Final report of the GARNet Advisory Committee on Arabidopsis Systems Biology in the UK, June 2006.

This report was prepared for the UK Biotechnology and Biological Science Research Council (BBSRC) Integrative and Systems Biology (ISB) Strategy Panel in 2005-2006, by the Advisory Committee (AC) of the Genomics Arabidopsis Resource Network (GARNet). It was informed by the AC members, representing the UK Arabidopsis research community, and by input from international systems biologists. Community input to the AC was gathered in response to email surveys, a newsletter article (appendix 8), two dedicated workshops and three ‘town hall’ meetings. An interim report was submitted in January 2006. Contributors and workshop reports are detailed in section 7.

1. Key recommendations

Biological research using *Arabidopsis thaliana* is adopting systems approaches in some areas. Arabidopsis researchers must build the capacity for systems analysis, if the UK’s strength in Arabidopsis biology is to be reflected in sub-systems models that, ultimately, will contribute to a whole-plant model. The committee’s key recommendations for new initiatives in the UK are:

- A distributed, 4-year PhD training school in Arabidopsis systems biology, with central facilities
- Mechanisms to support early-stage theoretical research involving Arabidopsis research groups
- A £5M networked research activity including multiple Arabidopsis research groups
- Mechanisms to support the acquisition of reference data sets in plant systems biology.

In addition, existing mechanisms should fund networking with theoreticians and with plant scientists modelling at larger scales, demonstration projects in plant systems biology, and the development of new tools and technologies particularly for data acquisition from single, homogeneous cell types.

2. Goals

Ultimately, systems biology will result in the development of a whole-plant model, a ‘virtual Arabidopsis’ or ‘computable plant’. This will aim to account quantitatively for all stages of the Arabidopsis life cycle from seed to seed, together with variations in growth and cellular function due to environmental responses. One proposal sets a target date of 2020 for this goal (www.cmp.uea.ac.uk/ivis; see appendix 1). The whole-plant model will integrate multiple, subsystem models that are appropriate to address different biological questions, from the scale of macromolecular complexes and metabolite pools upwards. It will be based upon observed biochemical and biophysical mechanisms and validated by experimental data at multiple levels.

3. Benefits, challenges, needs

3A. Benefits

The benefits of generating predictive models for particular biological processes are obvious to the community. A virtual Arabidopsis will be a significant and highly visible achievement in biology, with spinoff benefits across multiple disciplines. Unique advantages of Arabidopsis as a platform for systems biology include:

- fundamental plant biology: spatio-temporal models of plant growth at the cellular level will be significantly easier to develop in plants compared to animals, because cell migration and changes in cell shape are very limited in plants.
- status and organisation of the Arabidopsis research community: there is a well-established culture of data sharing and collaborative development of open-access biological and bioinformatics resources, including the underpinning resources for functional genomics.

- extension from the organism to larger scales: a plant model may readily be connected to well-established ecological and crop science models on the field to landscape scale (see appendix 4).

3B. Technical challenges and solutions

Plant systems biology shares many challenges with systems biology in other species, particularly multicellular organisms, which are not rehearsed here. However, data from defined plant cell types will be important for many projects. Data on whole organs or seedlings remains relevant in generating preliminary models and also for modelling subsystems that are similar in all cells. Potential solutions to provide cell-type specific data (see appendix 8) include:

Use of current cell cultures for baseline studies. Cell culture systems can produce diverse data that are readily amenable to modelling, including timecourse data after chemical interventions. These have been used to study plant cell cycle and senescence processes. However, plant cell suspension cultures do not obviously represent a particular cell type in the intact plant, in terms of metabolism, receptor expression or differentiation. Development of stable, differentiated cell cultures would greatly facilitate plant systems biology, but there are alternatives:

Cell purification or extraction of components from defined cell types. Certain plant cells can be physically separated at high purity, including pollen and stomatal guard cells. Other techniques exist to extract RNA, but potentially also protein and metabolites, from single cell types of intact plants: these include laser-capture microdissection, fluorescently tagging and purifying protoplasts of a single cell type, sampling single cell contents using micropipettes, or the extraction of mRNA from polyribosomes that carry a cell-type-specific protein tag. Limited amounts of sample are recovered, permitting transcriptomic assays that include amplification, as well as assays for specific molecular species. Other profiling methods (proteomic and metabolomic) may be limited by the sample amounts, so continued technical development to increase the sensitivity of these methods is desirable.

Establishment of unicellular model species. An experimentally tractable species from the lower plants or green algae, which could be cultured at steady state in a homogeneous or unicellular form, will facilitate some aspects of plant systems biology. Examples include the development of baseline metabolic or signalling models. There is no current consensus on the species of choice. Genome sequencing programmes continue to provide new candidates to supplement the longer-established systems, such as the alga *Chlamydomonas reinhardtii*. Evaluating new model species is desirable.

3C. Structural and organisational challenges

Research in plant systems biology is currently carried out in relatively few biology laboratories in the UK, in diverse geographical locations. Systems for collaborative interaction and discussion will be required to share experiences, solutions and diverse data across the UK Arabidopsis community. The reductionist bias identified at the highest levels of international plant science "constitutes a drag on the adoption of systems biology by the research community" (see 7). BBSRC's initiatives, including the preparation of this report, have contributed to mitigate this bias in the UK. However, BBSRC's aims in translation to crop research could conflict with the necessary focus of systems biology upon a model organism (see also 4Bi, below). Most importantly, many more theoreticians will be required in plant science to achieve any large-scale goals in Arabidopsis systems biology.

3D. Examples of needs perceived by the community (see also appendices 2-6)

1. Interdisciplinary training, especially of PhD students, was emphasized above all other items. See 4A below.

2. Access to 'biologist-friendly' theoreticians. See 4B below.
3. Technical support for modelling, in the form of a helpdesk or service facility. Given the diversity of modelling approaches required, such generic support cannot easily be provided.
4. Focus on biological questions rather than technological approaches. See 4C below.
5. Standards for experiments and formats for data/model exchange. Partly addressed in 4C. GARNet could participate in a coordinating role.

4. Near-term objectives

In the Advisory Committee (AC) view, three processes are paramount to achieve a leading international position in Arabidopsis systems biology, in addition to suitable funding levels and timescales:

1. Increased interdisciplinary training and theoretical expertise in Arabidopsis systems biology.
2. Widening the adoption of systems-level methods among Arabidopsis biologists.
3. Coordinating the UK community's research to provide coherent input to international projects.

Each of these aims should be addressed by multiple mechanisms, including those below. Existing funding for networking and discipline-hopping, for example, should be extended. GARNet's role in facilitating the adoption of functional genomics approaches was noted by the community as an example to follow. GARNet can also facilitate some recommended actions here, as noted below.

The AC's interim report covered two objectives, demonstration projects and tools/technology development, that are not included below. Each received community support but they can likely be achieved using existing (possibly expanded) funding mechanisms. Technology development is required particularly in the areas highlighted under 3C above. The AC recommends the following new initiatives, which were proposed in the interim report. They have since been supported by the community and refined based upon community input (especially the proposed network, item 4C):

4A. Interdisciplinary teaching, training and re-training.

A national, 4-year PhD training programme in Plant Systems Biology is proposed, similar to an EPSRC/BBSRC Doctoral Training Centre. Students would be based across a distributed network of Arabidopsis biology laboratories with appropriate, preferably local, theoretical co-supervisors. A training hub should provide computing support, some theoretical training for supervisors as well as students, and networking among groups. Supervisors without current systems biology funding must be included to broaden participation. Possible mechanism: ISB panel funding channelled via the Studentships and Fellowships panel, co-funding with EPSRC.

4B. Networking with the physical sciences, crop science and ecology

i. Researchers in crop and ecosystem modelling can potentially offer research expertise and validated models. Many such models use complex traits (e.g. nitrogen use efficiency) as simple input parameters. An Arabidopsis systems biology project that "deconvolved" one of these traits in terms of the underlying molecular networks might thereby connect Arabidopsis systems models to models of plant growth on the field to ecosystem scale. A database of relevant models, translated into a common format, would facilitate such translational research and might be combined with a summer training school in modelling methods (initiative of Prof. H. Thomas and Drs. H. Ougham and A. Gay (IGER), with GARNet).

ii. 'Study Groups' provide a week-long opportunity for theoreticians to tackle specific problems in an area of science, to give an informed view of the opportunities and approaches for further research. A regular study group on Mathematics in Plant Science should be established (initiative of Dr. Marcus Tindall (Oxford Maths), with GARNet).

Each of these mechanisms will broaden participation and attract theoreticians to Arabidopsis research.

4C. A co-ordinated research network in Arabidopsis Systems Biology

From Phase 3 systems biology funding, the UK could realistically contribute to at least one ambitious goal of lesser scope than the whole-plant model. The CPIB's focus on the plant root is one such project. A likely structure for a networked activity would comprise a shared core or coordinating component, with the majority of activity in 4-5 interdisciplinary collaborations. Each collaboration would consist of about 3 partners, with a mix of Arabidopsis biology groups, other experimentalists, and theoreticians, focussed on a particular biological problem. The core component would provide project management and shared infrastructure, probably with more involvement in modelling and bioinformatics than experimental facilities. The core might, for example, be responsible for establishing common standards. Applications for the coordinating functions could be separate from the research network, though a complete project that comprised both the core and the research collaborations would be better. The minimum level of funding required to make this an internationally competitive activity is ~£5M over 5 years (ca. 10 staff, including 6 post-doctoral research associates and 2 technicians). This assumes that some of the participating groups have additional, directly relevant project funding and/or substantive input from external collaborators. Increased core activity or generation of a large data set would increase this figure. The 5-year duration is also a minimum.

A research network in Arabidopsis systems biology should focus principally on a biological question. Though technology development or resource provision may be a natural component of some projects they should not be the major focus. Our consultation did not establish a consensus on any biological question: different groups supported various goals, including Systems Biology of "The Leaf" and of "Plant Growth". The conclusion was that the biological topic should not be prescribed. The applicants should have a single, clear goal to promote visibility and the projects in the network should be closely related to enable meaningful integration of the resulting models. In addition to the ISB guidelines for systems biology projects, the suggested guidelines for selecting a research network in plant systems biology are that the network should:

1. Focus on *Arabidopsis thaliana*.
2. Address biological topics that are clearly related, producing a subsystem model for each one.
3. Integrate the subsystems models for each topic into a coherent, larger model.
4. Validate the resulting model, for example using the methods of synthetic biology.
5. Coordinate with international projects to ensure that the resulting models are integrated into a larger framework, most likely a whole-plant model.
6. Incorporate some existing UK systems biology groups, to ensure rapid startup.
7. Nucleate the adoption of systems approaches in additional groups, both among and beyond the initial applicants.

Coordinating research in a dispersed network of laboratories will have significant benefits in addition to the added value in direct research outputs. The network will broaden the coordination and standardisation of experiment, data and model formats through the community more effectively than a single centre, facilitating development of public repositories and maximizing the value of data and models from many groups. The visibility of a large-scale project with a clear goal should facilitate interactions with other research communities (e.g. increase engagement by theoreticians), funding agencies, policy-makers and the public.

It is common experience that the most productive multidisciplinary interactions take several years to develop. The committee has learned of UK Arabidopsis laboratories at various stages of such interactions, often with limited or no funding specifically for systems biology. This suggests that there will be sufficient collaborations to support competing proposals for a network in plant systems biology.

4D. Baseline data sets underpinning Plant Systems biology

Global (genome-/ proteome-/ metabolome-wide) data sets, together with the biological resources to produce them, the informatics to distribute them and the pathway or network models from their analysis, are the starting resources for hypothesis generation as well as hypothesis testing in systems biology. Resources that are provided commercially for mammalian systems biology, such as pathway databases, risk being neglected for plants so public funding will be required in this area. Possible mechanism: the ISB Panel and the Tools and Resources Panel should jointly establish ways to fund the new resource provision for Systems Biology.

5. Contribution of NASC, CISBs and international comparisons

A networked research activity should make best use of the bioinformatics resources and materials at the European Arabidopsis Stock Centre (NASC), and those to be provided by the BBSRC Centres for Integrative and Systems Biology (CISBs). The CISBs already provide good contacts to international alliances in systems biology: for example, CPIB intends to collaborate with the large-scale FP6 project "Agron-omics" that focuses on the developing leaf. The scale of the proposed network could be comparable to that project. Only by incorporating significant additional personnel and resources from the participating laboratories could it approach the scale of centres in continental Europe that are strongly engaged in plant systems biology, such as the Max Planck Institute of Molecular Plant Physiology in Golm, Germany, or the Department of Plant Systems Biology in Ghent, Belgium. The UK network's initial funding would be larger than the NSF FIBR grants (US$2-3M) awarded to interdisciplinary collaborations focussing on particular questions in plant systems biology, such as the 'computable plant' project focussing on the shoot meristem (led by Mjolsness, UC-Irvine and Meyerowitz, Caltech), but the Computable Plant is growing in scale with additional international partners. Looking ahead, a recent NSF workshop considered the establishment of an interdisciplinary 'Synthesis Centre' or 'Cyber-Infrastructure Centre' in plant science, which would provide core support and technical development in informatics and modelling rather than focussing on a biological question.

6. Conclusion

Multiple initiatives will be necessary to develop Arabidopsis systems biology in the UK on the scale required to tackle the challenge of a whole-plant systems model. A distributed research activity, including a small fraction of the 200+ Arabidopsis research groups in the UK together with a strong cohort of interdisciplinary collaborators, could establish a leading position in plant systems biology if it were suitably coordinated.

7. Contributors, Acknowledgements, References and Reports.

The AC is grateful to BBSRC for funding this report and the associated meetings, to many participants at the meetings, and to individuals who have contributed their time and expert opinion, including: Steve Russell, Elliot Meyerowitz, Chris Somerville, Pierre Hilson, Mark Stitt, Stefan Hohmann, Igor Goryanin, Hanspeter Herzel, Boris Kholodenko, Hiroaki Kitano. We are grateful to Prof. Thomas and Drs. Ougham and Gay (IGER) for organising the Swindon workshop, and especially to the GARNet administrator, Dr. Ruth Bastow, for organising and reporting on the other actual and virtual committee meetings and the various workshops.

GARNet Advisory Committee members (for interim report): Ian Furner (Chair), Philip Gilmartin, Julie Gray, Claire Grierson, Nick Harberd, Jonathan Jones, Marc Knight, Ottoline Leyser, Keith Lindsey, Andrew Millar (Co-ordinator), Simon Turner, Sophie Laurie (BBSRC contact). For the final report: Brendan Davies, Paul Dupree, Zoe Wilson, Clare Rushowski (BBSRC contact).

References and appendices

(most are also accessible online at http://garnet.arabidopsis.info/garnet_reports.htm)

1. Progress report from Prof. Andrew Bangham on the *In Vivo-In Silico* Grand Challenge proposal, May 2006. (attached, page 7)
2. Town Meeting summary from SEB Annual Meeting, Canterbury, 5 April 2006. (attached, page 9)
3. Workshop report from Systems Biology Forum, John Innes Centre, 20 March 2006. (attached, page 10)
4. Workshop report from "Interfacing Systems Biology with Crop and Ecosystem Modelling", Swindon, 13-14 March 2006. (final report is attached, page 11; online version includes annexes)
5. Town Meeting summary from the New Phytologist meeting on Plant Networks, London, 27 January 2006. (attached, page 16)
6. Report from 'Single Cell Technologies' satellite workshop, GARNet annual meeting, Norwich, September 2005. (attached, page 18)
7. Workshop report from BBSRC workshop "Succeeding in Plant Systems Biology", Edinburgh, 25-26 July 2005. (online only)
8. World Techology Evaluation Center (WTEC) report on International Systems Biology, 2005, available online at http://www.wtec.org/sysbio/report/SystemsBiology.pdf

A briefing article, soliciting input for the report, was published in the November 2005 issue of GARNet's community newsletter, GARNish, which is distributed to most UK Arabidopsis laboratories in hard copy. It is attached as appendix 8, page 19 and is also available for download at http://garnet.arabidopsis.org.uk/garnish_nov05.pdf

Briefing materials were prepared for participants at each of the workshops, summarising the progress to date on the interim report and previous meetings. A sample briefing pack for the JIC meeting is attached as appendix 9, page 20 and ff.

The GARNet AC's interim report, submitted to BBSRC in January 2006, is included at the end of appendix 9, and is also available online at the address above.

Appendix 1.

Progress report from Prof. Andrew Bangham (University of East Anglia) on the *In Vivo-In Silico* Grand Challenge proposal, May 2006

"Report on the EPSRC, British Computer Society, Grand Challenges in Computing meeting, Glasgow, Spring 2006

In Vivo – In Silico (IVIS) lays down a major challenge to computing scientists: to model living organisms. It is a pressing problem that is becoming just feasible. Following the success of the Human Genome project, techniques in biology are yielding huge amounts of data from which computational models of yeast and other single celled organisms are likely to be derived. From a human viewpoint, however, it is higher organisms that are particularly interesting. It is possible, even likely, that the hierarchical spatial organisation of such differentiated organisms will allow the modelling of the key high-level systems with no more dependence on the detail of the low-level cellular processes than the design of a Unix operating system has on the type of processor on which it is running. Moreover, an understanding of these high-level, spatially organised, systems is likely to impact the understanding of computing itself.

From an initial focus on the worm, *Caenorhabditis elegans*, attention in IVIS at the 2004 conference (GCC04) shifted towards higher plants. (Not exclusively as there will be advantages in simultaneously studying simpler but nevertheless structured systems like Streptomyces.) Plants have several advantages. They appear more tractable than animals. Already there are many individual computational models covering a range of scales from crops and ecologies through plants to metabolic processes, gene regulation networks, and reaction diffusion/polarisation systems. For example, L-systems have already successfully captured the developing topology of plants and these computational concepts are also impacting computational science itself. IVIS is a challenge to computing science, posed by biology, the solving of which will benefit both.

More importantly, understanding plants is crucial to mankind. Plants are our primary food source, a primary energy source, a source of medicines and are vital to the environment.

Biologists will need to be involved. Since GCC04 therefore, the major activity has been to establish common ground with biologists. IVIS now has support from biologists studying crops, plants, plant metabolism, genetics and gene regulation. Recent technical developments in genomics and metabolomics are being augmented by new methods for uncovering spatio-temporal patterns of activity and this will provide the information required for spatio-temporal modelling. The need for IVIS was anticipated in the US Arabidopsis 2010 proposal and is seen in projects such as the 'Computable Plant' in California. However, relevant biological research is a particular strength of the UK and this is reflected in recent meetings organised by, for example, GARNet (Genomic Arabidopsis Resource Network, exemplar for crops) and IGER (Institute of Grassland and Environmental Research, e.g. biofuels). In that community there is a strong interest in a 'Plant Computer Models' repository, in 'Computer Model based Spatio-temporal' databases, in new models of plants growth and behaviour from the 'Crops and Ecology to DNA and Back' and a framework for collecting and linking them together. Such projects are timely and complement recent investments by the BBSRC in Systems Biology.

The importance of a goal and a deadline is embedded in a new running title, In Vivo – In Silico: a 2020 vision of modelling living processes. The meeting identified the main challenge to computing to be creating a framework that handles subsystems over scales ranging from molecules to populations and developing interfaces between sub-systems (models). Although physical and chemical constraints on ‘energy systems’ such as crops, metabolism and fuels are understood, biological ‘information’ (control or signal) systems, for that is what biological systems are, require further theoretical and experimental insights. For example, new concepts will have to handle not only systems that behave correctly but they must also grow with no explicit blueprint (i.e. they must be autopoietic) and by self-replication and selection be honed by evolution.

IVIS is a major challenge, but initial forays are already underway both in biological and computing science, e.g. biological clocks, growth models and in artificial life projects. The next step is to set up a network to enable collaboration on an outline for a long-term research programme.

Appendix 2.

GARNet Town Hall Meeting (Summary notes)

5th April 2006, University of Kent, Canterbury (following SEB session 'Genomes to Systems')

Questions/Comments were invited on the GARNet Interim Report on Plant Systems Biology, on Biological Topics for a Networked Research Activity, and on Tools/Technology required (format as for 27 January 2006, with more focussed questions) – Presented and Chaired by Andrew Millar

30-35 participants including theoreticians and plant scientists with a range of interests.

Discussion covered many of the topics already raised in the January Town Hall Meeting and elsewhere, which are only summarised here:

1. Training is urgently required, especially of interdisciplinary PhD students (earlier training, e.g. undergraduate courses or summer schools, Masters courses, were also desirable).
2. Mechanisms to help plant scientists to access theoretical expertise were discussed, also provision of technical assistance service for modelling. The proposed Study Group in Mathematical Plant Science was supported.
3. Species used in plant systems biology: focus on Arabidopsis is necessary, but crop species will also benefit and should not be forgotten.
4. Biological topic for a large-scale activity in plant systems biology. Participants were keen that the topic should be as inclusive as possible. The 'Plant Growth' and 'Evolution/Adaptation' topics suggested from the JIC workshop were each supported by some participants, as was 'The Leaf'. The comparison to the proposed CPIB focus on the root was noted. AJM outlined the small number of staff that could actually be employed in a realistic Phase 3 proposal (~£5M). Focus would be necessary for visibility and to link models from sub-projects.
5. Organisation of large-scale research network. The proposed 'hub and spokes' structure was supported. Participants suggested a more federated structure than previously discussed, with a strong coordinating function in the 'hub', perhaps funded separately, and more independent 'spokes'. EU FP6 'coordinating actions' were mentioned as a possible funding model for the hub.
6. Development of standards for experimental plant research and for data exchange was seen as a potential benefit from a large network. Participants called for more data sharing methods, either a central data repository or a decentralised system with a single access point, e.g. via Moby-style web services. Coordination by NASC and/or CISBs and/or GARNet was suggested.
7. The model repository suggested at the Swindon workshop was supported.

Appendix 3.

John Innes Centre (JIC) Systems Biology Forum (Summary Notes)

An informal meeting to discuss the GARNet interim report on Arabidopsis systems biology in the UK was held at the JIC on the 20th March 2006.

JIC, UEA and IFR staff were invited to attend and provided with background information about the report (JIC Q&A V4.pdf).

Present at meeting:- Jonathan Jones, Nick Harberd, Ian Bancroft, Mike Bevan, John Turner, Andy Maule, Caroline Dean, Jan Kim, Mike Ambrose, Alison Smith and David Studholme.

Attendees were asked for guidelines to aid selection of a community wide project in plant systems biology

Guidelines suggested by GARNet

- More than one group
- Interdisciplinary; mix of wet and dry
- Enabling projects, foundation for further responsive mode funding

Guide lines suggested by JIC

- Transparency of criteria, panel members and abstracts of funded projects
- Hypothesis/question driven science preferred
- Diverse approaches to systems biology
- "Nucleation capacity" ; future proposals can "join in"
- ERA-PG approaches to systems biology to be encouraged
- Exemplars to be included in remit; e.g. How do plants grow? Evolution and adaptation

Comments

Projects should be organised and managed to facilitate longer term studies in systems biology, e.g. 5yr grant terms

Suggestions for "nucleation capacity" and modular format so others can join in and access expertise, personnel and data

All data produced should be readily available and in a standard format for others to download and use. Need to address cross-referencability of data and interoperability of databases

Many thought that current suggestions of an organ or pathway were not inclusive and appeared to be centred on data collection rather than a science/hypothesis driven. Big biological question such as plant growth would be more inclusive and accessible to the community.

Appendix 4.

Report from *Interfacing Systems Biology with Crop and Ecosystem Modelling* workshop, Swindon 13-14 March 2006.

SUMMARY

The main objective of the *Interfacing Systems Biology with Crop and Ecosystem Modelling* workshop was to bring together scientists from the systems biology, bioinformatics and agro-ecological modelling communities to exchange ideas, enhance awareness of each others' fields, explore synergisms and make recommendations on fruitful future directions.

The workshop was open to all UK plant scientists with an interest in using system-based approaches in crop and ecosystems modelling. The meeting was fully subsidised by GARNet and was organised in consultation with Howard (Sid) Thomas, Helen Ougham, Alan Gay of IGER and the co-ordinators of GARNet (UK Network for Arabidopsis functional genomics). Thirty eight participants attended the workshop, including speakers and staff.

This report will be communicated to BBSRC's Integrative & Systems Biology Strategy Panel and made publicly available on the download section of the GARNet website.

1. BACKGROUND

a) The BBSRC and EPSRC launched its first call to establish Centres for Integrative Systems Biology (CISB) in 2004. This initiative funded three centres in Newcastle, Imperial College and Manchester. However the ISB panel were concerned to find that relatively few plant biology-based applications were received to this first call and concerns were expressed that none of the plant-based applications was of sufficient quality to be short-listed.

b) To help address this problem BBSRC and GARNet held a workshop entitled 'Succeeding in Plant Systems Biology' in Edinburgh during July 2005 to look into the issues surrounding the application of systems biology in plant science research. A full report of this meeting is available at
http://www.bbsrc.ac.uk/about/gov/panels/isb/docs/PSB_workshop_report_sept05.html

c) During the Edinburgh workshop it was brought to GARNet's attention (by Sid Thomas, IGER) that there are of a number of crop modellers in the UK, with expertise in this area, who could provide valuable insights into this new field.

d) To facilitate the involvement of crop scientists in plant systems biology, GARNet invited Prof Sid Thomas and colleagues at IGER to organise a workshop, that would bring together modellers, bioinformaticians and plant researchers, to enhance awareness of current research and discuss future involvement in systems biology.

e) During the summer of 2005 BBSRC asked the GARNet committee to investigate how systems biology could best be approached in UK Arabidopsis research. GARNet has submitted an interim progress report on this area in January and output from the Modellers workshop will be included in GARNet's final report to the ISB panel later this year.

2. WORKSHOP PROGRAMME

a) Programme content was decided upon via consultations with Sid Thomas, Helen Ougham, Alan Gay of IGER and GARNet. The programme was split into several sessions with time included for in-depth discussion during the breakout sessions. The workshop programme is provided in Annex 1.

b) Each delegate was invited to submit an A4 poster describing his/her current work to promote awareness of research interests represented at the workshop and assist future interactions. Copies of these posters were provided in the delegate book (Annex 2).

c) Attendees were divided into three groups (Annex 3), asked to consider the same four questions below and produce a poster of responses. Each group was of mixed disciplines and assigned a chair to supervise discussions.

1. How can systems biology build on established and ongoing modelling approaches?
2. Are there technical or other barriers in applying modelling techniques to plant systems biology?
3. How can systems biology contribute to the development of models and modelling techniques?
4. What are the common targets for systems biology and modelling in plant and environmental research?

d) Delegates were also provided with a summary of BBSRC's position on Systems Biology from the ISB panel (Annex 4) and an article on Systems Biology from the New Phytologist journal (Annex 5).

e) During the final plenary session of the workshop each group presented a summary of their discussions to all attendees. This was followed by a general open forum discussion of all issues raised during the two days.

3. OUTCOMES

a) Summaries of outcomes to each of the breakout question are provided here. These are based on a comprehensive table of responses, provided in Annex 6.

b) Two groups translated question 1 into what lessons could systems biology learn from current modelling approaches.

c) Two groups also added an additional question, what is systems biology?

Q0 What is systems biology?

- Integration and scaling
  - Covers scales that range from molecular to ecosystem
- Addresses:
  - Fundamental plant processes
  - Crop, form, function and efficiency

Q1a Lessons to be learnt from other systems?

Common themes

- Expertise in handling large amounts of data and complex models.
- Parsimony; Model Simplification
- Experience in dealing with:-
  - Stochasticity
  - Scaling up
  - Transient and disturbed systems
  - Parameterisation
  - Model testing and validation
- Appreciation of environmental components and design of field experiments

Q1b How can systems biology build on established and ongoing modelling approaches?

Common themes

- Systems Biology can supply evolutionary and developmental aspects not typically addressed by classical dn/dt modelling
- Classical modelling treats plants as machines driven by the environment; systems biology can add sensing and signalling to these models
- Systems Biology can provide a spatial perspective i.e. dn/dxyz

Q2 Are there technical or other barriers in applying modelling techniques to plant systems biology?

<u>Common themes</u>

- Inability to profile at single cell level
- Volume of -omics data
- Management, quality control & interrogation of metadata
- Lack of data standards
- Lack of measurement and estimation of key biochemical parameters
- Lack of common language
- Appropriate training

<u>Additional comments</u>

- Lack of commercial pull through; where should these models be focused?
- Lack of bespoke software

Q3 How can systems biology contribute to the development of models and modelling techniques?

<u>Common themes</u>

- Tools and theories for scalability; particularly needed for non linear systems, stochasticity and uncertainty
- Approaches for continuous and discrete model building
- New technologies e.g. bioinformatics and spatial visualisation
- Experience in data collection and handling

<u>Additional comments</u>

- Deconvolution of classical traits
- Tools to introduce evolutionary and developmental processes into the modelling package
- Prediction via systems biology will reduce the number of field experiments that need to be undertaken

Q4 What are the common targets for systems biology and modelling in plant and environmental research?

<u>Common Themes</u>

- Organism -organism interactions e.g. parasites, symbionts
- Durable pest resistance
- Sustainability
- Non-food uses of crops e.g. biofuels
- Evolution and development

General Conclusions from Discussion Forum

<u>Common Themes</u>

a) Data Standards

Appropriate ontologies and standards will be needed to deal efficiently and effectively with the data generated.

The greatest challenge in this area is likely to be in the field: new tools or technologies may need to be designed to collect the appropriate information at source rather than setting standards after data capture.

MIAME (Minimal Information About a Microarray Experiment) is just that - minimal, need more information not less, one answer to this problem might be to collaborate in large groups to facilitate data collection en masse under standard conditions.

b). Data Storage and Submission
Need to encourage researchers to submit data to a central resource or make data publicly accessible via institution or personal website. Incentives will be required to ensure this occurs, for example if data submitted to a public resource is used by others this needs to be highlighted and recognised by the community, by internal and external research evaluations and by funding bodies.

Both a) and b) also apply to models: a major outcome of the workshop were proposed mechanisms to adopt common standards and a model repository.

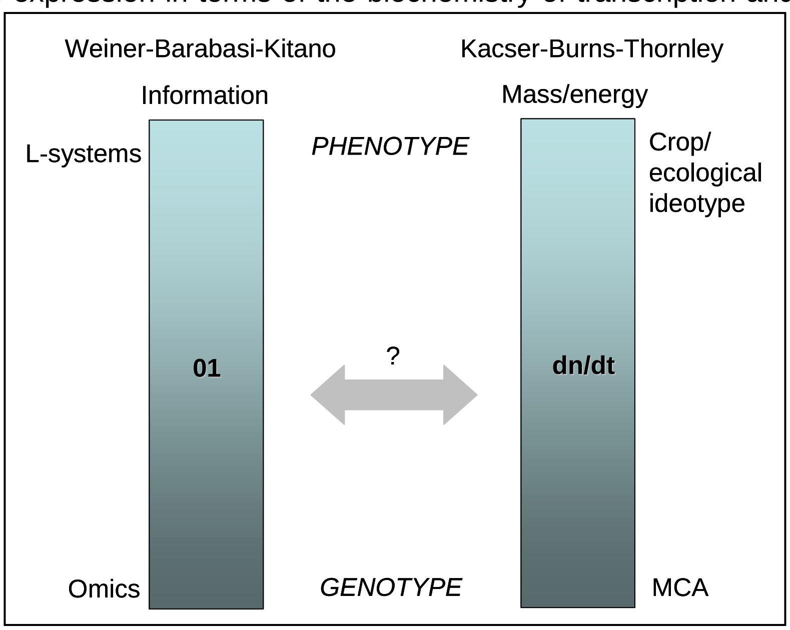
c) Model Focus
Models need to be aligned to practical problems with an economic impact. It would be advantageous if some Arabidopsis based models were centred on complex traits of interest and use to crop science. This would provide a focus for Arabidopsis systems biology and help link research in model species to that in the field.

If crop science models are to have a productive end point then ultimately GM issues will need to be considered. Instead of looking into yield or harvest index perhaps the focus should be changed to public good for example, diet, health and bioenergy.

4. MAIN FINDINGS

a) Emergent Properties
Qualitatively new behaviour emerges from increasing complexity as one moves up the scales. Predicting these new properties is a major challenge for modelling. At least as common as emergence is disappearance ("negative emergence") of hitherto significant processes above a certain threshold. In some cases (e.g. at catchment or landscape level) biology itself may shrink to insignificance as a contributor to model parameters. This is familiar to the crop/ecological genetics community as the propensity of the E term to dominate the G x E interaction.

b) Information versus Mass Energy
"Conventional" or "classical" process modelling in crop science and ecology typically takes the form of differential equations based on rates of mass or energy change (*dn/dt*). This approach is applicable at all levels from genotype to phenotype. At the molecular level metabolic control analysis (MCA) can model biochemistry, even including gene expression in terms of the biochemistry of transcription and translation.



By contrast, systems biology *sensu* Kitano could be thought of as based on information flows, decision networks and living organisms as cybernetic entities. Omics and bioinformatics provide the molecular-level modelling definition. Formalised conventions such as L-systems allow scale up to phenotype modelling.

It is not clear whether the *dn/dt* and cybernetic streams are truly distinct. It seems reasonable to expect emergent properties if the two approaches can be reconciled.

5. PRACTICAL SUGGESTIONS

a) Plant Specific Model Repository

Central resource for researchers to submit current plant based models for others to use and test. These models could be open to community review and annotation either via the web or at an annual event such as summer school where models would be tested and validated as part of a training exercise.

b) New Phytologist Meeting 2007.

A proposed follow-up to the present meeting is being considered by New Phytologist for the 2007 event in its annual Symposium series.

6. ACKNOWLEDGMENTS

Thanks to Sid Thomas, Helen Ougham, Alan Gay for organising the scientific programme for the workshop, and securing the participation of invited speakers; to Chris Rawlings, Sid Thomas and Malcolm Bennett for acting as Chairs of the breakout sessions; the New Phytologist for reproduction of the article on Systems Biology; and Enfys Lloyd and Suzy Shipman for organisational and web support.

Appendix 5.

GARNet Town Hall Meeting (Summary notes)

27th Jan 2006, Linnaean Society London

General Questions/Comments on Interim Report – Presented and Chaired by Andrew Millar
How complete does the model have to be?

Will it feed in to an international effort, if so who is doing what where?
Answer:- Single Institutions in the EU such as Ghent and Golm and in the US several FIBR grants such as Elliot Meyerowitz, computable plant.

The community should decide upon a single goal e.g. whole plant by 2017 and use what ever technology is available to us, to meet the goal across the whole of the community rather than let the technology determine what we do.

The leaf as a model was considered to be good proposal by a proportion of the audience including David Fell ISB Panel Member, who suggested that a single goal specific to plant would make it stand out and compete well against projects from other communities.
However a number of the audience were equally apposed the idea of a leaf and suggested other area such as homeostasis or fields linked by a generic technology.

There was general consensus among the audience that training was an essential part of any proposal.

David Fell reminded the audience that BBSRC can not allocate funds for PhD training but the EPSRC are able to supply funding in this area.

Engineers and Mathematicians present suggested that a 4yr PhD would not be attractive to them as this is rather a long time compared to the average PhD.

MSc courses were accepted as a useful way to assess students abilities in this area; exposing students to both theoretical and life sciences and providing experience of the problems they will face in the future. This type of approach is often not possible in a 3yr PhD that is focused on practical outputs rather than theoretical training.

More access for all lab members to attend taught University undergraduate courses.

Possible involvement in Marie Curie or other European programmes

Training will be an essential part of any centre that is created by Phase two of the CISB funding and any future plans will need to take this into account.

As with all training each person learns in a different way so no one course will be useful for all. Instead we should support a range of training programmes from taught courses to practical workshop and hotel facilities in relevant centres/institutions

Maths study groups with Medics at Nottingham have been very productive with 11 out of the 13 projects gaining further funding. This may be a good way to pump prime a project.

Many of the current Arabidopsis systems biology projects are at a very early stage and need pump priming before they can go any further. All this often requires is money for a theoretician to sit down with the biologist and discuss the pros and cons of a project. Often people are interested but don't have the time to spare. This type of approach can be covered by the BBSRC discipline hopping scheme

Suggestion for a service centre where people could send there data for theoretical treatment akin to NASC arrays for your micro array data

GARNet provided easy access to genomics technology in the past suggestions for a similar level of generic support for systems biology so that it can be taken up by the community as whole.
Mechanisms to achieve this could included

A retreat
Summer School for students
Centre to which labs can send personnel for training

Research Areas - Presented and Chaired by Simon Turner

Is it really feasible to find one single goal that we can all sign up to and would benefit from?
Unlikely instead a set of guidelines to choose a project would be useful.

David Fell stressed that what ever project or idea we put forward it must address an important scientific question, have some outline of what it will achieve and an understanding of the appropriate level of modelling.

It is possible for people to group together now and obtain systems level funding from current BBSRC responsive mode funding. So instead of personal goals we should be thinking of something that can offer outputs to both plants and other communities

The 30M in phase three is not going to fund a large project therefore we should not try to be inclusive as that is not feasible at this stage instead the project should have flexibility for further projects and spin outs for the future.

David Fell agreed that these proposals should be enabling and used as a foundation for other projects to build on in the future from responsive mode funding. Possible suggestions could include a coordinated approach to collect a data set e.g. micro array data under specific conditions that we could all use and benefit from.

Tools and Resources – Presented and Chaired by Keith Lindsey

Missing basic information and physiological data about plant functions is essential for a systems approach we should therefore work towards filling these gaps before we move onto obtaining new data.

Tools or technologies that are specific to plant biology should be provided by this type of community proposal. Examples might include looking at specific compartment or tissue or perhaps movement of transcription factors between cells.

David Fell suggested that we find tools distinct for system biology and should consider data access and curation e.g. a common data repository.

Localization of proteins/macromolecules activities in 3D

Appendix 6 - Workshop report from BBSRC workshop “Succeeding in Plant Systems Biology”, Edinburgh, 25-26 July 2005. (online only).
Reference 7 - World Techology Evaluation Center (WTEC) report on International Systems Biology, 2005, available online at http://www.wtec.org/sysbio/report/SystemsBiology.pdf

Appendix 8

Summary of discussion on single cell technologies at Garnet meeting Sept 2005

Present: J Jones, S Turner, K Lindsey, A Millar, P Gilmartin, J Doonan, C Turnbull, A Hall, Raju Datla, J Milner

The following points emerged.

Systems biology in yeast and other organisms is greatly simplified if only one cell type can be analyzed. However, even in yeast cultures, there can be considerable heterogeneity between cells in a culture.

In mammalian cells, when one cell type is analyzed, a gradual increase of an mRNA can be detected in response to an elicitor, but even this can be misleading, because when responses are monitored in individual cells, a wave-like response is detected, which is lost by lack of synchrony when responses of many cells are pooled.

Laser capture microdissection is difficult to apply to a single cell type because of yield and because of the need for fixation (unless cryosectioning can be used). However it can be applied to tissues (meristems, or specific regions of an embryo).

Solutions can be envisaged for cell types in the leaf and other tissues for transcription profiling; however, solutions do not currently exist for cell-type specific proteomics and metabolomics.

Enhancing detection sensitivity with new fluorescent dyes or tags to proteins of interest would be useful to follow their abundance and location in single cells over a time course.

The BBSRC seeks to encourage plant science into a more quantitative and “systems” level of analysis. This is a good aspiration; we all need to place our work on specific problems in the context of the biology of the whole organism.

However, for this aspiration to be realistic, we need considerable technology development. The Institute for Systems Biology in Seattle deploys much of its resources on technology. Garnet strongly encourages BBSRC to include support for technology development in its funding for systems biology in plants. Among possible targets for such technology development, establishing systems for monitoring the properties of and changes in single cell types is important.

Systems biology meets Arabidopsis

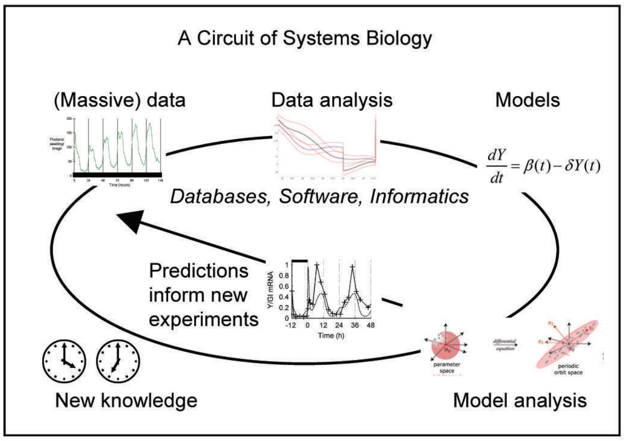
Systems biology has a grand vision – understanding all the components of a biological system and their interactions, across all relevant levels of organisation. In the plant context, the vision might translate into a whole-plant computer model that accounts for seed-to-seed development and environmental responses, starting from the scale of macromolecular and metabolic processes. GARNet's Advisory Committee is currently considering how this could best be approached in UK Arabidopsis research. BBSRC's strategy panel on Integrative and Systems Biology asked GARNet to produce a report that incorporates the community's views on this area for June 2006, with an interim report in January 2006. This article sets our questions (below) in context.

So what is Systems Biology?

The ISB panel has avoided an exclusive definition of Systems Biology, preferring to list the characteristics of a systems approach and noting that the balance among these will vary among projects:

- An integrative approach to the subject
- Large experimental data sets (we would add, of high quality)
- Predictive capability based on modelling
- A mix of inputs not only from across biology but also from the chemical, physical, engineering, mathematical and/or computational sciences.

The result is an iterative interaction between experiments and modelling. Informatics infrastructure underpins the approach as much as experimental methods. Scalability is a major advantage. Mathematical models make it possible to analyse and understand much larger data sets than the simple, logical models that are implicit in all scientific reasoning. Modelling from preliminary data is an excellent way to inform experimental design, to ensure that later experiments provide data that are suitable to test and extend the models.



These guidelines are broad enough to cover everything from data-led ("omics") projects that generate new hypotheses, to hypothesis-led projects that quantify and test current understanding. For example, the inference of preliminary network models from high-throughput data (not only transcriptomics – see Sachs et al. 2005) can be data-led, while the development of segmentation models for Drosophila was more hypothesis-led (von Dassow et al. 2000). In plant biology, examples would include Birnbaum et al. 2003 or Hirai et al. 2004, compared to Jonsson et al. 2005 or Locke et al, 2005, with Rolland-Lagan et al. 2003 between them. There is no dichotomy between these approaches: large data sets contribute to testing existing hypotheses and generate future hypothesis-led projects, perhaps on a broader scale than before.

Systems Biology in Arabidopsis

Systems biology in any multicellular organism can face additional problems, for example the need to acquire data from specific cells or cell types. Arabidopsis will be harder to work with than yeast for many approaches, and including other plant species risks diluting the effort. Relatively few theoreticians are working on plant science questions and we have few models to start from. On the other hand, the limited cell movement and cell shape changes make development in Arabidopsis easier to model than in many animals.

The Arabidopsis community has great strength to offer but our 200+ laboratories are dispersed around the UK. Systems biology has worked best when there is a close and continuous interaction between researchers with different backgrounds. BBSRC has funded three Centres for Integrative and Systems Biology (CISBs, in Newcastle, Imperial College and Manchester) to establish this type of joint research environment. Further institutions are now shortlisted for up to another three CISB awards, to be announced in early 2006. GARNet's report will provide input to "Phase 3" of systems biology funding (ca. £30M): this will not fund CISBs but might include coordinated activity across multiple locations.

Our questions to you

In addition to comments on the above, we are seeking input on these questions. Earlier versions have been circulated by email:

1. What concrete steps would most facilitate systems biology approaches in the UK Arabidopsis community? Examples might include: a national 4-year PhD programme with shared training and computing support; 'study group' events that allow plant scientists to present their research questions to theoreticians for preliminary analysis; a set of demonstrator projects that show what can be accomplished by the systems approach in plants; a 'virtual centre' that links interdisciplinary teams at different sites working on complementary topics.

2. Agreeing a single research goal might have significant advantages. What would be a suitable goal for a coordinated, national project? Examples might include: systems biology of the leaf, of the guard cell, of the auxin signalling pathway. For comparison, the systems biology centre in Cambridge proposes to have 100 researchers focussing on the Notch and wnt signalling pathways in Drosophila.

3. One aim of a large-scale systems biology project could be to collect a particular data set or to provide a particular resource in support of a national project, in addition to gathering researchers with the relevant expertise. Which data sets or resources would convince you to work in a new area, or which could you contribute to providing? Examples might include: a protein-protein interaction map, completion of the metabolic map, imaging of protein localisation throughout development, integrating biological databases, informatics that automates the link between data and models.

4. Which international or industrial partners would you be most interested to work with, and do you already have links to them in systems biology ?

Please respond either to ruth@arabidopsis.info, to one of the committee members, or in person at one of the town meetings.

Town Meeting Dates

1. 27th January 14:30-16:30 at The Linnean Society London, after the New Phytologist 'Networks in Plant Biology' meeting.
To register for this meeting see
http://www.newphytologist.org/networks/default.htm

2. 5th April at the SEB Canterbury Meeting
http://www.sebiology.org.uk/

References

1. von Dassow et al. (2000). Nature 406, 188-192.
2. Sachs et al. (2005). Science 308, 523-529.
3. Birnbaum, K. et al. (2003). Science 302, 1956-1960.
4. Hirai, M.Y. et al. (2004). Proc Natl Acad Sci U S A 101, 10205-10210.
5. Jonsson, H.et al. (2005). Bioinformatics 21 Suppl 1, i232-i240.
6. Locke, J.C.et al . (2005). Mol Sys Biol 1. doi: 10.1038/msb4100018.
7. Rolland-Lagan et al. (2003). Nature 422, 161-163.

The Future of UK Plant Systems Biology

Community Forum

GARNet Genomic Arabidopsis Resource Network

John Innes Centre

March 20th 2006

http://garnet.arabidopsis.org.uk

Background

Systems biology has a grand vision – understanding all the components of a biological system and their interactions, across all relevant levels of organisation.

A laudable aim but how do we achieve this goal?
To date the EPSRC and BBSRC have funded three Centres for Integrative Systems Biology (CISB) in Newcastle, Imperial College and Manchester and will fund another three from the second round of the initiative this year. These centres are intended to concentrate all the expertise and facilities required for systems biology in a single location, even a single building.
http://www.bbsrc.ac.uk/science/initiatives/cisb_phase2.html

But is this a successful way to approach to system biology in a widely distributed community such as plant research, which has over 350 laboratories spread across the UK? In addition to this approach would it possible for groups of plant laboratories in the UK to coordinate their work, in order to tackle large-scale projects? Or can we set up a distributed centre for plant systems biology?
These are the types of questions posed to GARNet by BBSRC's panel on Integrative Systems Biology (ISB) which has asked GARNet to produce a report on how systems biology can best be approached in UK Arabidopsis research. This report must incorporate the communities view and be with the BBSRC by June 2006.

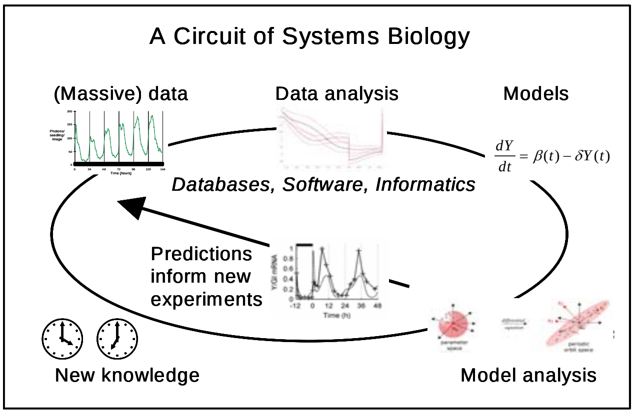
No specific objectives were set by the ISB panel for this report though the following were provided as potential areas for discussion:-

1. What might be the realistic research objectives of a UK Arabidopsis systems biology programme over the next 5/10 years?
2. What capabilities does the community currently have, what additional capabilities (skills, equipment, tools etc) need to be developed, and how should this be done?
3. What benefits might be expected from this?
4. What potential is there for international engagement?

So what is Systems Biology?

The ISB panel has avoided an exclusive definition of Systems Biology, preferring to list the characteristics of a systems approach and noting that the balance among these will vary among projects:

- An integrative approach to the subject
- Large experimental data sets (we would add, of high quality).
- Predictive capability based on modelling
- A mix of inputs not only from across biology but also from the chemical, physical, engineering, mathematical and/or computational sciences.



The result is an iterative interaction between experiments and modelling.

The story so far

A) Edinburgh Workshop

To kick start discussions in this area the BBSRC and GARNet held a workshop entitled 'Succeeding in Plant Systems Biology' in Edinburgh in July 2005. The workshop aimed to investigate systems biology approaches to plant science and consider the advantages and problems in applying such approaches to research. A full report of this meeting is available on line http://www.bbsrc.ac.uk/about/gov/panels/isb/docs/PSB_workshop_report_sept05.html

Delegates attending the workshop were asked to consider the current challenges and barriers to applying systems biology to research by addressing four main questions. A summary of these discussions are provided in Appendix A.

B) Interim Report

During the autumn of 2005 the GARNet committee deliberated over the issues raised by the Edinburgh workshop and carried out consultations with international researchers to generate an interim report on UK plant systems biology (submitted to the ISB panel in January). A full copy of the report is provided in Appendix B

The GARNet committee discussions for this report were wide ranging, covering a variety of topics including;

Can systems biology be carried out in a multicellular organism?

Is a collegiate effort possible in the context of the diverse research objectives and priorities of individual scientists and their institutions?

Should a community driven project focus on a single biological research goal or just provide the necessary new theory, data analysis and modeling tools?

What type of training is required to promote closer collaboration of biologists with theoreticians and modellers?

How will data quality, capture, storage and distribution be assured in a multi site project?

How do we attract the appropriate theoreticians to work in plant science?

Potential solutions to some of these questions are provided in the report whilst others required further community input.

C) Town Hall Meetings

To canvass community opinion on the interim report and the questions it raises GARNet has and will be holding town hall meetings during 2006.
The first of these meetings was in conjunction with the New Phytologist conference 'Networks in Plant Biology' 26th-27th January in London.
http://www.newphytologist.org/networks/

Those attending the meeting were asked to read the interim report and consider the following

1. What concrete steps would most facilitate systems biology approaches in the UK Arabidopsis community? Examples include: a national 4-year PhD programme with central facilities; 'study group' events that allow plant scientists to present their research questions to theoreticians for preliminary analysis; a set of demonstrator projects that show what can be accomplished by the systems approach in plants; a networked research activity focussing on a longer-term goal; tools and resources for the acquisition of large-scale data sets.

2. Agreeing a single research goal might have significant advantages. What would be a suitable goal for a coordinated, national project? Examples might include: systems biology of the leaf, of the guard cell, of the auxin signalling pathway. For comparison, the systems biology centre in Cambridge proposes to have 100 researchers focussing on the *Notch* and *wnt* signalling pathways in Drosophila.

3. One aim of a large-scale systems biology project could be to collect a particular data set or to provide a particular resource in support of a national project, in addition to gathering researchers with the relevant expertise. Which data sets or resources would convince you to work in a new area, or which could you contribute to providing? Examples might include: a protein-protein interaction map, completion of the metabolic map, imaging of protein localisation throughout development, integrating biological databases, informatics that automates the link between data and models.

There was a lively debate over the report and those attending the meeting agreed that:-

1. Training was an essential part of a community led effort in systems biology. However, it is not part of BBSRC remit to allocate initiative funds for studentships (a policy that the audience were critical of).

2. A single goal which all the community could sign up to was not agreed upon. Instead there were suggestions that all proposals in this area should be enabling and used as a foundation for other projects to build on in the future from responsive mode funding.

3. Delegates requested a generic level of systems biology support in a similar manner to the support and advice GARNet previously provided for genomic technologies. Examples provided by the floor to achieve this included; an annual retreat for life and theoretical scientists, a summer school for students and a centre to which labs could send personnel for training.

The next town hall meeting open to the whole community will take place during the SEB Canterbury Meeting on the 5th April 2006.

Norwich is a site for cutting-edge plant science research, and since two Garnet committee members are located here (Jonathan Jones and Nick Harberd), we have decided to implement a "town meeting" at JIC to enable questions and answers with, and feedback from, JIC, UEA and IFR scientists.

Appendix A
Edinburgh Workshop - Discussion Summary

Q1 What advantages might systems biology offer to plant sciences?

Formal representation as a 'system' allows:

- Better comparison between organisms
- Integration of disciplines
- Scalability from pathways and subsystems to landscapes
- Common structures for data collection, analysis and design
- Mathematical abstraction with predictive power
- Development of new tools

Q2 What areas of the plant sciences might be particularly amenable to such approaches?

All areas were seen as amenable to a systems approach from single cells to ecology, and even the landscape examples included; environmental responses, the leaf as a system, plant pathogen interactions

Q3 What are the current technical and other barriers in applying systems biology approaches to plant research?

- Lack of multidisciplinary training (MSc, PhD)
- Poor communication with other disciplines, need to develop common languages.
- Lack of "best practice" for data handling
- Spatial heterogeneity (many cell types contribute to each data point)
- Lack of examples of good predictive models that have been demonstrated to work.
- Project Timescales are too short to develop interdisciplinary research.

Q4 What scientific expertise can the UK in particular offer in this context?

- Broad excellence in plant science, computational biology and informatics
- (not always well connected)
- Numerous resources
  - Germplasm collections in a range of species
  - Large data sets e.g. Transcriptomic data sets (NASC)
  - Genomic databases
  - Metabolomics at Rothamsted and Aberystwyth
  - Large *in silico* crop breeding databases,
- Wide range of species studied

Interim report of the GARNet Advisory Committee on Arabidopsis Systems Biology in the UK

1. Key recommendations

Biological research using *Arabidopsis thaliana* can profitably adopt systems approaches and is doing so in some areas. Arabidopsis researchers must build the capacity for systems analysis, if the UK's strength in Arabidopsis biology is to be reflected in systems models that, ultimately, will lead to a whole-plant model. Challenges include the acquisition of data from single cells or cell types (see 3B) and increasing theoretical research on Arabidopsis biology (see 3C). The committee's interim recommendations are to establish:

- A distributed, 4-year PhD training school in Arabidopsis systems biology, with central facilities (see 4A)
- Networking activities for Arabidopsis researchers with physical, computer and mathematical sciences (see 4B)
- Demonstration systems research projects on tractable subsystems (see 4C)
- Additional of tools and resources, contributing to the acquisition of large-scale data sets (see 4D, 4E)
- A networked research activity focussing on a longer-term goal, comprising 20-25 research groups (see 5)

The GARNet committee's consultation will continue in 2006, particularly on the longer-term goals in Arabidopsis systems biology (5A), on mechanisms to address these goals (5B) and on involvement with international institutions (5D).

2. Goals

Systems biology in plants, as in other organisms, aims to understand all the components of a biological process, together with their interactions and emergent behaviour, at all relevant levels of biological organisation. This approach emphasizes an integrated understanding of biological systems. Modelling is central as a philosophy and as a tool to support data integration, simulation, analysis, prediction, knowledge delivery and the design of new experiments. One aim is to test the mutual compatibility of ideas derived from reductionist studies of individual components or small subsystems. Ultimately, systems biology will result in the development of a whole-plant model, a 'virtual Arabidopsis' or 'computable plant'. This will aim to account quantitatively for all stages of the Arabidopsis life cycle from seed to seed, together with variations in growth and cellular function due to environmental responses. One proposal sets a target date of 2017 for this goal (www.cmp.uea.ac.uk/ivis).

The whole-plant model will probably integrate multiple, subsystem models that are appropriate to address different biological questions. The model will describe macroscopic functions from the level of intracellular processes at the scale of macromolecular complexes and metabolite pools. It will be based upon specific, observed biochemical and biophysical mechanisms and validated by experimental data at multiple levels. For some topics, the Arabidopsis model will be extended to larger scales, incorporating interactions in plant populations, or to smaller scales, extending to the atomic description of particular protein surfaces, for example. Crop modelling on the field scale, where a plant population develops under well-characterised conditions, may allow plant systems models to extend above the level of the individual more easily. Crossing scales from the macromolecular to the individual organism will, however, be a central activity of plant systems biology for the next decade or more.

3. Benefits and challenges

3A. Benefits

The UK Arabidopsis community is generally enthusiastic about the prospects for systems research and is realising the opportunities for systems insight across many areas of Arabidopsis biology. The benefits of generating predictive models are obvious to the community. A virtual Arabidopsis will be a significant and highly visible achievement in biology, with spinoff benefits across multiple fields of scientific research, in addition to direct applications in areas ranging from pharmacobotany to global climate change. Basing this activity on a plant, rather than another organism, has major advantages that arise from fundamental plant biology as well as from the current status and organisation of the Arabidopsis research community. Spatio-temporal models of growth and development will be significantly easier to develop at the cellular level for plants compared to animal species, because cell migration and changes in cell shape are very limited in plants. Whole-cell models of microbial species, which are currently in development, should greatly facilitate the modelling of intracellular processes in plant cells. The technical and logistical issues facing plant systems biology are also well recognised by the community. In sections 3B and 3C, we highlight specific issues and propose solutions in sections 4 and 5, below.

3B. Technical challenges

Experimental screening methods that are readily applicable to single cell systems will often require further technical development to return data on single cell types within a complex multicellular organism, such as Arabidopsis. A workshop at the 2005 GARNet meeting specifically considered this issue (see report by Prof. Jones, at garnet.arabidopsis.info/garnet_meeting.htm). Data from defined cell types will be important for many but not all projects. Data on whole seedlings remains relevant for subsystems that are similar in all cells or conversely for processes that are active in only one cell type, and more generally in generating preliminary models that average across heterogeneous cell types.

Cell purification and cell cultures. Certain plant cells can be physically separated at high purity and in high numbers, including pollen and stomatal guard cells, permitting a wide range of experimental methods. Cell culture systems can produce data amenable to modelling, for example time-course data following application of a low-molecular-weight compound. They have been used in transcriptional and proteomic profiling of cell division and senscence pathways, for example, and could be used to provide baseline data, such as a protein-protein interaction network, that contribute to network inference. However, plant cell suspension cultures are not analogous to animal cell cultures, because they do not obviously represent a cell type that is present in the plant: they typically fail to express secondary metabolic pathways, may be unable to differentiate, and are genetically and physiologically heterogeneous. Many biological questions cannot be addressed in plant cell cultures, so a major focus remains on the intact plant. Purified cell types may be an option for some of these studies.

Extraction from defined cell types. Techniques exist to allow RNA and, potentially, protein and metabolite profiling of single cell types, if not single cells. These include laser-capture microdissection, fluorescently tagging and purifying single cell-type protoplasts, the use of micropipettes to sample the cytoplasmic contents of single cells, or the extraction of mRNA from polyribosomes that carry a cell-type-specific protein tag. These have been used in plants for the analysis of high-resolution transcriptional changes, exploiting linear amplification systems for RNA populations. While such amplification techniques are not applicable for protein and metabolite analysis, advances in dye technology and sensitive mass-spectrometry are allowing higher resolution identification of proteins and metabolites. There is clearly the potential to integrate transcriptome, proteome and metabolome data from defined plant cell types, particularly if there is continued technical development in this area.

3C. Logistical challenges

Research in plant systems biology is currently carried out in relatively few biology laboratories in the UK, in diverse geographical locations. Systems for collaborative interaction and discussion will be required to share experiences and solutions among groups, in addition to systems for sharing diverse data. This will be essential to develop systems-level research in the UK Arabidopsis community.

Related to this is the requirement for mathematical and computational expertise in an accessible form for plant biologists. Not every institution will be able to provide the necessary theoretical/modelling expertise or indeed enthusiasm for problems in plant biology. At the project level, individual biologists must identify collaborators, who may be in other institutions. This is not necessarily a barrier to progress, as many already have collaborators at distant locations, but ready access is clearly advantageous. At the community level, a significant increase in the number of theoreticians working on questions in plant science will be required to achieve any large-scale goals in Arabidopsis systems biology. The BBSRC workshop in July 2005 identified this as a major issue for the field (see http://149.155.200.17/about/gov/panels/isb_docs/PSB_workshop_report_sept05.html).

4. Near-term objectives

Projects that can and should be tackled now will contribute directly to the overarching goal of a whole-plant model: here we outline the near-term objectives that have particular importance. Training and networking activity (4A and 4B) must be as widely spread as possible and should result in the adoption of some systems methods in many Arabidopsis research projects. Methodological research (4D) clearly requires the participation of multiple disciplines; most research coordination required in this area will not be specific to plant research. The benefits from systems research projects will be significantly enhanced if they are coordinated: in section 5, we address longer-term goals, including a large-scale network for Arabidopsis systems biology. The requirement for theoretical expertise in plant systems biology must be addressed by multiple mechanisms to attract sufficient numbers of researchers into interdisciplinary Arabidopsis projects, both in the short and longer term.

4A. Interdisciplinary teaching, training and re-training.

The GARNet Advisory committee recommends:

i. A national, 4-year PhD training programme in Plant Systems Biology, along the lines of an EPSRC/BBSRC Doctoral Training Centre, to provide trained researchers. Students would be based across a distributed network of Arabidopsis biology laboratories, linking major centres of Arabidopsis research and several locations with less activity. Central investment in computing support and theoretical training, together with networking activities, would address the need for communication among Arabidopsis systems biology groups (see 3C) and provide added value.

ii. Increased investment in discipline-hopping and sabbatical visiting awards (including international awards) should continue, to address the need for training and retraining at the postdoctoral and senior levels. These awards should be on flexible terms to allow for the wide range of disciplines required in systems biology.

4B. Networking with the physical, computer and mathematical sciences

GARNet is pursuing two actions in this key area, to promote the traditional, distributed interaction among individual biology laboratories and their collaborators.

i. Researchers in plant bioinformatics and crop modelling provide a pool of research expertise that is relevant to plant systems biology. There is relatively little networking in this community, so the national potential for input to systems biology is hard to judge. Prof. Howard Thomas and Dr. Helen Ougham (IGER) have agreed to organise a workshop to bring together researchers in this area before Easter 2006, to enhance awareness of current research and discuss future involvement in systems biology.

ii. 'Study Groups' provide a week-long opportunity for mathematicians to tackle specific problems in an area of science, to give an informed view of the opportunities and approaches for further research. GARNet is working with Marcus Tindall (Oxford) to develop a proposal for a mathematics study group in Arabidopsis biology, along the lines of the Mathematics in Medicine Study Groups.

In the longer term, any coordinated funding for Arabidopsis systems biology should invest in central bioinformatics, programming and/or modelling support as a core activity (see 5B).

4C. Systems biology approaches to tractable subsystems in Arabidopsis

Systems-oriented projects that can be tackled immediately will achieve several, related objectives in the context of plant systems biology, in addition to the immediate objectives of each project:

- to complete case studies that plant biologists can turn to as exemplars of the systems approach, demonstrating concrete benefits to plant science from systems biology
- to promote interdisciplinary collaborations (see 4B above)
- to accelerate model building and validation in Arabidopsis biology, which is currently minimal but will contribute directly to the whole-plant model
- to increase the scope of Arabidopsis research objectives progressively from a focus on a particular gene or pathway to a sub-system and ultimately full-system approach
- to increase the volume and diversity of data that can be routinely analysed, interpreted and modelled in Arabidopsis research groups

It is therefore doubly important that early efforts in this area are appropriately reviewed and that the evaluation committees include members with suitable expertise. Completing a small number of highly interdisciplinary research projects may be more beneficial for the take-up of systems research than a broader spread of more conventional activity, particularly given the increased funding and longer timescales required for systems research even on a sub-system. Awards greater than £1M should be anticipated.

There is scope for such projects across many areas of plant science and at this early stage it is undesirable to be selective based on subject area. It is desirable that the projects make best use of the current engagement of theoreticians and informaticians. This could be achieved by extending current modelling/bioinformatics projects into related biological areas, widening the group of Arabidopsis biologists that collaborate with each modelling/bioinformatics group. Coordination by GARNet might facilitate such a process.

4D. Technological development of Systems-level tools, both experimental and theoretical

The tools required will be common to other model species and to conventional plant research. Particular goals might include methods to:

- obtain cell specific, or cell type-specific, biochemical data that are suitable for systems modelling (high quality, high time resolution, genome-scale)

- monitor multiple plant and plant cell functions *in vivo*
- allow dynamic manipulation of multiple biological components *in vivo*
- maintain the differentiation status of Arabidopsis cell cultures
- model Arabidopsis growth patterns at the cellular level (already underway)
- visualise experimental data in the context of plant architecture
- capture data from the experimental literature
- facilitate public access to quality-controlled experimental data of multiple types
- integrate multiple data types, both in individual research projects and in public databases
- infer network models from 'omics data
- convert static to dynamic models
- optimise experimental testing of models, exploiting the available tools and resources

Arabidopsis biologists must engage strongly with infrastructural and methodological development in other research communities, to make most rapid progress and to avoid duplication of resource development and provision. Structures that allow international and European coordination in Arabidopsis research should be harnessed for this purpose. Greater integration with resources developed in biomedical research will likely be beneficial, not limited to sequencing facilities.

4E. Baseline data sets underpinning Plant Systems biology

This activity aims to provide the global data sets, the biological resources to produce the data and the network models from analysis of the data. Genome-wide analysis of DNA- and RNA-binding protein binding sites, proteome-wide protein interaction analysis, and development of plant-specific ontologies are already underway. Other data sets might include:

- Proteome-wide interaction analysis of key ligands
- Completion of a core metabolic map
- Description of all Arabidopsis cell types and their numbers in different organs
- Image timeseries of Arabidopsis development

The provision of such data could be coordinated internationally; the UK will be in a strong position to contribute to some elements. These data sets, together with data analysis tools and results (e.g. network maps) in public access knowledge bases, are the starting resources for systems biology research. Their importance for systems biology is analogous to the importance of populations of sequence-mapped insertion lines for Arabidopsis functional genomics. The Arabidopsis community, with BBSRC, has a good track record of balancing resource provision with hypothesis-led research. The GARNet committee recommends that the Integrative and Systems Biology Panel and the Tools and Resources Panel jointly establish ways to fund the new resource provision for Systems Biology.

5. Goals and organisation of Arabidopsis systems biology in the UK

5A. Longer-term goals

Setting a number of ambitious goals of lesser scope than the whole-plant model will be important if the international plant research community is to achieve a whole-plant model within a reasonable timeframe. The goal-oriented philosophy developed in such projects is some distance from current research practice in most parts of the community, but will produce concrete benefits to justify the necessary changes. For example, there will be growing benefits from the coordination and standardisation of experiment, data and model formats, maximizing the inter-lab comparability and thus the value of data and models from different groups. The visibility of larger-scale projects should facilitate interactions with other research communities (e.g. increase engagement by theoreticians), funding agencies, policy-makers and the public.

The UK Arabidopsis community could realistically undertake one or possibly two large-scale projects, each of which would be much larger than the subsystem projects (see 4C), involving many groups together with their interdisciplinary collaborators. There is widespread acceptance that a set of guidelines should be provided to select projects in plant systems biology that are consistent with an agreed goal. The nature, scope and particular benefits of such goals are the subject for ongoing consultation by GARNet. As described above, the characteristics of Arabidopsis development provides a major advantage over other multicellular models for spatial modelling, so a goal that exploits this advantage might be attractive. In other areas, Arabidopsis systems biology may be at least as tractable as other multicellular model species, for example in understanding environmental response networks. Further areas may be less tractable but have strong relevance to global geochemical processes and to end users, for example photosynthetic metabolism and yield measures.

The possibilities under discussion include different combinations of these factors, ranging from organ-scale goals, e.g. "The Leaf" (see box), through projects focussing on a single, experimentally tractable cell type, e.g. "The Guard Cell", to projects focussing on a single signalling pathway or subsystem, e.g. "The Photoperiod Response System". The balance between inclusiveness and research focus will be crucial. A strongly-focussed activity has great potential for synergy among groups but must also offer attractive research questions together with powerful resources and baseline data, in order for sufficient Arabidopsis researchers to align their current activity with the broader goal.

> A sample project: Systems Biology of the Leaf
>
> One way to encourage the community to work more coherently on a broad problem is to focus everyone's attention on a physical entity, such as the leaf.
>
> Many features of leaves commend themselves to the systems approach. The cell number of the Arabidopsis leaf has been defined (Pyke & Leech [1991] J Exp Bot 42:1407), and the proportion of mesophyll, epidermal and vascular tissue measured. The growth of leaves from primordia on the apical meristem is already a subject of mathematical modelling. Leaves also perform a wide range of important metabolic and signalling processes, studied by many Arabidopsis research groups in the UK. Leaves first import then export sugars over their life times. Leaves receive and export hormone signals. When challenged by pathogens and pests, attacked leaves export and systemic leaves receive signals such as oxylipins. Stomatal guard cells regulate the flow of CO2 into and water out of the leaf, under the control of various environmental signals. Differing light intensities will also affect aspects of leaf performance, and differentially at different times of day.
>
> Other levels of biological organization could be chosen; but the choice of an organ such as the leaf would force researchers to consider how the intracellular pathways they study play out in the context of whole tissues and organs, and whether different cell types are behaving in the same way. This choice would force integration in studies of physiology and development, and of intra- and inter- cellular signalling. A long-term UK project in this area would be a major contribution to any international effort to develop a whole-plant model.

5B. Networked research activity to tackle a longer-term goal

How would research teams be organised, in order to address one of the longer-term goals outlined in section 5A? Possible structures for a networked Arabidopsis systems biology were considered by the GARNet advisory committee. The logistical issues in coordinating research across groups even within the UK will be significant. However, experience of and mechanisms for national coordination will be a major advantage when international coordination is required on the whole-plant model. A single overarching network linking the current Arabidopsis biology community, with inputs from chemical, physical, engineering, mathematical and computational sciences, would be too diffuse (would lack focus) and would be too cumbersome to manage effectively. A more likely structure would comprise a network of collaborations with a shared core component. The majority of activity would be in (say) 4-5 interdisciplinary collaborations, each collaboration consisting of (say) 4-5 partners, with a mix of 'traditional' Arabidopsis biology labs and theoreticians, focussed on a particular biological problem. The problems selected should be sufficiently close to allow integration of the results and models to achieve a larger goal (see 5A) and would probably include one or more of the subsystems project areas (see 4C). The core component would provide project management and infrastructure that are shared by all the sub-groups, with infrastructure to support modelling and bioinformatics rather than experimental facilities. Interactions with the core and across collaborations, together with regular meetings of the entire network, would facilitate the exchange of modelling tools and experiences. The minimum level of funding required to make this an internationally competitive activity with minimal activity in the core is ~£5M over 5 years: this assumes that some of the subsystems funding has been allocated in the same research area. Increased core activity or generation of a large data set would increase this figure. The 5-year duration is also a minimum.

5C. Contribution of academic groups, NASC and the CISBs

A networked research activity would comprise a wide variety of groups in universities and BBSRC-funded institutes, with some participation of other institutes, such as those funded by the Scottish Executive. It should make best use of the bioinformatics resources and materials at the European Arabidopsis Stock Centre (NASC), and those to be provided by the BBSRC Centres for Integrative and Systems Biology (CISBs). Outreach of the various CISB research programmes to plant science will be encouraged by greater activity in Arabidopsis systems biology, particularly if a CISB is funded in phase 2 to focus on research in an area of plant science. A plant-focussed CISB would complement and not preclude the successful operation of a networked activity in Arabidopsis systems biology.

5D. International context

The extent to which interdisciplinary Arabidopsis research is termed systems biology varies with the working definition of systems biology in force locally. There is generally greatest enthusiasm where the research areas that are of most interest locally are also tractable for systems approaches. The GARNet committee will consult additional international sources in the months ahead. Locations in continental Europe that are strongly engaged in plant systems biology will be natural partners for a large-scale activity in the UK. These include the Max Plank Institute of Molecular Plant Physiology in Golm, Germany, and the department of Plant Systems Biology in Ghent, Belgium. At least one integrated proposal (coordinated from Ghent) has recently been submitted to the EU under Framework Programme 6 to support large-scale, coordinated acquisition of systems data in Arabidopsis, with data analysis and limited modelling.

Activity in the US is significant, particularly through NSF funding. There is as yet no national coordination specific to plant systems biology, nor is there a consensus on the need for such coordination. The NSF 2010 programme on Arabidopsis functional genomics recently produced its mid-course report, recommending a focus on the provision of biological data and informatics resources (compare section 4E, above) and on understanding "exemplary networks" to facilitate the transition to systems biology. Several NSF FIBR grants of US$2-3M have been awarded to interdisciplinary collaborations focussing on particular questions in plant systems biology (compare section 4C above), including projects on meristem architecture (led by Mjolsness, UC-Irvine and Meyerowitz, Caltech), on nitrogen and carbon signalling (Coruzzi, NYU) and on flowering time (Schmitt, Brown). A recent NSF workshop considered the establishment of an interdisciplinary 'Synthesis Centre' or 'Cyber-Infrastructure Centre' in plant science, to provide core support and technical development in informatics and modelling (compare 5B, above).

6. Potential risks and alternatives

Implications of the adoption of systems approaches in plant research include:

- Selection of model species for systems research. At this early stage of plant systems biology, focussing on more than one model species risks losing the interaction between research areas that will be crucial in order to combine sub-systems models within a reasonable timeframe. Achieving this synergy even within one model species requires collaboration across research areas that may never have previously occurred. The only model species that can realistically be supported is therefore Arabidopsis. Arabidopsis systems biology projects should clearly make their data, tools and resources available for systems-oriented research in other species (plant and non-plant).
- Conflict of research priorities between translational and fundamental research in Arabidopsis labs. The 'models to crops' agenda will have significant effects on the UK Arabidopsis community at the same time as the systems biology effort develops. Smaller academic labs may be unable to sustain activity at both interfaces. For larger groups and the community more generally, there is a risk of diluting research effort in both areas.

7. Conclusion

Multiple initiatives will be necessary to develop Arabidopsis systems biology in the UK on the scale required to tackle the challenge of a whole-plant systems model. A distributed research activity, including a fraction of the 200+ Arabidopsis research groups in the UK together with their interdisciplinary collaborators, could establish a leading position in plant systems biology if it were suitably coordinated.

Acknowledgements: The committee is grateful to BBSRC for funding the Plant Systems Biology workshop held in Edinburgh, in July 2005, to many participants at the workshop, and to individuals who have contributed their time to this interim report, including: Steve Russell, Elliot Meyerowitz, Chris Somerville, Pierre Hilson, Mark Stitt, Stefan Hohmann, Igor Goryanin.

GARNet Advisory Committee (for interim report): Ian Furner (Chair), Philip Gilmartin, Julie Gray, Claire Grierson, Nick Harberd, Jonathan Jones, Marc Knight, Ottoline Leyser, Keith Lindsey, Andrew Millar (Co-ordinator), Simon Turner, Sophie Laurie (BBSRC contact). We are extremely grateful to the GARNet administrator, Dr. Ruth Bastow.

GARNet

Genomic Arabidopsis Resource Network