A Complex System

The applications of complex system theory are wide-ranging within the analysis of biological systems. Considering the cell as a complex system could open up new possible opportunities in drug discovery

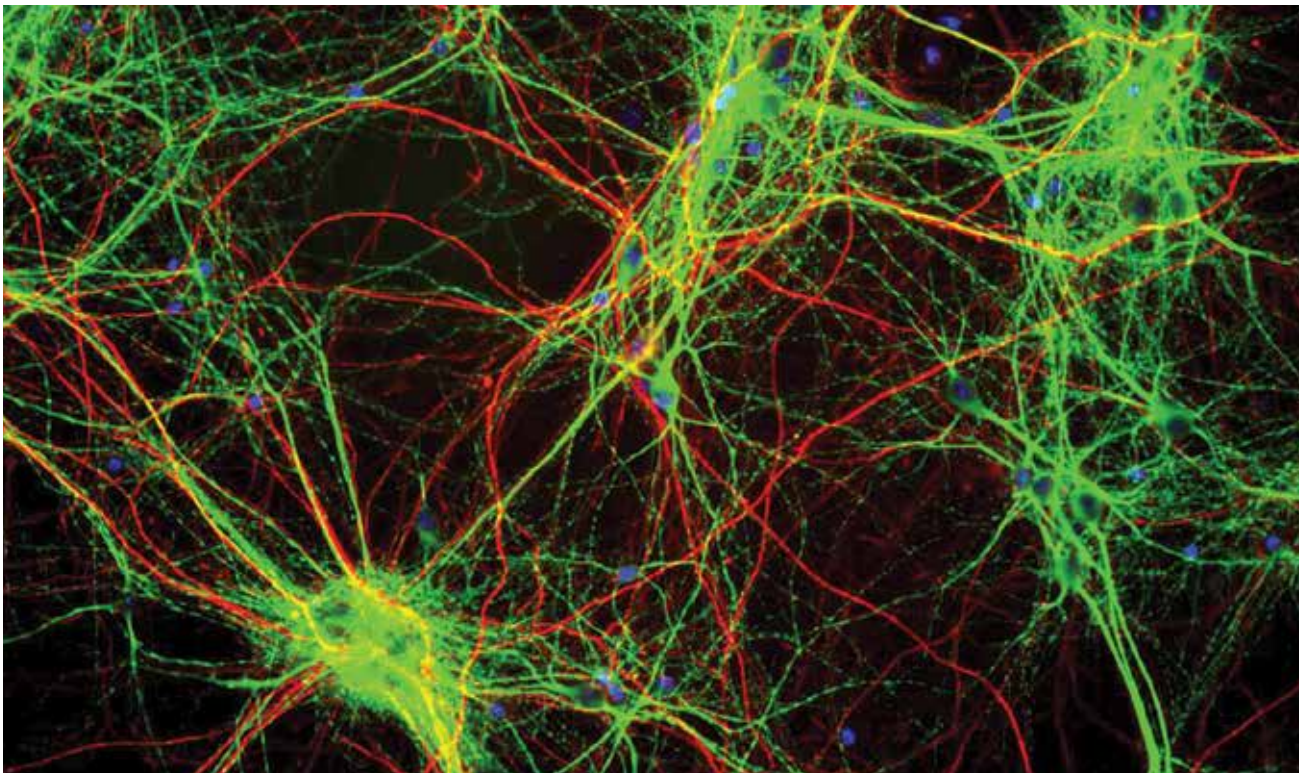
Jonny Wray at e-Therapeutics

The concept of complex adaptive systems has a long history of use in describing and analysing the behaviour of biological systems across multiple different scales – from collective animal behaviour to the emergence of cellular function from the interaction of large collections of molecules. The key characteristics of any complex adaptive system are that the behaviour of the system emerges due to the interactions between multiple – typically much simpler – entities, and that the behaviour of the system cannot be predicted by linear combinations of the isolated behaviour of the individual elements.

In addition, these systems adapt to changes in both the environment and underlying elements. Complex system theory aims to provide a framework in which non-linear behaviour can be understood and predicted. When applied to biological systems, and considered in the context of natural selection, these properties are thought to give rise to system resilience, robustness and evolvability (1,2).

Consideration of the cell as a complex system has consequences for drug discovery. Cellular function, including disease, arises due to interactions between proteins, second-messengers and intrinsic small molecules within the cell (3). Cellular mechanisms of disease can be considered as a collection of pathological interactions that only occur in that disease state (4). This view of disease mechanism has found multiple recent applications including the hallmarks of cancer, immuno-oncology, cardiovascular disease and more.

If we can consider the cellular mechanisms underlying disease as making up a complex system, then drug discovery aimed at combating those diseases can be regarded as the identification of agents that have a significant effect when used to perturb those systems. Approaching the discovery of new therapeutic agents from this direction has several theoretical benefits. A number of recent investigations into the reasons for drug development failure in the clinic highlighted that many of these failures are due to efficacy which dominate in late-stage clinical trials (5,6). Approaches that explicitly consider biological complexity during the compound discovery phase have the potential to vastly reduce the number of such failures, which will be of benefit to the industry and patients alike (7). Explicitly targeting system function during drug



discovery should be better placed to address complex diseases that arise due to interactions between multiple components, to tackle inherent cellular robustness mechanisms, reduce the capacity to develop resistance, and aid in the discovery of personalised therapeutics (2,8,9).

Systems-driven approaches have an implicit history within drug discovery. Phenotypic screening – the historically traditional approach to discovery – is effective since it recognises that testing within intact cellular or animal systems is likely to be more predictive of what will happen in man than the isolated protein assays typically used in more modern high-throughput screening (10). However, the complexity of these assays has limited the practical application to the screening of very large compound libraries, highlighting the need for a rational approach to compound selection for phenotypic screening.

The robustness and resilience properties of complex systems imply they can withstand the failure, or functional perturbation, of small numbers of their constituent elements. Thus, substantial levels of change in system behaviour can only occur when multiple elements are perturbed. Nonlinearities imply that the identification of such element collections is not trivial, and is certainly not as simple as choosing those that appear most important individually. In the context of drug discovery, this leads to the concept of the identification of target sets or collections of multiple proteins that, when agitated simultaneously, can have a large effect on biological function. It is important to note that drug discovery driven by the search for optimal target sets does not imply the need for small molecules that actively bind to all the proteins in those target sets. In fact, the concept of a target set should be considered as the desired downstream, pleiotropic footprint of a small molecule that could well assert its influence via a single binding target.

Practical Application

The experimental identification of effective target sets within an intact cellular system is tricky. State-of-the-art experimental techniques are able to perturb a small, on the order of three or four, number of genes or proteins simultaneously (11). Combinatorial explosion, along with practical limitations, hamper the investigation of larger numbers (12). Conversely, computational approaches are ideally suited for problems plagued by combinatoric issues.

Three major components are required for the implementation of a practical, in silico, systems-based approach to drug discovery. First, the ability to construct computer models representing the cellular disease mechanisms to be targeted. Second, analysis techniques that can utilise these models to identify effective protein target sets. Finally, the ability to find compounds whose downstream, pleiotropic protein signature matches those identified in the previous step. The goal is to take advantage of the inevitable pleiotropic effects of drug application and use that footprint explicitly during discovery, rather than consider these effects as undesirable (13).

Networks can be considered a mathematical abstraction of a complex system and have become a very useful tool for modelling the molecular interactions within a cell and guiding the understanding of integrated biological function (14,15). They are, therefore, an ideal computational approach to use for modelling cellular disease mechanisms. Percolation theory applied to complex networks is concerned with exploring the change in network structure and behaviour due to perturbation of collections of system elements. As such, it forms a suitable theoretical framework to develop analysis approaches aimed at the identification of effective target sets in disease networks. It should be noted that the network biology approach is distinct to the common definition of systems biology, which is the application of dynamical systems theory to cellular modelling.

It has been possible to design and implement a discovery platform based on just such an approach. Cellular disease mechanisms are modelled as protein interaction networks, capturing the pathological relationships believed to give rise to the disease phenotype. Analysis techniques from network science and optimisation theory are used to identify target sets capable of significant effects on the structural integrity of the disease networks. These approaches are motivated from the underlying assumption that causing structural change to the network will predict large effects on disease phenotype.

Finally, compounds are identified whose protein footprint is predicted to be close to the desired footprints identified by network analysis. This step utilises a compound bioactivity database consisting of both empirical evidence and evidence from machine learning models that predict compound activity. The output of the computational process are lists of compounds believed to be enriched in actives, which are then tested in a suite of cell-based, phenotypic assays representing the disease mechanisms being targeted. Hits, defined as compounds active across the whole assay suite, are repeated and then taken into a phenotypic-driven medicinal chemistry programme for hit to lead optimisation. The fact that the computational platform produces a statistical output

“The goal is to take advantage of the inevitable pleiotropic effects of drug application and use that footprint explicitly during discovery, rather than consider these effects as undesirable”

(a collection of compounds enriched in actives), rather than predicting activity in small numbers of compounds, reflects the explicit consideration throughout the computational process of the effects of noisy and missing data.

The core motivation of this systems-driven approach to discovery is to produce molecules much more likely to be active in man. Ultimate validation of the approach will, therefore, be clinical results on candidates that are produced from the platform in the coming months and years. However, several internal discovery programmes are under way and have produced large numbers of hits across a range of biologies.

In addition to explicitly tackling the complexity of disease biology, the approach opens up some novel potential avenues for discovery. For example, gene regulation has long been considered a desirable intervention target but transcription factors are generally considered undruggable (16,17). Regulation is well known to be a network effect caused by collections of molecular regulators interacting with each other. Reconstruction of those regulatory networks from diseased cells, followed by their use in network-driven discovery, opens up a novel way of approaching mechanisms previously thought to be undruggable. Recent work using network-based analysis of next-generation sequencing data from human disease tissue suggests cancer subtypes are better explained at a network level than an individual gene or protein level (18). Consideration of cancer subtypes in such a network framework would facilitate patient segment-specific discovery.

References

1. Kitano H, Biological robustness, *Nature Reviews Genetics* 5: pp826-837, 2004
2. Tian T *et al*, The origins of cancer robustness and evolvability, *Integrative Biology* 3: p17, 2011
3. Hartwell LH *et al*, From molecular to modular cell biology, *Nature* 402: pp47-52, 1999
4. Schadt EE, Molecular networks as sensors and drivers of common human diseases, *Nature* 461: pp218-223, 2009
5. Kola I and Landis J, Can the pharmaceutical industry reduce attrition rates?, *Nature Reviews Drug Discovery* 3: pp711-715, 2004
6. Cook D *et al*, Lessons learned from the fate of AstraZeneca's drug pipeline: A five-dimensional framework, *Nature Reviews Drug Discovery* 13: pp419-431, 2014
7. Ahn AC *et al*, The limits of reductionism in medicine: Could systems biology offer an alternative?, *PLoS Medicine* 3: e208, 2006
8. Kitano H, A robustness-based approach to systems-oriented drug design, *Nature Reviews Drug Discovery* 6: pp202-210, 2007
9. Chen R and Snyder M, Systems biology: Personalized medicine for the future?, *Current Opinion in Pharmacology* 12: pp623-628, 2012
10. Zheng W *et al*, Phenotypic screens as a renewed approach for drug discovery, *Drug Discovery Today* 18: pp1,067-1,073, 2013
11. Sahin Ö *et al*, Combinatorial RNAi for quantitative protein network analysis, *Proceedings of the National Academy of Sciences* 104: pp6,579-6,584, 2007
12. Lambeth LS *et al*, A direct comparison of strategies for combinatorial RNA interference, *BMC Molecular Biology* 11: p77, 2010
13. Davignon J, Beneficial cardiovascular pleiotropic effects of statins, *Circulation* 109: pp39-43, 2004
14. Albert R and Barabási A-L, Statistical mechanics of complex networks, *Reviews of Modern Physics* 74: p47, 2002
15. Barabási A-L and Oltvai ZN, Network biology: Understanding the cell's functional organization, *Nature Reviews Genetics* 5: pp101-113, 2004
16. Darnell JE, Transcription factors as targets for cancer therapy, *Nature Reviews Cancer* 2: pp740-749, 2002
17. Yan C and Higgins PJ, Drugging the undruggable: Transcription therapy for cancer, *Biochimica et Biophysica Acta (BBA) - Reviews on Cancer* 1,835: pp76-85, 2013
18. Le Van T *et al*, Simultaneous discovery of cancer subtypes and subtype features by molecular data integration, *Bioinformatics* 32: pp445-454, 2016

About the author

Jonny Wray is currently Head of Discovery Informatics at e-Therapeutics and has led the conceptual development of the network-driven discovery approach and the design of the software platform embodying that method. He was a Fellow in Theoretical Neuroscience at The Neurosciences Institute, US, where his research was aimed at understanding how system function arises from network structure.
Email: contact@etherapeutics.co.uk