

The Environmental Genomics programme

Science and implementation plan

Contents

- Introduction
- Scientific guidance
- Programme deliverables
- Programme funding and implementation
- Training and community development
- National and international collaborations
- Data management plan
- Ownership and exploitation of intellectual property (IP)
- programme management

Introduction

The Environmental Genomics research programme aims to facilitate the exploitation and, where appropriate, acquisition of whole or partial genome sequence data in environmental biology. Central to this aim is the establishment of a community of researchers, comprising empiricists and theoreticians, who are proficient in the development and application of post-genome technologies to the study of organism and ecosystem function. The use of model systems, where substantial genome sequence already exists, will be considered where an ecological relevance can be demonstrated. The study of specific genes within a single species falls outside the remit of the Environmental Genomics programme.

The integration of molecular genome science with ecology and evolution will generate exciting and novel interdisciplinary research opportunities. The precise nature of these opportunities is impossible to anticipate without restricting the scope of applications. An extensive list of prescriptive objectives has therefore been avoided in the interest of encouraging a broad spectrum of imaginative and innovative proposals. Nevertheless, there are key areas where the marriage of genomics and ecology will ensure substantial progress.

Scientific guidance

1. Phenotypic and genotypic bases of fitness

Whole genome sequence, in conjunction with technologies for functional analysis, enables ecological performance (fitness) to be studied using "bottom-up" (genes to population) strategies. Objectives may include:

- Identification of gene loci affecting ecological performance.
- Functional characterisation of traits affecting ecological performance.
- Individual, population, community and ecosystem responses to the environment.
- The extent and significance of allelic variation among traits affecting ecological performance.

2. Ecological significance of molecular variation

Chip and other technologies enable variation among populations to be quantified rapidly at multiple loci. Objectives may include:

- Determining the consequence of variation in the levels of gene expression at multiple loci on ecological performance.
- Defining the extent of spatial and temporal variation at regulatory versus structural loci.
- Identifying numbers of loci (QTLs) and their genomic distribution.

- Determining community structure in unculturable organisms and investigating the response to environmental change.

3. Integration of genotype and phenotype

One of the major challenges facing biologists is the understanding of how phenotype maps on to genotype. Many opportunities exist for exploiting whole or partial genome data to address aspects of this problem. Objectives may include:

- Responses of genomes to genetic and environmental stimuli.
- The impact the environment plays on the phenotype.

4. Phenotypic evolution

The relationship between DNA sequence, form and function is poorly understood. Whole or partial genome sequence, in combination with the tools of functional genomics, offers opportunities to develop a better understanding of phenotypic evolution. Objectives may include:

- Explaining (at mechanistic levels) the evolution of novel traits.
- Determining rules of adaptive evolution at a theoretical level.
- Understanding the contribution of regulatory pathways to adaptive evolution.

5. Genome architecture

The position of genes within genomes, their number, arrangement and the nature of their interactions determine the expression of phenotypic traits. Objectives may include:

- The stability of genetic architecture.
- Organisational constraints.
- Responses of genome architecture to selection.

6. Collation, management and mining of data

Data generated from genome sequencing projects, but particularly from the application of functional genomic strategies, are extensive and require careful collation. Software packages for management and analysis of genome sequence are available, but few exist for the collation and analysis of functional genomic data. Objectives may include:

- Development of "smart" data collection software for the collation and analysis of large data sets. Such tools must enable the cross-comparison of data generated under different experimental regimes.
- Development of a theoretical and computational framework for genome sequence data that incorporates and integrates the transcriptome, proteome and metabolome.

7. Technological developments

A wide range of tools has been developed specifically to address the question of gene activity. However, few of these strategies can currently be used with success in complex natural environments, such as soil. An important objective will be the modification of existing, and the development of new, technologies capable of measuring gene expression in complex environments.

Programme deliverables

The performance targets of the programme are:

- to attain the scientific objectives stated above, through an integrated programme based on interdisciplinary collaborations, an international perspective and cross-research council linkages.
- to develop a data management system for the archiving, collation and mining of environmental genomic data. Cross-research council linkages with EPSRC and BBSRC Bioinformatic initiatives will be established.

- to provide high quality training to develop national expertise in the application of genomic-based technologies to environmental samples and the management of such data.
- to demonstrate to the wider environmental community, in particular environmental regulators and the industrial user community, the potential application of post-genomic technologies.

Programme funding and implementation

The Environmental Genomics research programme is funded by the Natural Environment Research Council (NERC) at £16.5m over c. 5 years. Approximately £2.5m of these funds will be set aside to ensure the development and implementation of the data management plan.

There have been two calls for proposals: under Round 1, 17 applications (graded a-4 & above) to the value of c £8M were funded ; under Round 2, 10 applications (graded a-3 high & above) to the value of c £3.5M were funded. A third round is planned which will be committed to engaging interested stake-holders, including industry and environmental policy makers. NERC have set aside c 500K to leverage additional funds to address areas of mutual interest

It is anticipated that the second phase of projects will be funded to coincide with the start of the 2002/2003 academic year.

Two Environmental Genomics Workshops have been held thus far, which highlight progress in the application of post-genomic technology to answer ecologically relevant questions, and to continue cross-disciplinary communication. Specific workshops will also be held to present the achievements of the Environmental Genomics programme to the user community.

An Environmental Genomics Workshop was held on 19th & 20th December 2000 at Scarman House, Warwick University. This enabled the aims and objectives of the NERC Environmental Genomics programme to be presented to the scientific community. The workshop also helped to foster collaborative links within the scientific community.

Subsequent workshops will be held to highlight progress in the application of post-genomic technology to answer ecologically relevant questions, and to continue cross-disciplinary communication. Specific workshops will also be held to present the achievements of the Environmental Genomics programme to the user community.

Training and community development

The Environmental Genomics initiative will establish a new scientific research community. To ensure that this develops rapidly with appropriate collaborative links, regular workshops will be held for established researchers, post-doctoral scientists, post-graduate students and the user community.

Numerous training courses exist, nationally and internationally. Attendance of established researchers, post-doctoral scientists and post-graduate students at these courses will be particularly encouraged, to aid their professional development, and the establishment of the Environmental Genomics research community,. Approximately £200K will be set aside to support such training.

It is anticipated that any centres set up under the Environmental Genomics programme to provide technical support to award holders will provide appropriate training. Regular project updates and other pertinent information will be posted on the NERC Environmental Genomics web pages to ensure effective transfer of knowledge to the wider community.

National and international collaborations

Joint proposals between research groups with complementary expertise, and with involvement of industry, are strongly encouraged within the programme, to stimulate technology transfer and linkages between fundamental and applied science. Given the global effort in genome sequencing projects and the application of post-genomic technology, international linkages are also encouraged.

Data management plan

It is NERC policy that research programmes ensure the long-term availability of data collected by award holders to maximise the application and exploitation of their results. All data and reports submitted under the programme must be made available in an electronic format. This will maximise ease of future retrieval and exploitation.

In most cases, NERC Designated Data Centres are used, both for quality control and archiving, with costs covered from central programme funds. After a period of sole access by PIs for publication preparation, such data are made available to other programme participants and the wider community.

Under the Environmental Genomics programme there is the potential for vast amounts of data to be generated. The collation, archiving, storage and mining of these data will require data management systems that are currently not present within NERC. A provisional sum of £2m has been set aside for proposals to manage the data generated, through the development of appropriate data management systems.

A Data Sub-Committee has been established with the approval of the Environmental Genomics Steering Committee to draw up a detailed Data Management Plan. There will be communication with other research programmes including the Soil Biodiversity and Marine & Freshwater Microbial Biodiversity (M&FMB) programmes. Other cross-research-council Data Management systems will be examined.

Ownership and exploitation of intellectual property (IP)

Ownership of IP and IP Rights arising from NERC awards lies with the grant holding body (university, NERC Research Centre or other NERC-approved institution). Whilst all recipients of NERC funding must endeavour to ensure that the outcomes of their research are used to the advantage of the UK, NERC retains the right, for a limited period, to exploit IP in partnership with grant recipient organisations to the benefit of the UK and the organisations. This is to avoid circumstances where fragmentation of IP would reduce the likelihood of exploitation.

Many possibilities exist for the commercial exploitation of discoveries. Workshops, involving end-users, such as industry, environmental regulators, NGOs, conservation scientists and IP companies, will be held to discuss mutually beneficial means of maximising benefit from novel discoveries.

Programme management

The **Environmental Genomics steering committee** is responsible for directing and integrating the programme science and for recommending the award of research funds. The Committee also has responsibilities for co-ordinating data management arrangements, and for promoting the transfer of the outcomes of the research to the user community. The steering committee is chaired by **Professor Chris Payne OBE** (School of Plant Sciences, The University of Reading, Whiteknights, Reading, Berkshire, RG6 6AS; e-mail: C.C.Payne@reading.ac.uk; tel: 0118 931 8071; fax: 0118 975 0630.

The **Environmental Genomics Programme Administrator** has responsibility for all aspects of the day to day programme administration and for enquiries concerning the programme. Contact information: **Will Thomas**, NERC Polaris House, North Star Avenue, Swindon SN2 1EU; e-mail: wpht@nerc.ac.uk; tel: 01793 411698; fax: 01793 411545.

The **Environmental Genomics Science Co-ordinator** assists the steering committee in the co-ordinated implementation of the programme's science objectives, through liaison with award holders and other bodies. Contact information for Science Coordinator: **Dr Jason Snape**, Brixham Environmental Laboratory, AstraZeneca, Freshwater Quarry, Brixham, South Devon, TQ5 8BA; e-mail: jason.snape@brixham.astrazeneca.com; tel: 01803 882882; fax: 01803 882974.