

Computational Systems Biology and Information Theoretic Approaches to Gene Regulatory Networks, Development, Evolution, and Sustainability

Jan T Kim, j.t.kim@herts.ac.uk

Evolution can formally be modelled as a process in which mutations, i.e. genotypic variation, result in variation in phenotypic traits, such as height, physical strength and agility, crop yield, and also sensory acuity such as colour vision and odour perception. Phenotypic traits can be determined by genomic information in a fixed way, as illustrated by the genetic determination of flower colour in peas, first described by Gregor Mendel. However, most traits are not “genetically hard-coded” in this way. Many traits vary during an individual’s life, often in response to environmental conditions. Human skin colour, which changes in response to exposure to sunlight, is an example of such controlled variation. Further examples include changes in musculature and body shape in response to nutrition and exercise in humans as well as many animals, and dramatic shifts in foliage and growth patterns displayed by plants in response to sunlight.

These observations highlight the genome (i.e. the complete genetic information inherited by an individual) has a more complex role than that of a “blueprint”, which determines phenotypic traits, possibly with limited precision resulting in random variation to some degree. From a computer science perspective, it is more adequate to conceive of the genome as a kind of program code which takes signals from the environment as its input and generates phenotypic traits that are adapted to the individual’s current environment as its output. Morphogenesis is a key biological process in which specific phenotypic traits are realised at each developmental stage, as highlighted by “evo-devo” research which combines evolutionary and developmental biology with genetics.

The code of genomic “programs” that realise morphology and other complex traits is unlike that of typical programming languages, however. Gene Regulatory Networks (GRNs), which are comprised of genes which activate or repress other genes, are a central mechanism realising adaptive flexibility. Some formalisations of GRNs are similar to Artificial Neural Networks (ANNs), suggesting that theoretical concepts of neurocomputing can be fruitfully transferred to GRN research. More specifically, a number of formalisms and computational methods to model GRNs have been developed, including the transsys framework for GRN modelling [<https://github.com/jttkim/transsys>].

The Adaptive Systems Research Group [<http://adapsys.cs.herts.ac.uk/>] has established an extensive information theoretic basis for the principled study of sensorimotor systems, sensor evolution, and other learning systems. Such approaches have enabled key insights into the evolution of transcription factors and their binding sites, which are key constituents of GRNs. More recently, we have introduced an approach for quantifying sustainability based on information theory in [this article](#).

My research focuses on advancing principled scientific understanding of biological information in evolution and GRNs, and on applying information theory at key interfaces of biological systems and technology, such as sustainability and rational design of (crop) plant morphology. PhD projects are available addressing the following research questions:

- What are the key properties of gene regulatory networks? Can new insights from artificial neural networks, including “deep learning” multi-layer networks, be used to advance our understanding of GRNs?
- How do GRNs and evolution interact? During an individual’s lifetime, only a fraction of its genomic potential may be activated and exposed to selection. How does that affect the dynamics of evolution, especially of the structure of GRNs? Can we predict GRN features shaped by different evolutionary forces? And can we detect them in GRNs inferred from high throughput data from molecular biology? What is the role of epigenetics, i.e. of transient reversible modifications of genetic information? What are the effects of GRN structures on evolvability?
- What is the potential of GRN engineering to improve food security in the face of climate change? Extreme conditions such as drought or flooding, constitutes major changes in the environment of crop plants, often resulting in crop failures. Responding with adaptive phenotypic traits could protect crop yield in such conditions. Can the potential of crop varieties, and of their wild relatives, be predicted from their GRNs? And can we use such predictions to inform breeding crop varieties adapted to climate change?
- How can the concepts of relevant information, empowerment, information bottlenecks, and other approaches rooted in information theory be applied to GRNs and their evolution?
- Can empowerment and other information theoretic measures, which have been used for modelling agents in Machine Learning, Artificial Intelligence and Artificial Life, be used to capture criteria of sustainability for systems requiring long-term viability?

Projects will be conducted in an interdisciplinary environment, comprised of scientists from the Adaptive Systems Research Group and bioscience collaborators based at the University or at partner organisations such as Rothamsted Research. Candidates will be required to develop interdisciplinary expertise. We would normally expect candidates to be suitably qualified in either computer science or the biosciences (possibly by degrees and experience in closely related areas such as mathematics, statistics or medical sciences), and to demonstrate their ability to acquire knowledge and expertise in any required areas they are not yet familiar with.

I welcome informal enquiries from anyone with an interdisciplinary profile and interest as outlined above. Please get in touch by email to j.t.kim@herts.ac.uk.