

Contents

Speakers Abstracts

| | |
|--|---|
| David Ferrier..... | 2 |
| Homeobox Genes as a Route to Reconstructing Animal Ancestors | |
| Esko Ukkonen | 2 |
| Computational Prediction of Gene Enhancer Elements by Comparative Genomics | |
| Francesca Ciccarelli | 3 |
| Low Duplicability and Network Fragility of Cancer Genes | |
| Aoife McLysaght | 3 |
| Interacting Gene Clusters and the Evolution of the Vertebrate Immune System | |
| Anastasia Gardiner | 4 |
| Chemoperception in the Fruit Fly – The Genome’s Story | |
| Ian Roberts | 4 |
| pSNPs: A New Tool for Comparative Genomics? | |
| Frank Wright | 5 |
| TOPALi Software for Evolutionary Analysis of Multiple Alignment Data | |
| Trevor Paterson | 5 |
| The ComparaGRID Project: Integrating Genomic Mapping Data using Semantic Web Technologies (Can Formal Ontologies with Logic Reasoners Help Biologists?) | |
| Madhu Bhattacharjee | 6 |
| A Bayesian Framework for Data Integration: Application to Comparative Genomics | |
| Joaquín Dopazo | 7 |
| Systems biology in functional genomics: a bioinformatics perspective | |

Poster Abstracts

| | |
|--|-----------|
| Leighton Pritchard | |
| Comparative and Functional Genomics Identifies Major Differences Between Genomic Islands in Soft Rotting Enterobacterial Plant Pathogens | 8 |
| Saif Ur-Rehman | |
| A eukaryotic phylogeny as an initial step to a functional annotation pipeline of protein coding gene products | 8 |
| Olga Vasieva | |
| The RAST Server: Rapid Annotations using Subsystems Technology | 9 |
| Maria Keays | |
| Preliminary survey of desaturase genes in 12 Drosophila species | 9 |
| Linda Cardle | |
| Relator: a visualization tool for relating unigene datasets | 10 |
| Urmi Trivedi | |
| A vision of Comparative Genomics through next generation sequencing | 10 |
| Alexey Goltsov | |
| Modelling of compensatory properties of large scale gene/metabolic networks carrying series of mutations. Application to purine biosynthesis of <i>E.coli</i> strains | 11 |
| Meeting Information | 12 |

Homeobox Genes as a Route to Reconstructing Animal Ancestors

David E. K. Ferrier

University of St Andrews, The Gatty Marine Laboratory, School of Biology, St Andrews, Fife, KY16 8LB, UK

The discovery of the homeobox in the 1980s was a major factor in the rejuvenation of the field of evolutionary developmental biology. Genes with this motif have key roles in many aspects of animal development and evolution. Allied to these prominent roles in animal evo-devo homeobox genes, particularly of the ANTP class, often exhibit a distinctive organization in the genome in gene clusters. The best known example of this is the Hox gene cluster. As well as providing insight into how genomic organization impacts on developmental gene function, homeobox genes also provide excellent markers for revealing major transitions in animal genome evolution, such as the 2 rounds of whole genome duplication early in vertebrate evolution. Here I will draw on some recent examples of comparisons of homeobox gene organisation from across the animals to illustrate the utility of these genes in understanding genome evolution and animal evo-devo.

Computational Prediction of Gene Enhancer Elements by Comparative Genomics

Esko Ukkonen

Department of Computer Science, University of Helsinki, Finland; <http://www.cs.helsinki.fi/u/ukkonen/>

The talk discusses the problem of modelling and finding putative gene enhancer modules (cis-regulatory elements) in the DNA of mammals. A comparative-genomics based model of enhancer modules is described and a Smith-Waterman type dynamic programming algorithm is given for finding modelled enhancers in DNA. We applied this algorithm on the human and mouse genomes and showed in vivo that some of the strongest predicted enhancer modules seem to have biological function. Some further questions related to our model will be discussed.

(Joint work with Outi Hallikas, Institute of Biomedicine, University of Helsinki; Kimmo Palin, Department of Computer Science, University of Helsinki; and Jussi Taipale, Institute of Biomedicine, University of Helsinki.)

Low Duplicability and Network Fragility of Cancer Genes

Davide Rambaldi¹, Federico M. Giorgi¹, Fabrizio Capuani², Andrea Ciliberto² and Francesca D. Ciccarelli¹

¹Department of Experimental Oncology, European Institute of Oncology, ²Istituto FIRC di Oncologia Molecolare (IFOM), IFOM-IEO Campus, Via Adamello 16, 20139 Milan, Italy

Large heterogeneity has been recently reported in the number and types of genes that can undergo driver genetic modifications during cancer development (1-3). Finding the biological basis of this heterogeneity is a key challenge of cancer genetics as it can contribute to a better understanding of the entire tumorigenic process. We applied a combination of genomic and network-based approaches to identify common biological properties within a group of 350 well-known cancer genes. We also extended our analysis to around 250 candidate cancer genes identified through large-scale unbiased mutational screenings. We found that, regardless of their molecular function, cancer genes retain significantly less genomic duplicates than other human genes. In addition, they code for protein hubs occurring within highly interconnected modules of the human protein-protein interaction network. Comparable genomic and network properties recur also within candidate cancer genes, while they differ significantly from those of singleton human genes not involved in cancer. The peculiar properties of cancer genes depict them as particularly fragile components of the human gene repertoire, whose modifications can have multiple deleterious effects. They also contribute to explain the heterogeneity of cancer genetics and support the interpretation of tumor as a “systems disease”.

1. T. Sjoblom *et al.*, *Science* **314**, 268 (2006).
2. C. Greenman *et al.*, *Nature* **446**, 153 (2007).
3. L. D. Wood *et al.*, *Science* **318**, 1108 (2007).

Interacting Gene Clusters and the Evolution of the Vertebrate Immune System

Takashi Makino and Aoife McLysaght

Smurfit Institute of Genetics, University of Dublin, Trinity College, Dublin 2, Ireland

Unravelling the “code” of genome structure is an important goal of genomics research. Co-localization of genes in eukaryotic genomes may facilitate epistasis or co-regulation of functionally-related genes. However, the presence of interacting gene clusters in the human genome has remained unclear. We systematically searched the human genome for evidence of closely-linked genes whose protein products interact. We find 83 pairs of interacting genes that are located within 1 Mbp on the human genome. This number of interacting gene clusters is significantly more than expected by chance, and is not the result of tandem duplications. Furthermore, we find that these clusters are significantly more conserved across vertebrate (but not chordate) genomes than other pairs of genes located within 1 Mbp in the human genome. In many cases the genes are both present but not clustered in older vertebrate lineages. These results suggest gene cluster creation along the human lineage. These clusters are not enriched for housekeeping genes, but we find a significant contribution from genes involved in “response to stimulus”. Many of these genes are involved in the immune response. That these clusters were formed contemporaneously with the origin of adaptive immunity within the vertebrate lineage suggests that novel evolutionary and regulatory constraints were associated with the operation of the immune system.

Chemoperception in the Fruit Fly – The Genome’s Story

Anastasia Gardiner¹, Daniel Barker¹, Roger K. Butlin², William C. Jordan³ and Michael G. Ritchie¹

¹School of Biology, University of St Andrews, St Andrews, Fife, KY16 9TH, UK, ²Animal and Plant Sciences, The University of Sheffield, Sheffield, S10 2TN, UK, ³Institute of Zoology, Zoological Society of London, Regent’s Park, London, NW1 4RY, UK

Chemoperception plays a key role in adaptation and speciation in animals, and the senses of olfaction and gustation are mediated by gene families which show large variation in repertoire size among species. In *Drosophila*, there are around 60 loci of each type and it is thought that ecological specialization influences repertoire size, some studies showed that specialists have fewer functional chemosensory genes than generalists. We analyse the size of the gustatory and olfactory repertoires among the genomes of 12 species of *Drosophila*. We find that repertoire size varies substantially and the loci are evolving by duplication and pseudogenization, with striking examples of lineage-specific duplication. The majority of loci are subject to purifying selection, but this is less strong in gustatory loci and in loci prone to duplication. A few loci show statistically significant evidence of positive selection. Overall genome size is strongly correlated with the proportion of duplicated chemoreceptor loci, but genome size, specialization and endemism may be interrelated in their influence on repertoire size, thus we find that island endemics lose chemosensory genes more rapidly than mainland relatives.

pSNPs: A New Tool for Comparative Genomics?

Ian N. Roberts¹, Michael J. T. O’Kelly², Alexander van Oudenaarden², Stephen A. James¹ and Robert P. Davey¹

¹National Collection of Yeast Cultures, Institute of Food Research, Norwich Research Park, Colney, Norwich, NR4 7UA, UK, ²Department of Physics, Massachusetts Institute of Technology, Cambridge MA 02139, USA

Current methods of genome sequencing do not allow assembly of ribosomal DNA (rDNA) repeats. It is generally assumed by genome projects that these tandem repeats will in any event have been homogenised by evolutionary forces such as sister chromatid exchange and gene conversion. We have investigated rDNA sequence variation in whole genome shotgun sequences from a variety of different strains of *Saccharomyces cerevisiae* and have found unexpectedly high levels of rDNA sequence variation amongst the repeats. We suggest the term pSNPs (partial Single Nucleotide Polymorphisms) to describe this variation. We propose that pSNPs may provide a valuable new measure of genomic interrelatedness and stability.

Acknowledgements: We thank Richard Durbin and David Carter (Sanger Institute) and Ed Louis and Gianni Liti (University of Nottingham) for scientific advice and the BBSRC for financial support.

TOPALi Software for Evolutionary Analysis of Multiple Alignment Data

Frank Wright

Biomathematics & Statistics Scotland, BioSS Office, Scottish Crop Research Institute, Invergowrie, Dundee, DD2 5DA, UK.

TOPALi version 1 specialised in the detection of recombination in multiple alignments. We have redesigned the package to utilize the increased computational power of high performance computing clusters and multicore desktops, and have added model selection and phylogenetic analysis methods, all providing rich, graphical output.

Multiple sequence alignment analysis methods in TOPALi version 2 now include substitution model selection and phylogenetic tree estimation using the Bayesian Inference and Maximum Likelihood approaches. Jobs are submitted from a graphical user interface as web services to either run remotely on high performance computing clusters or locally on a multi-core desktop. The best nucleotide or amino acid substitution model can be selected using accurate statistical criteria derived from Maximum Likelihood co-estimation of the tree and the substitution model, with graphical display of the model parameters. For protein-coding DNA, we have implemented a simple individual codon position model selection procedure that aids model choice for use with Bayesian codon position tree estimation. Phylogenetic software available includes PHYML, RAxML and MrBayes. Tree manipulation tools include midpoint rooting and editing to simplify the display of support values. Data and results are saved to an XML project file.

The ComparaGRID Project: Integrating Genomic Mapping Data using Semantic Web Technologies (Can Formal Ontologies with Logic Reasoners Help Biologists?)

Trevor Paterson¹, Tony Burdett², Robert Davey³, Andrew Gibson⁴, Matthew Pocock⁵, Jo Dicks⁶, Andy Law¹, Peter Rice², Ian N. Roberts³, Robert Stevens⁴ and Anil Wipat⁵

¹ The Roslin Institute, Royal (Dick) School of Veterinary Studies, University of Edinburgh, Roslin, EH25 9PS, UK. ² European Bioinformatics Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge, CB10 1SD, UK. ³ Institute of Food Research, Norwich Research Park, Colney Lane, Norwich, NR4 7UA, UK. ⁴ School of Computer Science, University of Manchester, Oxford Road, Manchester, M13 9PL, UK. ⁵ School of Computing Science, University of Newcastle upon Tyne, Newcastle upon Tyne, NE1 7RU, UK. ⁶ John Innes Centre, Norwich Research Park, Colney Lane, Norwich, NR4 7UH, UK.

Biologists create and maintain vast resources of genomic mapping data available from a plethora of online resources. These span species and application domains, and vary in quantity, quality, detail and format – providing a classic data integration problem.

A query and visualisation architecture that could integrate these disparate datasources would not only allow information to be accessed and navigated more simply, but would facilitate comparative mapping across species boundaries, allowing information in one species to inform the mapping process in another species. This will be particularly valuable for poorly characterised species, by exploiting the well characterised genomic data from ‘model species’.

The approach of the ComparaGRID project has been to provide a data integration architecture which maps individual data schema to a common OWL Domain Ontology and provides OWL Query interfaces which map queries and results through this Domain Ontology and can be accessed and displayed graphically through a browser.

The architecture and implementation of ComparaGRID will be outlined, demonstrating how we can publish datasources as OWL, map these published resources to the shared OWL Ontology and perform data queries, displaying the results graphically to the end user. The generic nature, simplicity, efficiency and usability of the approach will be discussed.

A Bayesian Framework for Data Integration: Application to Comparative Genomics

Madhu Bhattacharjee

School of Mathematics and Statistics, Mathematical Institute, North Haugh, St Andrews, Fife KY16 9SS, UK

In many situations comparison between two organisms can greatly benefit if information from multiple sources could be combined. Depending on the hypothesis of interest these sources can be various. In order to integrate and compare data from the different sources, say e.g. maps, a statistical model is needed which explains all of the relevant entities and their variability, e.g. markers and their locations on maps.

Bayesian statistical inference (and associated computationally intensive algorithms) is a fundamental technology for data integration, inference and prediction. It allows us to fit complex statistical models to diverse data sources, utilising available prior information. This framework can handle errors and uncertainties properly, can propagate uncertainty coherently, which is central to comparative genomics.

As an example we illustrate how we can integrate disparate information from maps to form a comprehensive and higher resolution view of a genome, which can then be used for comparative purposes.

Systems biology in functional genomics: a bioinformatics perspective

Joaquín Dopazo

Department of Bioinformatics, Centro de Investigación Príncipe Felipe,
c/ ep Avda. Autopista del Saler 16 (Junto al Oceanografico), 46013, Valencia, Spain

The ultimate goal of any genome-scale experiment is to provide a functional interpretation of the results, relating the available genomic information to the hypotheses that originated the experiment. Initially, this interpretation has been made on a pre-selection of relevant genes, based on the experimental values, followed by the study of the enrichment in some functional properties. Nevertheless, functional enrichment methods, demonstrated to have a flaw: the first step of gene selection was too stringent given that the cooperation among genes was ignored. The assumption that modules of genes related by relevant biological properties (functionality, co-regulation, chromosomal location, etc.) are the real actors of the cell biology lead to the development of new procedures, inspired in systems biology criteria, generically known as gene set methods. These methods have successfully used to analyze transcriptomic and large-scale genotyping experiments as well as to test other different genome-scale hypothesis in other fields such as phylogenomics.

Poster Abstracts

Comparative and Functional Genomics Identifies Major Differences Between Genomic Islands in Soft Rotting Enterobacterial Plant Pathogens

Leighton Pritchard

Plant Pathology Programme, SCRI, Errol Road, Invergowrie, Perth and Kinross, Scotland, DD2 5DA

The soft rotting enterobacterial plant pathogens *Pectobacterium atrosepticum* (Pba), *Pectobacterium carotovorum* (Pcc) and *Dickeya dadantii* (Dda) are closely-related but differ in their host ranges, geographical distributions, and survival in the environment. Each organism causes disease by a similar mechanism, namely the production of plant cell wall degrading enzymes, but the molecular interactions and processes that distinguish between the course of disease in each case are largely unknown. Here we present results from a study describing the extent of differential horizontal gene transfer in these pathogens. We used computational and microarray comparative genomic hybridisation (M-CGH) techniques to identify genomic islands in Pba strain SCRI1043 that are absent or divergent in Pcc strain SCRI193 and/or Dda strain 3937. Such islands may make a contribution to Pba1043-specific phenotypes, niche adaptation or pathogenicity in Pba1043. Many identified islands are composed mostly of genes with no annotated function, but several islands contain genes that make a known or potential contribution to pathogenesis. We demonstrate that there is no simple relationship between hybridisation affinity and sequence identity for M-CGH, and that the predictive accuracy of M-CGH for cross-species studies is low. We propose a HMM-based model for improving the predictive accuracy of M-CGH for cross-species studies.

A eukaryotic phylogeny as an initial step to a functional annotation pipeline of protein coding gene products

Saif Ur-Rehman and Daniel Barker,

Sir Harold Mitchell Building, School of Biology, University of St Andrews, St Andrews, Fife, KY16 9TH, UK

Functional annotation of gene products is a problem for which rapid, accurate computational solutions are particularly attractive. One method of computationally predicting the function of proteins is to localize them within modules of known function, on the assumption that genes showing correlated gain and loss over evolutionary time have functionally linked products. This allows annotation on the established principle of 'guilt by association'. As an initial step towards an evolutionary context-based annotation pipeline, a rooted phylogeny of 54 eukaryotic species including 19 animals, 4 plants, 16 fungi and 15 protists has been reconstructed. Nine low copy number orthologous proteins were selected through the use of all-against-all Smith-Waterman searches (SSEARCH software) and the Inparanoid clustering algorithm. Multiple sequence alignments were generated for each ortholog and then concatenated to form a super-matrix. The phylogeny was reconstructed using Bayesian Markov chain Monte Carlo (Bayesian-MCMC; MrBayes). It is planned to use maximum likelihood methods to examine gains and losses of genes on a Bayesian-MCMC sample of trees (bms_runner and BayesTraits). Genes of unknown function will be examined against genes within known biochemical pathways, thus providing novel functional annotation.

The RAST Server: Rapid Annotations using Subsystems Technology

Olga Vasieva (representing 22 authors*). Fellowship of Interpretation of Genomes

The number of prokaryotic genome sequences becoming available is growing steadily and is growing faster than our ability to accurately annotate them.

We describe a fully automated service for annotating bacterial and archaeal genomes. The service identifies protein-encoding, rRNA and tRNA genes, assigns functions to the genes, predicts which subsystems are represented in the genome, uses this information to reconstruct the metabolic network and makes the output easily downloadable for the user. In addition, the annotated genome can be browsed in an environment that supports comparative analysis with the annotated genomes maintained in the SEED environment.

The service normally makes the annotated genome available within 12–24 hours of submission, but ultimately the quality of such a service will be judged in terms of accuracy, consistency, and completeness of the produced annotations. We summarize our attempts to address these issues and discuss plans for incrementally enhancing the service. By providing accurate, rapid annotation freely to the community we have created an important community resource. The service has now been utilized by over 120 external users annotating over 350 distinct genomes.

*: Ramy K Aziz^{8,9}, Daniela Bartels³, Aaron A Best⁷, Matthew DeJongh⁷, Terrence Disz^{2,3}, Robert A Edwards^{1,2}, Kevin Formsma⁷, Svetlana Gerdes¹, Elizabeth M Glass², Michael Kubal³, Folker Meyer^{2,3}, Gary J Olsen^{4,2}, Robert Olson^{2,3}, Andrei L Osterman^{1,5}, Ross A Overbeek¹, Leslie K McNeil⁶, Daniel Paarmann³, Tobias Paczian³, Bruce Parrello¹, Gordon D Pusch^{1,3}, Claudia Reich⁶, Rick Stevens^{2,3}, Olga Vassieva¹, Veronika Vonstein¹, Andreas Wilke³ and Olga Zagnitko¹

¹Fellowship for Interpretation of Genomes, Burr Ridge, IL 60527, USA; ²Mathematics and Computer Science Division, Argonne National Laboratory, Argonne, IL 60439, USA; ³Computation Institute, University of Chicago, Chicago, IL 60637, USA; ⁴Department of Microbiology, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA; ⁵The Burnham Institute, San Diego, CA 92037, USA; ⁶National Center for Supercomputing Applications, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA; ⁷Hope College, Holland, MI 49423, USA; ⁸University of Tennessee, Health Science Center, Memphis, TN 38136, USA; ⁹Department of Microbiology and Immunology, Cairo University, Cairo, Egypt

Preliminary survey of desaturase genes in 12 *Drosophila* species

Maria Keays, School of Biology, Harold Mitchell Building, University of St Andrews, KY16 8TH

In *Drosophila*, desaturases are involved in cuticular hydrocarbon (CHC) production. As well as playing an important physiological role, CHCs act as sex pheromones and are thought to be important in mate recognition. The desaturase genes are therefore candidate “speciation genes”, as changes in their function could cause changes in mate recognition, eventually precipitating formation of new species. Studying the evolution of this family will provide insights into its role in the evolution of the *Drosophila*. The first step in this process is to assemble a dataset of orthologous sequences. This was done for the 12 species whose genome sequences are available, using TBLASTN and GeneWise. Nine putative desaturases from *D. melanogaster* were used to search the other 11 genomes. Given this initial dataset, MEME/MAST were used to detect any genes that may have been overlooked. Of particular interest is Fad2, which is thought to be responsible for gender-specific differences in *D. melanogaster* CHCs. It was found in eight of the 12 species, all of which belong to the *Sophophora* subgenus. The remaining three species, which appear to lack this gene, are not members of this subgenus. Fad2 is intronless, indicating that it is likely to be a retrogene.

Relator: a visualization tool for relating unigene datasets

Linda Cardle and David Marshall,
SCRI, Invergowrie, Dundee, DD2 5DA

For organisms that have not been completely sequenced, many researchers have turned to using unigene sets to design probes for microarrays and genetic maps. In the case of *Hordeum* and *Solanum* species there are several different versions of unigene sets, all of which have been used as a basis for probe design. As a result there is a need for a method of relating equivalent unigenes to allow data from different sources to be drawn together.

Several unigene sets were compared to each other and themselves using BLAST searches. Additional BLAST searches were carried out against appropriate genomes for which there was a full genomic sequence available in order to be able to identify gene families and annotate the unigene sets. The BLAST results were stored in a relational database and an interface was created to query the database and display the results in several different formats.

The Relator webtool provides a useful utility that allows researchers to draw comparisons between sequences from different sets of unigenes, both within and between species.

A vision of Comparative Genomics through next generation sequencing

Urmi Trivedi, Sujai Kumar, Marian Thompson, Nicola Wrobel, Jill Lovell, Andrew Gillies, Anna Mottazzam and Mark Blaxter
Institute of Evolutionary Biology, The Ashworth laboratories, The University of Edinburgh, King's Buildings Campus, Edinburgh EH9 3JT

DNA sequence data lies at the heart of the ongoing genomics revolution in biology. The Gene Pool at the University of Edinburgh offers DNA sequencing and genotyping services to the academic community, using various platforms. Along with medium throughput Sanger sequencing using ABI3730 capillary instruments, we also offer ultra high throughput sequencing based on Illumina-SOLEXA and Roche 454 instrumentation. These new DNA sequencing platforms deliver substantially increased volumes of data at dramatically lower costs, but demand new analytical methods to take full advantage of the short reads that they produce. The Gene Pool is developing solutions to these challenges, including de novo genome assembly, genome resequencing, finding polymorphisms, digital transcriptomics and ChIP-Seq. We will discuss what the data from next-generation sequencing projects look like when they come off the Illumina Solexa sequencer (quality values, probabilities of error towards the end of a read, etc.) and the analytic methodologies followed by the Gene Pool. Specifically, we will highlight (a) the informatics steps for comparison of bacterial strains from species that have not been completely sequenced, (b) initial analyses of a comparative nematode genomics study, and (c) best parameters for detecting SNPs in new strains when compared to a reference genome using Solexa data.

Modelling of compensatory properties of large scale gene/metabolic networks carrying series of mutations. Application to purine biosynthesis of *E.coli* strains

¹Goltsov, Alexey, ¹Lebedeva, Galina, ¹Goryanin, Igor ²Demin, Oleg

¹ Centre for Systems Biology at Edinburgh, University of Edinburgh, UK

²A.N. Belozersky Institute of Physico-Chemical Biology, Moscow State University, Moscow, Russia

Kinetic modeling based approach has been developed to simulate the effect of different genetic modifications (knockouts, mutations) on the dynamic properties of bacterial metabolic networks. The approach allows for comparative analysis of the metabolism in different bacterial strains and elucidation of the regulatory relationship between the series of mutations and the responses of metabolic networks to the external and internal perturbations. It also allows to obtain realistic estimates for the growth rate of the derived strains.

We focused on *in silico* screening of the mutation combinations leading to block/restoration of cellular growth: 1) growth block mutations in two genes, that lead, when combined, to growth arrest; 2) restoring mutations, among which one is a growth arrest, but when combined with the second one the cell growth is restored.

We applied our method to model dynamic properties of different *E.coli* guanosine auxotrophs carrying the following block/restoring mutations in the genes involved in purine nucleotide biosynthesis: *purE*⁻ ($T_2=78$ min); *purE deoD*⁻ (700 min); *gsk*^{*} (800 min); *purE deoD gsk*^{*} (111 min); *purE deoD gsk purF*⁻ (56 min); *purE deoD gsk prs*^{*} (58 min); and *gsk hpt gpt* (600 min). (T_2 - doubling time).

The results of modelling have shown, that the growth arrest of the guanosine auxotrophs with various sets of mutations results mainly from the disturbance of the gene and metabolic regulatory systems of purine biosynthesis. The origin and kinetic mechanism of growth restoring mutations was elucidated: it results from compensatory effects, when the disturbance in one feedback loop is compensated by the changes in the strength of another feedback.

We discuss possible applications of the developed method for evolutionary biology studies (gene transfer in the evolution of bacteria nucleotide biosynthesis) as well as for drug design, as a tool to predict promising drug targets and analyze/predict/overcome drug resistance effects due to compensation mechanism of gene/metabolic networks.

Meeting Information

Organised by

Daniel Barker, School of Biology, University of St Andrews,
Madhu Bhattacharjee, School of Mathematics and Statistics, University of St Andrews
Chris Janssen, Scottish Bioinformatics Forum,
Sandra Borthwick, Scottish Bioinformatics Forum.

Sponsored by

Scottish Bioinformatics Forum <http://www.sbforum.org>

Bute Medical School, University of St Andrews <http://medicine.st-andrews.ac.uk>

Centre for Evolution, Genes and Genomics, University of St Andrews
<http://biology.st-and.ac.uk/cegg>

Gatty Marine Laboratory, University of St Andrews <http://biology.st-andrews.ac.uk/gatty>

School of Mathematics and Statistics, University of St Andrews
<http://www.mcs.st-andrews.ac.uk>

With Support from

School of Biology, University of St Andrews <http://biology.st-andrews.ac.uk>