. On a scientific level, the discovery that *A.tumefaciens* has extensive similarity with *Sinorhizobium meliloti*'s circular chromosome through comparative genomic analysis (also one of Unicamp and ONSA's microbial genomics strengths) has proved doubly significant. Firstly, it has meant reclassifying the bacterium into the family of *Rhizobium*, and secondly it has opened up the study of the evolutionary divergence between pathogenic and symbiotic lifestyles, a divergence that in the case of these closely related bacteria, was relatively recent. Given that HIV is symbiotic in chimpanzees and pathogenic in humans, understanding of these biological processes and the nature of interactions between parasites and hosts may well be the basis of significant advances in biological understanding of infectious disease in humans.

As with the *Xylella*, the combination of organisation (here international scientific and technological collaboration) and organism (fundamental model organism and natural engineer), has provided Brazilian genomics with an additional innovation pathway to the global leading edge.

### **3. The Brazilian Sugar Express**

Another major development arising from the ONSA network has been the SUCEST project related to the transcriptome of the sugarcane. Given the size and complexity of the sugar cane genome, the objective has been to identify gene expression (Expressed Sequence Tags, ESTs) in specific traits rather than to engage in whole genome sequencing<sup>22</sup>. The SUCEST programme involves a distinct innovation pathway, partly in complementarity with other global sugarcane genome projects<sup>23</sup>,

---

<sup>22</sup> In Steve Oliver's view, this strategic decision to concentrate on ESTs was another example of gaining maximum leverage from available resources and skills.

<sup>23</sup> Sugarcane genome sequencing projects are being undertaken in Australia, South Africa and the USA.

and critically through a transatlantic gene functionality alliance for commercial exploitation of a patented database of sugarcane ESTs held at Unicamp. From a DIPs perspective, therefore, the distinctiveness of SUCEST is in the way that ONSA has linked with other economic agents and operates on different geographical scales.

The strategic significance of sugarcane is obvious as Brazil is responsible for 25% of world production, half of which is concentrated in São Paulo state. With Paulo Arruda as DNA coordinator and Joao Meidanis as bioinformatics coordinator, the project had identified 1,000 previously unknown genes, and had constructed a data base of Single Nucleotide Polymorphisms (SNPs) related to environmental and biotic stress by September 2001. As with drug discovery-related genomics, SNPs are considered to be a fast route to commercial exploitation, because they concern single genes related to the expression of one particular trait. Using the technologies of automated microarray analysis on material from old clones, they have attained expression profiles for key aspects of sugarcane cultivation under different environmental conditions. The services of a US patent attorney were engaged, and they have been able to successfully patent a unique database of sugarcane SNPs related to key aspects of sugarcane metabolism.

The SUCEST innovation process involved two other important institutional aspects, one domestic the other transatlantic. From the Brazilian industrial side, the Copersucar, the sugar producers' cooperative provided not only financial support, but also collaborated in datamining and physical genetics from their own research facilities.

The transcontinental alliance involved collaboration with CropDesign, a Belgian firm specialising in validating and developing gene expression processes throughout the lifecycle of the plant ([www.cropdesign.com](http://www.cropdesign.com)). Its proprietary

technology platform, TraitMill™ is a high-throughput technology that automates gene identification expressed in gene functionality, and trials them on laboratory plant specimens. It can also be employed for determining and improving the expression of new traits in transgenic plants. CropDesign in turn is involved in a number of pre-competitive projects partly financed by the European Commission, related to plant growth and metabolism (ECCO, GVE, and DAGOLIGN). In combination with this additional technology, TraitMill enables a fast track route from identification of sugarcane ESTs (in Unicamp) to product development. The functional genomics agreement between SUCEST and CropDesign was announced on September 4, 2001. The agreement gives CropDesign exclusive rights of access to SUCEST's database for a defined period, and rights to commercialisation of the results outside Brazil. For commercial exploitation in Brazil, there is joint responsibility and benefits to the partners. The choice of CropDesign by SUCEST was strategic, a relatively small biotechnology firm rather than a large Transnational Corporation. The power balance in the collaboration was further maintained by the controlled release of genes to CropDesign (only a dozen out of a total of 50,000 at the time of interview).<sup>24</sup> Thus, a specific institution of an innovation process involves arrangements over the property rights, where several forms of appropriation are involved: appropriation by technology behind the *public* science firewall of unpatented genes; patented genes owned by a public science spin-out in Unicamp, but not yet available for commercialisation; and exclusivity agreements for geographically divided property rights over commercialisation of genes between the spin-out and a biotechnology company in the open traded market-place. These complex arrangements have created specific 'instituted economies of knowledge' (McMeekin and Harvey, 2002). They

---

<sup>24</sup> Paulo Arruda, September 2001).

also involve the institution of a distributed innovation process between different classes of economic agent sustained by both public and private resource flows: ONSA, Copersucar, Unicamp, CropDesign, and the European Commission. Thus whether from standpoint of the different forms of appropriation of knowledge or from inter-institutional arrangements and resource flows, the SUCEST innovation process as a whole represents a complex of public and private domains. Viewed in terms of instituted processes and flows, a dichotomous view of the relation between public and private seems quite inadequate. It is a prime example of multi-economy processes, where market and non-market processes combine (Harvey and Randles, 2002).

#### **4. Genomics and therapeutics for tropical diseases**

The final instituted innovation pathway to be considered differs from the others in many respects. It is disease- rather than agro-oriented. It involves an international network funded by the World Health Organisation (Tropical Disease Research), the World Bank, and national science research organisations, the CNPQ in the case of Brazil. The Leishmania Genome Network (LGN) was launched in Rio de Janeiro in 1994, and followed the science and technology strategy of undertaking projects which had direct socio-economic significance for Brazil (Ribeira, interview)<sup>25</sup>. Worldwide leishmaniasis affects 12 million people with 1.5 to 2 million new cases a year, and depending on the variety of the disease, is lethal or causes severe lesions and ulcerations. 90% of all cases occur in Bolivia, Brazil and Peru (WHO, 2000). Of the 10 laboratories in the worldwide network, two are in Brazil, one in USP (Angela Cruz) the other in Belem (Schneider and Sampaio)<sup>26</sup>. The

---

<sup>25</sup> Other tropical diseases, such as schistosoma, are also being supported in genomics programmes of a similar kind.

<sup>26</sup> Of the other laboratories all in the 'developed' world, the United Kingdom has three, and France, Switzerland, Canada, Israel and the USA each host one.

coordination of this network is from the United Kingdom, at the Sanger Centre and Addenbrooks Hospital in Cambridge.

As a protozoan parasite, leishmania is much more complex than bacteria, containing 36 chromosomes, and is transmitted by a sandfly vector within which it develops prior to infectivity. Given this complexity, the early history of the LGN was engaged in the physical mapping of the genome through selective hybridisation and trait isolation (Ivens and Blackwell, 1999). It thus did not involve the DNA sequencing of genome projects or the bioinformatics that accompanied it until relatively recently. The early strategy, apart from physical mapping, was to concentrate on ESTs, in the same manner as the sugarcane and for similar reasons. Although it was thought that a complete physical map was a precondition for genomic sequencing, nonetheless the first chromosome was initiated in 1996. It is now anticipated that the whole genome sequencing will be completed by late 2002 (Almeida, Norrish, et al., 2002).

In terms of the scientific and technological trajectory, this innovation pathway is also promising new directions with the development of genetically engineered vaccines, based on the understanding of gene function and proteomics. In this development, research on biological systems of host-parasite interactions become significant. Moreover, two distinct routes have emerged from the proteomics of leishmania with potentially differing geopolitical significance. Vaccine development in the public science LGN has now been complemented by the discovery that Zantac, an anti-inflammatory digestion drug recently acquired by GlaxoSmithKline, is effective in treating some varieties of leishmaniasis. This discovery was the consequence of using proteomic analysis of the side effects of a pharmaceutical product. So the preventative-medicine public science route has now been joined by a

big pharma therapeutic route. Whilst there is no reason to believe that one is in competition with the other, nonetheless each demonstrates different configurations of economic agents in the innovation process, and different combinations of public and private science from different geographical regions.

### Conclusion: emergent strength through diversity

The four cases of emergent Brazilian genomics and bioinformatics have been selected to demonstrate the institution of innovation processes in a global context. Many important developments, such as the INOVAR programme and the national genomics programme with its bioinformatics centre, the National Laboratory of Scientific Computation at Petropolis, have not been covered. But the evidence of the four contrasting cases poses some difficult questions for a national or sectoral systems of innovation approach, or for technology transfer and development approaches. The four cases demonstrate different configurations of economic agents, cutting across geographical space in different ways, and sustained by various different flows of knowledge and resources between public and private sector. The table below summarises the main differences in terms of scientific and technological trajectories and geopolitical significance. Of course, the key that unites these different innovation pathways is their Brazilian point of origin. They are all distinctively Brazilian pathways, without being components of a Brazilian national system of innovation.

	<b>Science and technology</b>	<b>Institutional organisation</b>	<b>Geopolitical significance</b>
1. <i>A model bug – Xylella f.</i>	First whole genome of a phytopathogen. From plant to generic pathogenicity	ONSA network, local state based, producer cooperative.	Brazil at a leading edge, sought after by the USDA and California Grape industry
2. <i>The natural genetic engineer – A.tumefaciens</i>	Understanding the evolutionary divergence between symbionts and pathogens. Researching the main technological tool for agricultural genetic engineering.	ONSA network, Washington University, Pioneer Hi-Bred Inc. Precompetitive collaboration.	Brazil engaged with global players with a globally strategic organism.
3. <i>The Brazilian Sugar Express - SUCEST</i>	Development of functional genomics of sugarcane Platform for genetic engineering for enhancement of crops without use of chemicals.	ONSA, Unicamp spin-out, transatlantic alliance, and European Union.	Enhancing a global crop of which Brazil is a globally significant producer.
4. <i>Tropical disease genomics - Leishmaniasis</i>	Understanding parasite-host interactions. Developing preventative technologies through genetically engineered vaccines	International public science network of laboratories, coordinated from UK, but with Brazil as the only ‘developing’ country participant.	Tackling a disease afflicting 12 million people worldwide.

The approach developed here is one of ‘instituted economic processes’ (Polanyi, 1957), of which distributed innovation processes are the particular focus of the paper. Innovation is shown to be sustained by differently instituted flows of resources and knowledge, where new classes of economic agent emerge, and existing classes of economic agent are transformed, leading to new types of interaction between them. The ONSA network involved in three of the cases was shown to be a

distinctively Brazilian redesigning of an originally European public science template. In each of the three cases, moreover, ONSA is itself the point of origin of very diverse distributed innovation processes. SUCEST is one case in point, where the public science nature of a university (Unicamp) is transformed by its generation of forms of private appropriation of knowledge combined with commercial arrangements with a new type of biotechnology firm in Europe, itself partly sustained by supranational funding from the public sector (the European Union). As the table illustrates, each case presents a different science and technology trajectory, a different configuration of flows of knowledge and resources, and a different geopolitical significance and impact.

Thus, the geopolitical dimension points to the way in which differently distributed innovation processes achieved global leading-edge science and technologies with quite different impacts: crop protection, generic pathogenicity, genetic engineering of plants, sugarcane cultivation, or genetically engineered vaccines. Each of these Brazilian ‘building blocks’ (Derengowski et al.) has utilised relatively limited resources and capabilities to attain global leverage on critical but narrow fronts. In global terms, Brazilian genomics and bioinformatics may not be on the same scale as clusters and centres in Europe, Japan, and the USA, where major genome and proteome projects have been and are being undertaken. But, in the long term, opening up distinctive pathways that attain global significance offers at least the possibility of a geopolitical redistribution of innovation activity and the benefits of a multiplicity of emerging centres of gravity. It is a scenario of increasing diversity following the logics of new innovation pathways that arise from the exigencies and orientations of different societies and economies.

## References

- Almeida, R.; Norrish, A., Levick, M., Vetrie, D., Freeman, T., Vilo, J., Ivens, A., Lange, U., Stober, C., McCann, S., and Blackwell, J.M. 2002. 'From genomes to vaccines: *Leishmania* as a model.' *Philosophical Transactions of the Royal Society*, 357, 5-11,
- Bornberg-Bauer, E., Harvey, M. and McMeekin, A. (2002), 'Bioinformatics and education: skills shortage or fundamental shift in skill creation and demand.' *European Biopharmaceutical Review*, May
- Brown, N, Nelis, A, Rappert, B., Webster, A., van Ommen G.J.B. (1999), *Bioinformatics. A technology assessment of recent developments in bioinformatics and related areas of research and development including high throughput screening and combinatorial chemistry*. Final Report for the Science and Technology Options Assessment, European Parliament. May.
- Coombs, R., Harvey, M. and Tether, B. 2001. 'Analysing Distributed Innovation Processes.' *CRIC Discussion Paper*, 43. ESRC Centre for Research in Innovation and Competition. University of Manchester.
- Danchin, A. 2000. 'A brief history of genome research and bioinformatics in France' *Bioinformatics*, Vol 16:1, 65-75
- Dasgupta, P. and David, P. 1993 'Towards a new economy of science', *Research Policy*, 23, 487-521.
- Derengowski, M. G., da Silveira, J.M. and Salles-Filho, S. 2001 'Recent biotechnology development: challenges and opportunities to the consolidation of its knowledge "building blocks"'. Mimeo.
- David, P., Hall. B. and Toole, A. 2000 'Is Public R&D a compliment or substitute for private R&D? A review of the econometric studies', *Research Policy*, 29, 1165-1183.
- European Commission. 2001. 'The competitiveness of European biotechnology: a case study in innovation.' *Competitiveness Report*, Chapter V, 97-136.
- Fransman, M. 2001, 'Designing Dolly: interactions between economics, technology and science in the evolution of hybrid institutions.' *Research Policy*, 30, 263-73.
- Gavin, A-C., Bösche, M. et al. 2002. 'Functional organisation of the yeast proteome by systematic analysis of protein complexes.' *Nature*, 415, January, 141-147.
- Goodner, B., Hinkle, G. et al. 2001. 'Genome Sequence of the plant pathogen and biotechnology agent *Agrobacterium tumefaciens* C58', *Science*, 294, 2323-2328.
- Goujon, P. 2001. *From Biotechnology to Genomes. The Meaning of the Double Helix*. World Scientific. New Jersey.
- Grindrod, P. 2001. *UK Bioinformatics for Functional Genomics: Watching the Detectives*. Numbercraft.
- Harvey, M. and McMeekin, A. 2002. *UK bioinformatics. Current landscapes and future horizons*. Department of Trade and Industry Report. March.  
<http://les.man.ac.uk/cric>
- Harvey, M. and Randles, S. 'Markets, the organisation of exchanges and 'instituted economic process' – an analytical perspective.' CRIC Discussion Paper 51.  
<http://les.man.ac.uk/cric>

- Ivens, A.C. and Blackwell, J.M. 1999. 'The *Leishmania* genome comes of age.' *Parasitology Today*, 15:6, 225-231
- Joly, P.B. and Mangematin, V. 1997a. 'How long cooperation in genomics is sustainable.' Wheale, P. et al. *Science Technology and Human Values*. Avebury. London.
- Joly, P.B. and Mangematin, V. 1997b. 'Strategic games of appropriation during scientific production: the example of molecular biology.' Paper to the Conference of New Economics of Science, Notre Dame, USA, March 13-16.
- Lyll, A. 1997. 'Informatics – the key to redesigning drug discovery.' *EBR*, 55-62.
- McMeekin, A., Harvey, M. and Gee, S. 2002. 'Emergent bioinformatics and newly distributed innovation processes.' Paper to the Bioscience and Biotechnology Workshop, Gothenberg, September.
- McMeekin, A., and Harvey, M. 2002 'The formation of bioinformatic knowledge markets: An 'economies of knowledge' approach.' CRIC Discussion Paper 52, July. <http://les.man.ac.uk/cric>
- Oliver, S. 2002. 'Today we have the naming of parts.....' *Nature Biotechnology*, 20, News and Views, 1-2.
- Oliver, S. et al. 1992. 'The complete DNA sequence of yeast chromosome III'. *Nature*, 357, 38-46.
- Parekh, R.B. and Lyll, A. 2000. 'Proteomics as an emerging technology in pharmaceutical R&D.' *Journal of Commercial Biotechnology*. 6, 284-291.
- Polanyi, K. (1957) "The Economy as Instituted Process" in: Polanyi, K., Arensberg, C.M. and Pearson, H.W. (ed.s) *Trade and Market in the Early Empires*. The Free Press. New York.
- Rybak, B. (1968), *Psyché, Soma, Germen*. Gallimard. Paris.
- Rybak, B. (1978), 'Bio-Informatics and Bio-Process in the Physiology of Communication.' *Biosciences Communications*, 4, 3-4, 158-9.
- Senker, J., and van Zwanenberg, P. 2001. *European Biotechnology Innovation Systems. Final Report*. TSER Project (SOE1-CT98-1117), SPRU, University of Sussex.
- Vassarotti, A., Dujon, B., Mordant, P., Feldmann, H., Mewes, W. and Goffeau, A. 1995. 'Structure and organisation of the European Yeast Genome Sequencing Network.' *Journal of Biotechnology*, 41, 131-137.
- World Health Organisation. 2000 *Report on Global Surveillance of Epidemic-prone infectious diseases*. Chapter V. Leishmaniasis and Leishmania/HIV coinfection.' WHO.
- Wood, D.W., Setubal, J.C. et al. 2001. 'The genome of the natural genetic engineer *Agrobacterium tumefaciens* C58', *Science*, 294, 2317-2323.
- Wolkenhauer, O., (2001), 'Systems Biology: The reincarnation of systems theory applied in biology?' *Briefings in Bioinformatics*, vol. 2, no. 3.