Pareto front a system lies on could inform us about which tasks it needs to achieve well. Going back to where we started, as hairball-like high-dimensional datasets become more and more complete, they can perhaps help us identify the objectives, or tasks, for which a system has been optimized. Indeed, the set of connections described in these hairballs have arisen via the sampling of a large phenotypic space where tradeoffs have been made, and the topologies chosen should outline the Pareto fronts of the relevant sets of tasks or goals (Hart et al., 2015). It will be fascinating in the future to consider how combinations of network motifs perform together and how they might arise from an evolutionary perspective. Do systems’ tasks acquire new meanings when network motifs work together; are there higher-level tradeoffs? When considering biological systems as dynamic entities that are optimized over evolutionary timescales to perform certain tasks, we still have a lot to learn.

REFERENCES

Synthetic Gene Circuits Learn to Classify
Andriy Didovyk1,2 and Lev S. Tsimring1,2,*
1BioCircuits Institute, University of California, San Diego, 9500 Gilman Dr., La Jolla, CA 92093, USA
2San Diego Center for Systems Biology, 9500 Gilman Dr., La Jolla, CA 92093, USA
*Correspondence: ltsimring@ucsd.edu
http://dx.doi.org/10.1016/j.cels.2017.02.001

An efficient computational algorithm is developed to design microRNA-based synthetic cell classifiers and to optimize their performance.

Classification and pattern recognition problems are traditionally the domain of computer science. Typically, a classification system is built to choose one of a finite repertoire of outputs based on the analysis of multiple inputs. Such classifiers are usually designed and trained using sophisticated machine-learning algorithms. In recent years, synthetic biologists have begun to explore the possibilities of engineering living cells to perform classification tasks. There are many important potential applications of this approach—from building automated water quality monitoring systems to identifying abnormal cells in situ in living organisms. Designing a synthetic gene circuit for a particular task is usually done by trial and error and requires laborious manual parameter tuning to achieve acceptable performance. In this issue of Cell Systems, Yaakov Benenson and his collaborators (Mohammadi et al., 2017) put forth a computational workflow to streamline and automate this tedious and highly non-trivial process (Figure 1). They use this workflow to design a classifier gene circuit that responds to complex microRNA (miRNA) expression patterns for recognition and selective targeting of pathological cell types.

For a synthetic gene circuit with sufficiently complex functions, the choice of optimal topology may not be obvious, thus calling for a systematic automated approach to find it. Overall, a complete gene circuit design is a two-step process. One must decide on a circuit topology (discrete optimization) and also select biochemical parameters that optimize the circuit performance (continuous optimization). In general, these two tasks are not independent of each other, since different circuit topologies may yield optimal performance at different values of parameters. François and Hakim (2004) were, to our knowledge, the first who proposed a computational evolutionary algorithm for the automatic design of gene networks with desired properties. They illustrated the feasibility of this approach by generating robust bistable switch and oscillator circuits in silico. However, evolutionary algorithms operating on comprehensive mathematical models of synthetic gene circuits were computationally costly and only applicable to designing relatively small networks. To alleviate this problem, Marchisio and Stelling (2011) developed an algorithm that could efficiently design large logical gene networks, i.e., networks where inputs and the output only take “0” or “1” values.

Classifiers form an important class of synthetic gene circuits with potentially
high-impact applications. The possibility of making synthetic miRNA-based classifier circuits that can be delivered directly into the host to identify and destroy cancer cells “from the inside” was recently demonstrated experimentally by Xie et al. (2011) and Wroblewska et al. (2015). However, they were designed and optimized “manually.” Until now, there have been no computational frameworks for designing synthetic gene classifiers de novo. Unlike purely binary circuits, the challenge in designing classifiers automatically is to engineer a gene circuit which maximizes the chance of producing a correct digital classification answer (e.g., a binary decision whether a cell is cancerous or healthy), given graded and noisy inputs.

To address this challenge, Mohammadi et al. (2017) develop an integrated workflow that allows one to automatically design such miRNA-based classifiers for cancer detection by addressing the dual nature of the optimization process described above. First, the authors introduce mechanistic models of elementary logical gene circuits that generate scalar outputs from small subsets of inputs (miRNA levels). Using these models, they compute the circuit parameters that maximize the On/Off ratio of positive and negative responses. The optimal parameters generated for miRNA regulation of individual gene expression were nearly the same for all elementary circuits, thus allowing the authors to avoid iterative parameter optimization after evolutionary changes of classifier topology and to significantly simplify and streamline the overall optimization strategy. Of course, real gene circuits may not completely conform to the model assumptions or allow the parameters to be tuned to their optimum values, and so accounting for the variability in parameters in the evolutionary search for the best topology may still emerge as an issue.

In the second part of the paper, Mohammadi et al. (2017) develop two evolutionary training algorithms for finding the optimal classifier circuit topology using large sets of positive and negative examples. In the first algorithm, specific computational models of the individual logical modules are tested and “evolved” to maximize the area under the operating characteristic curve of the circuit. In the second, general implementation, the authors use a Boolean approximation of the classifier circuit’s logical structure. Since the circuit parameters have already been tuned to provide maximally sharp response to the respective miRNA inputs, the general optimization algorithm yields nearly as good results as the more computationally expensive model-specific optimization. Overall, both algorithms demonstrate nearly errorless and very robust performance in silico with both simulated and real-world miRNA expression datasets.

The general approach developed by Mohammadi et al. (2017) opens the possibility to automate the design of various synthetic classifiers for biomedical or biotechnological applications, far beyond miRNA-based circuits. Of course, such in silico circuit optimization relies on the availability of sufficiently accurate computational models of elementary biochemical reactions involved in its function, which so far is not always the case. Other potentially important factors affecting the optimization performance are the cell-to-cell variability, stochastic fluctuations in miRNA levels, and the molecular noise in the classifier circuit itself. The choice of the optimal topology and the corresponding parameters can be strongly affected by such noise-performance considerations. Although identification of cancer in individual cells is an important task, in other applications, it can be beneficial to take advantage of distributed multi-cellular decision-making to reduce the effects of single-cell variability and noise and hence increase the robustness of classification. Some recently proposed synthetic circuits for detecting cancer tumors (Anderson et al., 2006; Danino et al., 2015; Din et al., 2016) use quorum sensing to trigger a collective decision by the population of engineered cells to produce either a diagnostic marker or attack the detected tumor in situ. One can also engineer non-monoclonal cell populations to solve
even more complex classification and decision-making problems, when single-cell classification is difficult. Boosting the performance of a classifier by combining inputs of multiple distinct weak classifiers is a well-known meta-algorithm in machine learning (Freund and Schapire, 1999). Theoretically, the outputs from individual cells can be appropriately combined to produce a robust classification output (Didovyk et al., 2015). Of course, design and training of such distributed classifiers present additional challenges that will need to be overcome in the future. Such classifiers, similarly to their intra-cellular counterparts, can be “trained” in silico using the computational algorithms akin to those described by Mohammadi et al. (2017) and then implemented in vivo. However, the real challenge is to figure out practical methods for continuous training and adaptation of cell-based synthetic classifiers to changing environments directly in vivo, bypassing the need for in silico training.

In summary, the paper by Mohammadi et al. (2017) establishes a novel computational platform for designing synthetic gene classifiers and demonstrates its utility by generating promising miRNA-based classifier circuit topologies to be tested experimentally. We expect that future experimental and computational work will validate, refine, and expand this design paradigm.

REFERENCES


Protein Networks’s Alzheimer’s Disease

Eva Meier Carlsen1 and Rune Rasmussen2,*
1Department of Neuroscience and Pharmacology, Faculty of Health and Medical Sciences, University of Copenhagen, Blegdamsvej 3, 2200 Copenhagen N, Denmark
2The Danish Research Institute of Translational Neuroscience – DANDRITE, Nordic EMBL Partnership for Molecular Medicine, Department of Biomedicine, Aarhus University, Ole Worms Allé 3, 8000 Aarhus C, Denmark
*Correspondence: runerasmussen@dandrite.au.dk
http://dx.doi.org/10.1016/j.cels.2017.02.006

Overlap of RNA and protein networks reveals glia cells as key players for the development of symptomatic Alzheimer’s disease in humans.

Memory loss, and the other devastating symptoms of Alzheimer’s disease, are likely the result of complex molecular changes in the brain. Importantly, these molecular changes may precede outward clinical symptoms of the disease by years. In a recent article in Cell Systems, Seyfried et al. (2017) compared protein expression in human brain tissue from healthy individuals versus tissue from patients with Alzheimer’s. The authors showed that measurements of protein expression complement gene expression data, which have been collected in other studies, to help understand the etiology and consequences of disease. Moreover, unravelling networks of cell-type-specific proteins associated with the onset and progression of Alzheimer’s disease may be useful for early diagnosis.

A critical player in Alzheimer’s disease pathology is amyloid-beta (Aβ) deposition, resulting in Aβ plaques in the brain. But lately it has become clear that Aβ is not alone in causing Alzheimer’s disease. The severity of symptoms does not scale linearly with the amount of plaques in the brain and Aβ plaques are also found in healthy elders without any apparent cognitive impairment (Elman et al., 2014; Bennett et al., 2006). Thus, our understanding of Alzheimer’s disease pathology still seems too simplistic and limited, and Alzheimer’s disease is likely determined in a multi-factorial parameter space.

The current available pharmacological treatments aimed to protect Alzheimer’s disease patients are aimed largely at reducing Aβ (Yiannopoulou and Papa-georgiou, 2013), but the results so far have been disappointing. This underscores the view that Alzheimer’s disease etiology and progression are not simply a reflection of the amount of Aβ in the