# Integrative Biology

Accepted Manuscript






ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/ibiology

Synthetic biology involves the engineering of life either to construct functional biological devices or to test the functioning of biological systems, thus achieving countless insights into the inner workings of the cells. Here, we provide an overview of the impact that the field of synthetic biology has had in the areas of gene expression, cell heterogeneity (noise), coupling of growth and energy usage to expression, and spatiotemporal dynamics of cellular machinery. We compare bacterial and mammalian systems, which currently provide some of the most-developed engineering frameworks. Over the last decade, the many insights that have arisen from synthetic biology have triggered an exciting transition from "creating in order to understand" towards "creating in order to cure."

# From noise to synthetic nucleoli: can synthetic biology achieve new insights?

**Authors**: Marta Ciechonska[1,2], Alice Grob[1,2] & Mark Isalan[1*]

1. Department of Life Sciences, Imperial College London, London SW7 2AZ, UK.
2. Authors contributed equally

*Corresponding author: E-mail, m.isalan@imperial.ac.uk (MI)

**Main text: 4844 words**

## Abstract (141 words)

Synthetic biology aims to re-organise and control biological components to make functional devices. Along the way, the iterative process of designing and testing gene circuits has the potential to yield many insights into the functioning of the underlying chassis of cells. Thus, synthetic biology is converging with disciplines such as systems biology and even classical cell biology, to give a new level of predictability to gene expression, cell metabolism and cellular signalling networks. This review gives an overview of the contributions that synthetic biology has made in understanding gene expression, in terms of cell heterogeneity (noise), the coupling of growth and energy usage to expression, and spatiotemporal considerations. We mainly compare progress in bacterial and mammalian systems, which have some of the most-developed engineering frameworks. Overall, one view of synthetic biology can be neatly summarised as "creating in order to understand."

## Introduction

The application of synthetic biology, via the study of engineered genetic circuits and networks, has led to many important insights into native gene network functions, gene expression (including transcriptional and translational noise), cell physiology (including growth and metabolism), and the elucidation of basic biological design principles that are analogous to fundamental laws in physics. The influence of engineering principles in biology, and recent advances in DNA synthesis and sequencing technologies, has additionally led to the construction of predictive mathematical models based on experimental data, which have, in turn, guided the design of more robust, predictable, and even entirely novel networks.

In this review, rather than focusing on the technological and engineering advances in synthetic biology, we will concentrate on seeing how re-engineering biology has given us new understanding. Synthetic biology has been instrumental for gaining insights into inter- and intra-cellular behaviour, from accurate temporal quantification of transcription and translation, controlled by defined stimuli, to rewiring or *de novo* reconstruction of native circuits, to directed evolution of synthetic networks and the manipulation of multicellular interaction, via modulation of multi-circuit dynamics.

Many synthetic biologists have been inspired by a statement from Richard Feynman, found after his death and written in chalk on his blackboard: "*What I cannot create, I do not understand.*" This has served both as a rallying call for

synthetic biology and as a central unifying feature for a field that encompasses diverse disciplines. Here, we will explore bacterial and mammalian systems in particular, with a view to seeing the limits of what we can engineer and understand, particularly for synthesising predictable spatiotemporal gene expression.

## 1. Bacterial gene circuits: from synthesis to insight

Some of the first developments in synthetic biology included the generation of simple synthetic gene expression circuits, such as the repressilator and the toggle switch (1, 2). These networks contained sets of inducible promoters and transcription repressors, with fluorescent markers serving as quantitative outputs, and were developed based on electric circuit design rules. Primarily, they served as a proof of principle that non-native genetic circuits can be expressed and stably maintained in bacteria. In a sense, the field has expanded from simple models to more complex ones (**Figure 1**). Nonetheless, the concepts of using controllable transcriptional and translational regulation, as well as employing quantifiable readouts for gene expression, still form the building blocks of many complex synthetic networks today (3). The development of such circuits has been reviewed extensively elsewhere (4, 5). Here, we will rather focus on exploring two questions:

- What are the biological insights that synthetic approaches have achieved?
- How do they help us to understand underlying cellular processes?

### 1.1 Lessons from bacteria I: the role of noise

Initially considered a major roadblock to overcome, stochastic gene expression, or ‘noise’, has arguably become one of the main biological insights contributed by the field of synthetic biology, transforming our understanding of how, when, and why specific genes are expressed.

Gene expression in isogenic bacterial populations varies from cell-to-cell, even when the population reaches a stable average expression level or steady state. This is due to variations in the number of 'hardware' units, such as transcriptional-translational machinery and regulatory molecules (resulting in extrinsic noise), as well as the inherent stochasticity attributed to the random nature of single-molecule kinetics (resulting in intrinsic noise). These two measures of stochasticity contribute greatly to the phenotypic heterogeneity documented in genetically identical populations (6, 7). Moreover, fluctuations resulting from noise seem to be an intrinsic property of gene expression, as seen in artificial cells composed of cell membrane-mimetic vesicles, containing only transcription and translation machinery (8).

Noise-dependence also plays a prominent role during dynamic cell responses to changing environmental conditions; this involves pulses in activation and deactivation of regulatory factors, in a stochastic manner within a population (9). Occurrences of such pulse regulation include the control of the energy stress response, entry into competence, and the sporulation of *Bacillus subtilis* (10-12). For example, activation of the energy stress response involves

frequency modulation (FM) pulsing of the alternative sigma factor $\delta^B$, with increasing stress positively correlating with increased pulse frequency (10). Synthetic rewiring of the native circuit involved in this response revealed that noise activates a phosphorylation pulse switch that can be propagated, or turned off, by a combination of a mixed positive and delayed negative transcriptional feedback loop, thus allowing for amplification and subsequent pulse termination (10).

In another example of pulse regulation, rewiring the ComS-ComK competence-control circuit of *B. subtilis* demonstrated that stochastic gene expression plays multiple roles leading up to, and during, differentiation to this transient state. With fluctuation in ComK expression controlling the initiation of competence, noise plays a central role in driving its expression and accumulation. Additionally, noise provides a wider range of response duration, which is controlled by the degradation of ComS; as ultra low levels of ComS are necessary for exit from competence, the system becomes more sensitive to stochastic fluctuations of this regulator (11).

A different type of noise-related pulsing is a phenomenon encountered during transcriptional 'bursting', which describes stochasticity involved in global gene expression, particularly in highly-expressed bacterial genes. The bursting pattern of transcription was identified in various synthetic circuits, however, its molecular mechanism has only recently been characterized: it results from the reversible binding and dissociation of gyrase to resolve the buildup of highly supercoiled DNA (6, 13, 14). The latter system serves as an excellent example of

how cellular behaviours, such as stochastic gene expression, are brought into focus via synthetic biology approaches and can subsequently lead to the identification of novel mechanistic insight.

While noise appears to be a global variable in gene expression, it does not affect all genes equally. In *E. coli*, highly-expressed genes have been shown to contain sequences associated with higher extrinsic noise levels (15), while promoters associated with functionally important genes were characterized by decreased noise (16). Additionally, a recent analysis of a synthetic promoter library evolved *de novo* revealed that, while initial promoter noise levels are generally low, selective pressure placed on these regulatory sequences routinely results in increased expression noise. Promoter-associated noise is further propagated by noise associated with regulators, and may act as the primary step to the evolution of more finely tuned regulation (17).

Overall, 'noise' contributes to many levels of gene and network regulation, and has been studied through elegant synthetic circuit design. Thus, we are finally beginning to unravel and control the molecular mechanisms responsible for this phenomenon.

### 1.2 Lessons from bacteria II: growth laws and metabolic cost

An equally important insight contribution from the field of synthetic biology is the concept of cellular energy allocation. While it may now seem intuitive, the quantitative co-dependent relationships of bacterial gene expression, metabolism and growth have only recently begun to emerge. The introduction of synthetic

tools, such as inducible promoters and tunable circuits, has allowed for the quantification of the cellular 'cost' of producing various component parts, thus defining a relationship between expression and growth (18, 19).

One phenomenological theory of bacterial growth suggests that the cellular proteome may be subdivided into one unchangeable core, and two adjustable components that contribute to cell growth: ribosomal proteins and proteins responsible for the influx of nutrients. These components are in constant flux with respect to each other. The insertion of synthetic circuitry, and the addition of extraneous protein expression, is thought to appropriate ribosomal resources from the adjustable proteome fraction and, therefore, to exert a negative effect on the growth rate (20, 21). The cellular contribution to gene expression is therefore not limited to regulatory mechanisms and noise, but also to a global effect on expression based on the physiological constraints of the system (22, 23).

The quantitative nature of this kind of work has contributed to the concept of introducing bacterial growth laws, by applying analogies from Kirchoff's and Ohm's laws to predict the relationship of gene expression and growth rate (24). However, while differences in growth rate and ribosome availability of *E. coli* strains are certainly factors of gene expression, strain-specific nitrogen and carbon metabolism capacity add another layer of complexity in predicting circuit output (25). Additionally, the function of a given protein product can also exert an effect on the growth rate and on other cellular factors. For example, not only can stochastic expression of catabolically active enzymes be propagated and cause

growth fluctuations, noise can also be transmitted to other genes via cell growth (26). Heterogeneity in protein expression, resulting in metabolic stochasticity, can be propagated and can ultimately influence phenotypic heterogeneity (26, 27).

The cellular growth rate can itself affect the dynamics of a gene circuit, and may be responsible for selecting favourable phenotypes, such as the development of bistability for persistence under antibiotic challenge (28, 29). Conversely, synthetic circuits can also affect the growth rate, as they place an energetic burden onto the chassis, or host organism, with greater output generally associated with decreased expression capacity and growth rate (30). However, additional protein cost may be normalized after several generations, as the cell adjusts its ribosomal content to new steady-state levels. (31). Interestingly, construct output is not simply inversely correlated with host capacity, and circuit design plays an important role in maximizing efficiency, with strong RBS-containing circuits being least efficient (32). This is perhaps not surprising as the availability of cellular translation machinery plays a decisive role in cell growth rate (33).

The application of synthetic biology to develop robust growth and expression models, via the generation of quantitative datasets in multiple bacterial species and conditions, promises to transform our predictive understanding of gene expression, with global cell properties being equally important to signalling via transcription factors.

### 1.3 Lessons from bacteria III: multicellular synthetic gene circuits for complex environments

Synthetic biology approaches have been applied to understanding the principles of not only intra-, but also inter-bacterial signalling networks, involving spatial as well as temporal dimensions. Understanding how networks have gained their physiological function, discovering multiple solutions to the generation of identical phenotypes, and analysing the evolutionary decisions taken to arrive at particular natural networks are necessary for the elucidation of common design principles, if they exist, and for the design of novel networks and behaviours.

Multicellular spatiotemporal patterns have been observed during bacterial swarming and chemotaxis, and have been recognised as self-organising behaviours (34, 35). The autonomous nature of the formation of these spatiotemporal patterns was recapitulated using a synthetic circuitry that coupled motility with cell density. Stripes with variable periodicity were obtained by changing expression levels of the motility-associated signal within the circuit (36). Such proof-of-concept reconstruction provides an excellent tool for the analysis of mechanistically complex and uncharacterised natural stripe-forming phenotypes, and may be extrapolated to other periodic population phenomena, such as reaction-diffusion, where spatiotemporal dynamics are controlled by a diffusible signal (36).

Another excellent model for the study of spatial and temporal signalling is the generation of 'stripe' patterns in bacterial populations (37). A recent and powerful approach is to consider all potential network designs that a given circuit may adopt in order to form a stable phenotype. This carries the benefit of discovering multiple, equally attractive solutions, including ones that are not found in native biological systems, giving new mechanistic insight (**Figure 2**)(30, 38). Thus, by adopting a computational approach to generate a genotype-phenotype network map, and by experimentally evaluating prediction results of theoretically-predicted stripe-forming networks, Schaerli et al. probed the entire 'design space' of 3-node gene networks that can form a central stripe in response to a chemical signalling gradient (30). Synthetic biology therefore allows the systematic analysis of a whole network design space and can lead to the identification of common design principles of circuit and network function.

Reconstruction of network dynamics through circuit engineering must also consider the context of the native mechanism. While an alternative synthetic network may simulate the behaviour of a naturally evolved one, it may not be as stable or robust. For example, an engineered circuit with an alternative architecture was compared to the differentiation to competence circuit in *B. subtilis* (39). While the desired phenotype was reconstituted, the circuits differed in noise profiles, which, in turn, resulted in distinct competence dynamics (39). Differences in circuit response may point to selection between high versus low variability in environmental cues, and circuit noise must be considered during network design and phenotype analysis.

The competence circuit has also been used to demonstrate that dynamical behaviour results from the integration of multiple inputs affecting distinct circuit locations, as opposed to the two inputs acting on the same circuit element (40). This brings to light the idea that quantities of environmental cues may also play a role in bacterial circuit dynamics. Exploring transcription network input-output rewiring on a large scale, via the construction of a combinatorial library of promoters and transcription or σ-factor genes, has shown that *E. coli* can tolerate, and indeed take advantage of, extensive rewiring regimes (41). Additionally, the surprising finding that expression of most newly constructed promoter-gene combinations has little effect on growth, is indicative of an underlying mechanism responsible for the buffering of changes in host genetics (41).

Combining systems and synthetic biology approaches can lead to the determination of specific rules for network stability within a complex cellular environment. The randomisation of entire circuits followed by evolutionary selection has shown that fitness decreases with the increasing number of added genes, high expression levels and repeated sequences (42). This randomisation and selection approach was superior to rational circuit design by yielding a robust circuit with low gene expression and no repeated sequences. Moreover, an expression “threshold maximum” was established for ensuring increased evolutionary stability of the circuit (42). Overall, it is clear that synthetic biology is changing our view of the control of gene expression in bacteria, by showing that transcription factor network topologies alone are insufficient to give us a

predictive understanding of cell phenotypes. As we will see in the next section, similar conclusions are likely to hold for eukaryotic cells.

## 2. Rise of the mammals

### 2.1 Sub-cellular organisation and stability of mammalian synthetic circuits

There is great potential in mammalian systems not only to engineer useful functions or devices, but also to gain insights into the mechanisms behind native cell systems. Whereas bacterial synthetic biologists have worked hard to engineer and standardize modularity (43) as well as component orthogonality (44-46) in gene circuits, so far there have not been similar concerted attempts to standardize parts for mammalian systems on a large scale. Nonetheless, mammalian synthetic biology is progressing rapidly (reviewed in (5, 47)), benefiting directly from the large amount of pioneering gene circuit engineering in bacteria (reviewed in (4)). For example, Francois Jacob and Jacques Monod's study of the lactose operon in *E. coli* (48) enabled the first synthetic inducible promoters in mammalian cells (49-51). Since then, toggle switches (52), Boolean logic gates (53), hysteretic switches (54); oscillators (55-57), and even electrically-induced transcriptional switches (58) or light-induced opto-genetic switches (59-62) have been successfully developed in mammalian cells (reviewed in (47, 63, 64)). Despite this progress, mammalian synthetic circuits can be less reliable, more context sensitive, and necessitate longer design-testing iteration cycles than their bacterial counterparts. At the most extreme, this

implies making stable cell lines to test each new gene circuit reliably in a fixed genomic context (e.g. (65)).

A key consideration is the subcellular compartmentalisation of eukaryotic cells, which creates its own engineering demands. For example, eukaryotic cells have highly compartmentalised nuclei, which tightly pack their genomes into chromatin and chromosome territories. Thus, nuclei are functionally organised into domains such as the nuclear lamina, nucleoli, speckles, histone locus bodies (HLBs), PML bodies and Cajal bodies (CBs) (66-68). Unpredictable long-range genome interactions (revealed by HiC studies), occurring at the border of chromosome territories, are responsible for “spatial effect variegation” and stochasticity observed in the mammalian transcriptome (69). The resulting strong expression modulation provided by the nuclear locus makes it difficult to 'plug-and-play' gene expression functions, such as inducible promoters for synthetic circuits.

Moreover, mammalian nuclei are highly dynamic with nuclear components constantly diffusing passively through the nucleoplasm (70) and architectural disruption upon every cellular division (71, 72). In the past, our understanding of the tight spatiotemporal regulation of gene expression has relied largely on the lac operator (LacO) / lac repressor (LacI) system, which allowed the tethering of genes to specific sub-nuclear domains and/or the visualisation of gene movement using GFP-tagged LacI (73). These studies have highlighted that the location of genes and their chromatin states are strongly connected (74), and that the mammalian genome is regulated in 3D (75). Generally, silenced

heterochromatin is found at the nuclear and nucleolar periphery, while active euchromatin is more central.

Synthetic biology is now allowing us to investigate comprehensively how the critical organisation of mammalian nuclei is established and propagated through open mitosis. Indeed, *de novo* biogenesis of synthetic nuclear domains demonstrates our knowledge of the basic mechanisms involved: tethering of a CB constituent, fused with LacI, to an ectopic array of LacO appears to be sufficient to nucleate ectopic CBs via protein-protein interactions (76). Similarly, using ectopic *Xenopus* enhancer (XEn) from ribosomal genes (rDNA) in human cells resulted only in the *de novo* formation of nucleoli inner centers, i.e., synthetic fibrillar centers (FCs) called pseudo-NORs (**Figure 3 and 4**) (77, 78). XEn arrays constitute strong binding sites for nucleolar HMG box transcription factor UBF (Upstream binding factors). Interestingly, during mitosis, the resulting UBF-loaded chromatin forms a prominent chromosomal feature, i.e. a secondary constriction of under-condensed chromatin, constituting the first step in nucleolar formation (77, 78).

As sites of ribosome biogenesis, nucleoli are functionally structured around arrays of ribosomal genes (rDNA) into three distinct compartments: the fibrillar centers (FCs), the dense fibrillar components (DFCs) and the granular component (GC) (**Figure 3**) (79, 80). While the role of FCs in ribosome biogenesis remained uncertain for organisms having DFC/GC bipartite nucleoli (81), synthetic pseudo-NORs now suggest that FCs are essential for the cell cycle propagation of mammalian nucleoli through “open” mitosis (**Figure 4**) (66,

78). They contain unengaged transcription factors and are devoid of processing factors. Transcription of rDNA only occurs at the interface between FCs and DFCs, producing pre-rRNA into the DFCs where early processing takes place. Subsequent late processing of pre-rRNA and assembly of ribosome subunits occurs within the GC. In humans, rDNA arrays are found within the short arms of the five acrocentric chromosomes and nucleoli can derive from more than one acrocentric chromosome (**Figure 4**). Thus, nucleoli are both dynamically and functionally compartmentalized and at the convergence of multiple chromosomes. As such, they are the ultimate paradigm of nuclear functional organization (80). Therefore it was reasonable to ask the question: do we understand nucleoli well enough to engineer them?

Pseudo-NORs reproduced the first step in the formation of artificial nucleoli and their mitotic memory (66, 78). However, the *de novo* construction of fully functional synthetic nucleoli, termed neonucleoli, required not only UBF binding sites but also rDNA transcription units assembled in arrays called neo-NORs (78) (**Figure 4**). Their construction revealed that a chromosomal context, ensuring perinucleolar heterochromatin (82), is implicated in maintaining the genomic stability of rDNA arrays and the efficiency of their transcription (78). Moreover, pseudo-NORs, neo-NORs and neonucleoli provide compelling evidence that nucleolar biogenesis and propagation through 'open' mitosis is a staged process, where UBF-dependent mitotic bookmarking precedes transcript-dependent nucleolar assembly (**Figure 4C**).

Similarly, the generation of ectopic HLBs, speckles, paraspeckles and nuclear stress bodies absolutely requires the synthesis or tethering of an RNA constituent (83-85). Thus, *de novo* construction of nuclear bodies has highlighted the key role of architectural RNAs in nuclear "self-organization". The memory of such organization and its rapid re-establishment after cell division could be driven by mitotic bookmarking (66). Such epigenetic marks retain the memory of active genes through mitosis to ensure an early G1 reactivation of essential genes and/or the maintenance of cell lineage transcriptomes and phenotypes (86).

The development of synthetic biology has also enabled a more systematic study of epigenetic memory and spatial regulation of gene expression by nurturing the development of designed DNA binding factors such as zinc fingers, TALEs and CRISPR/Cas9-derived systems, and their application in editing the epigenome (**Figure 5**) (87-90).

The rise of mammalian synthetic biology has highlighted the crucial role of spatial and epigenetic regulation of gene circuits for their stability. One strategy toward more network reliability is to target the genomic integration of synthetic circuits to genome safe harbours to minimise undesired interference (91). For ultimate reliability and predictability, synthetic circuits may even require the construction of well-characterised mammalian artificial chromosomes (MACs). The first MACs were formed in cells transfected with a DNA cocktail including a centromeric α-satellite array, genomic DNA, and telomeric sequences (92). However, MACs were only observed after uncontrolled rearrangement events, resulting in MACs larger than the input DNA (93). Next-generation MACs could

benefit from the inclusion of rDNA chromosomal context, providing a protective heterochromatin shell, since isolation from the surrounding nucleoplasm and interfering activities would improve the efficiency and reliability of mammalian synthetic circuits. Moreover, progress in the development of MACs will certainly continue improving our understanding of kinetochores and centromeres. Indeed, while mammalian systems have been challenging for synthetic biology, its principles and methodology have provided the most powerful means of understanding the basics of mammalian spatiotemporal genomic organization.

With the means of understanding and overcoming challenges like spatiotemporal stability, mammalian synthetic biology is a rapidly expanding field now turning to synthetic signalling pathways and exciting multicellular networks.

### 2.2. Multicellular mammalian synthetic networks and synthetic therapies

In the process of reducing noise and coordinating gene expression at a cell population level, synthetic biologists have turned to engineering synthetic signalling pathways, thus developing mammalian receiver cells (65, 94-96). Such populations of cells, capable of sensing environmental metabolites, constitute powerful tools to better understand the molecular spatiotemporal dynamics of essential biological processes, including embryonic pattern formation, paracrine or autocrine signalling and cellular migration.

Inspired by the extensive synthetic biology work on spatiotemporal patterning in bacteria (e.g. (30, 37, 97), mammalian analogues have appeared. For instance, the classical “French flag” model of stripe formation (98) has been a

common engineering goal in both bacterial (37) and mammalian synthetic biology. In such systems, cells in a field receive 'high', 'middle' or 'low' concentrations of a morphogen diffusion gradient, and express flag-like stripe genes accordingly. Thus, a mammalian concentration detection network has successfully been engineered in response to a tetracycline gradient (99). Furthermore, a toolkit to establish morphogen diffusion gradient in cysts has been developed to extend the study of pattern formation to 3D cultures in collagen (65). Building artificial sender-receiver systems in cells also allows the use of information theory to study communication between cells quantitatively (reviewed in (100)).

Next-generation networks will likely provide sophisticated synthetic therapeutics, with high precision control devices coupling sensing and delivery mechanisms. Already, mammalian synthetic biology has provided circuits oriented toward therapeutics. RNA-based multi-input Boolean logic cell classifiers can specifically recognize and kill cancer-derived HeLa cells, rather than HEK293 and MCF7 cells, according to their miRNA expression patterns (101, 102). Refinement of the cell classifiers using RNA replicons and RNA-binding proteins provides an even more robust and finely tuneable system (102)

Another complex device with therapeutic potential consists of sender and receiver cells, contained in alginate-poly-L-lysine-alginate capsules, circulating as prosthetic delivery units in the mouse bloodstream, and responding to an acetaldehyde signal (103). These units form multicellular networks, which behave as hormone-like information-processing devices: exogenous signals trigger

quorum-sensing cross-talk to finely tune the blood level of a critical protein, the human placental secreted alkaline phosphatase (103).

Mammalian synthetic therapies can also take advantage of lymphocyte T cells, which can be collected and genetically engineered before being injected back into patients. Thus, chimeric antigen receptors (CARs) of T cells have been successfully rerouted to specifically recognize and kill cancer cells (104, 105).

Finally, perhaps the most exciting development in the last decade has been that of CRISPR/Cas systems for genome editing and a wide range of other applications (106). Already, CRISPR systems have been successfully incorporated in the design of mammalian synthetic networks (107, 108). In the short term, these systems are likely to revolutionise *ex vivo* therapies, where cells are removed, modified and reimplanted, such as in pioneering HIV treatment (109). In the longer term, *in vivo* therapies should become well established, as long as specificity and host immunity complexities are overcome (106, 110). Widespread genome editing will transform therapeutics and genetics, likely leading to the greatest insights of all into the relationship between genotype, phenotype, and the functioning of cells.

## Conclusions

Synthetic biology is the most powerful tool to dissect the limits of basic cell biology. In concert with systems biology, artificial systems can be constructed to obtain high quality quantitative data for computational modelling, and to transform

our ability to generate a predictive understanding of complex cellular processes. In this review, we illustrated this crossover between disciplines by taking examples from bacterial synthetic biology. This field is maturing and constructs are increasing in size and complexity, thereby generating new concepts and insights into gene regulation and cellular physiology.

While synthetic biology has been pioneered in bacteria, the past decade has seen mammalian engineering expanding from sub-cellular to multicellular therapy-orientated networks at an incredible pace. Mammalian cells have their own unique structures, compartmentalisations and regulatory mechanisms that are already yielding unique insights of their own. For instance, by building gene expression factories such as nucleoli, we are already gaining valuable insights into how the corresponding natural systems are organised.

Whereas the last decade focused mainly on single cell biology, it is likely that the next decade will bring advances into multi-cell and organ synthetic biology, which will open up new therapeutic possibilities and will give new biological insights along the way. Such complexity is not restricted to multicellular organisms, such as mammals, but can also apply to microbial populations and ecosystems. For example, in the near future, synthetic biology techniques will be applied to engineer phage and bacteria to explore and to modify the gut microbiome (111, 112)

Therapeutic insights, as well as biological ones, are therefore a natural outcome of re-engineering biology. Nowhere is the potential greater than in the

new genome editing technologies, based on CRISPR systems (106), which have the promise to revolutionise gene and cell therapies in all manner of cell types. For the first time, we have at our fingertips a technology that can introduce almost any sequence or mutation into almost any cell type. Thus perhaps the most exciting change for synthetic biology in the years to come is the transition from "creating in order to understand" towards "creating in order to cure."

**Acknowledgements**

Authors were funded by New Investigator Award No. WT102944 from the Wellcome Trust UK (MI).

## References


1. Elowitz MB, Leibler S. A synthetic oscillatory network of transcriptional regulators. Nature. 2000;403(6767):335-8.
2. Gardner TS, Cantor CR, Collins JJ. Construction of a genetic toggle switch in Escherichia coli. Nature. 2000;403(6767):339-42.
3. Moon TS, Lou C, Tamsir A, Stanton BC, Voigt CA. Genetic programs constructed from layered logic gates in single cells. Nature. 2012;491(7423):249-53.
4. Michalodimitrakis K, Isalan M. Engineering prokaryotic gene circuits. FEMS microbiology reviews. 2009;33(1):27-37.
5. Wieland M, Fussenegger M. Engineering molecular circuits using synthetic biology in mammalian cells. Annu Rev Chem Biomol Eng. 2012;3:209-34.
6. Elowitz MB, Levine AJ, Siggia ED, Swain PS. Stochastic gene expression in a single cell. Science. 2002;297(5584):1183-6.
7. Ozbudak EM, Thattai M, Kurtser I, Grossman AD, van Oudenaarden A. Regulation of noise in the expression of a single gene. Nat Genet. 2002;31(1):69-73.
8. Nishimura K, Tsuru S, Suzuki H, Yomo T. Stochasticity in gene expression in a cell-sized compartment. ACS Synth Biol. 2015;4(5):566-76.
9. Levine JH, Lin Y, Elowitz MB. Functional roles of pulsing in genetic circuits. Science. 2013;342(6163):1193-200.
10. Locke JC, Young JW, Fontes M, Hernandez Jimenez MJ, Elowitz MB. Stochastic pulse regulation in bacterial stress response. Science. 2011;334(6054):366-9.
11. Suel GM, Kulkarni RP, Dworkin J, Garcia-Ojalvo J, Elowitz MB. Tunability and noise dependence in differentiation dynamics. Science. 2007;315(5819):1716-9.
12. Levine JH, Fontes ME, Dworkin J, Elowitz MB. Pulsed feedback defers cellular differentiation. PLoS Biol. 2012;10(1):e1001252.
13. Chong S, Chen C, Ge H, Xie XS. Mechanism of transcriptional bursting in bacteria. Cell. 2014;158(2):314-26.
14. Pedraza JM, Paulsson J. Effects of molecular memory and bursting on fluctuations in gene expression. Science. 2008;319(5861):339-43.
15. Taniguchi Y, Choi PJ, Li GW, Chen H, Babu M, Hearn J, et al. Quantifying E. coli proteome and transcriptome with single-molecule sensitivity in single cells. Science. 2010;329(5991):533-8.
16. Silander OK, Nikolic N, Zaslaver A, Bren A, Kikoin I, Alon U, et al. A genome-wide analysis of promoter-mediated phenotypic noise in Escherichia coli. PLoS Genet. 2012;8(1):e1002443.
17. Wolf L, Silander OK, van Nimwegen E. Expression noise facilitates the evolution of gene regulation. Elife. 2015;4.
18. Klumpp S, Hwa T. Bacterial growth: global effects on gene expression, growth feedback and proteome partition. Curr Opin Biotechnol. 2014;28:96-102.
19. Shahrezaei V, Marguerat S. Connecting growth with gene expression: of noise and numbers. Curr Opin Microbiol. 2015;25:127-35.

20. Scott M, Gunderson CW, Mateescu EM, Zhang Z, Hwa T. Interdependence of cell growth and gene expression: origins and consequences. Science. 2010;330(6007):1099-102.
21. Carrera J, Rodrigo G, Singh V, Kirov B, Jaramillo A. Empirical model and in vivo characterization of the bacterial response to synthetic gene expression show that ribosome allocation limits growth rate. Biotechnol J. 2011;6(7):773-83.
22. Berthoumieux S, de Jong H, Baptist G, Pinel C, Ranquet C, Ropers D, et al. Shared control of gene expression in bacteria by transcription factors and global physiology of the cell. Mol Syst Biol. 2013;9:634.
23. Weisse AY, Oyarzun DA, Danos V, Swain PS. Mechanistic links between cellular trade-offs, gene expression, and growth. Proc Natl Acad Sci U S A. 2015;112(9):E1038-47.
24. Scott M, Hwa T. Bacterial growth laws and their applications. Curr Opin Biotechnol. 2011;22(4):559-65.
25. Cardinale S, Joachimiak MP, Arkin AP. Effects of genetic variation on the E. coli host-circuit interface. Cell Rep. 2013;4(2):231-7.
26. Kiviet DJ, Nghe P, Walker N, Boulineau S, Sunderlikova V, Tans SJ. Stochasticity of metabolism and growth at the single-cell level. Nature. 2014;514(7522):376-9.
27. Labhsetwar P, Cole JA, Roberts E, Price ND, Luthey-Schulten ZA. Heterogeneity in protein expression induces metabolic variability in a modeled Escherichia coli population. Proc Natl Acad Sci U S A. 2013;110(34):14006-11.
28. Klumpp S, Zhang Z, Hwa T. Growth rate-dependent global effects on gene expression in bacteria. Cell. 2009;139(7):1366-75.
29. Deris JB, Kim M, Zhang Z, Okano H, Hermsen R, Groisman A, et al. The innate growth bistability and fitness landscapes of antibiotic-resistant bacteria. Science. 2013;342(6162):1237435.
30. Schaerli Y, Munteanu A, Gili M, Cotterell J, Sharpe J, Isalan M. A unified design space of synthetic stripe-forming networks. Nat Commun. 2014;5:4905.
31. Shachrai I, Zaslaver A, Alon U, Dekel E. Cost of unneeded proteins in E. coli is reduced after several generations in exponential growth. Mol Cell. 2010;38(5):758-67.
32. Ceroni F, Algar R, Stan GB, Ellis T. Quantifying cellular capacity identifies gene expression designs with reduced burden. Nature methods. 2015;12(5):415-8.
33. Cardinale S, Arkin AP. Contextualizing context for synthetic biology--identifying causes of failure of synthetic biological systems. Biotechnol J. 2012;7(7):856-66.
34. Woodward DE, Tyson R, Myerscough MR, Murray JD, Budrene EO, Berg HC. Spatio-temporal patterns generated by Salmonella typhimurium. Biophys J. 1995;68(5):2181-9.
35. Matsuyama T, Takagi Y, Nakagawa Y, Itoh H, Wakita J, Matsushita M. Dynamic aspects of the structured cell population in a swarming colony of Proteus mirabilis. J Bacteriol. 2000;182(2):385-93.
36. Liu C, Fu X, Liu L, Ren X, Chau CK, Li S, et al. Sequential establishment of stripe patterns in an expanding cell population. Science. 2011;334(6053):238-41.

37. Basu S, Gerchman Y, Collins CH, Arnold FH, Weiss R. A synthetic multicellular system for programmed pattern formation. Nature. 2005;434(7037):1130-4.
38. Rodrigo G, Carrera J, Jaramillo A. Computational design of synthetic regulatory networks from a genetic library to characterize the designability of dynamical behaviors. Nucleic Acids Res. 2011;39(20):e138.
39. Cagatay T, Turcotte M, Elowitz MB, Garcia-Ojalvo J, Suel GM. Architecture-dependent noise discriminates functionally analogous differentiation circuits. Cell. 2009;139(3):512-22.
40. Espinar L, Dies M, Cagatay T, Suel GM, Garcia-Ojalvo J. Circuit-level input integration in bacterial gene regulation. Proc Natl Acad Sci U S A. 2013;110(17):7091-6.
41. Isalan M, Lemerle C, Michalodimitrakis K, Horn C, Beltrao P, Raineri E, et al. Evolvability and hierarchy in rewired bacterial gene networks. Nature. 2008;452(7189):840-5.
42. Sleight SC, Sauro HM. Visualization of evolutionary stability dynamics and competitive fitness of Escherichia coli engineered with randomized multigene circuits. ACS Synth Biol. 2013;2(9):519-28.
43. Canton B, Labno A, Endy D. Refinement and standardization of synthetic biological parts and devices. Nat Biotechnol. 2008;26(7):787-93.
44. Stanton BC, Nielsen AA, Tamsir A, Clancy K, Peterson T, Voigt CA. Genomic mining of prokaryotic repressors for orthogonal logic gates. Nat Chem Biol. 2014;10(2):99-105.
45. Rhodius VA, Segall-Shapiro TH, Sharon BD, Ghodasara A, Orlova E, Tabakh H, et al. Design of orthogonal genetic switches based on a crosstalk map of sigmas, anti-sigmas, and promoters. Mol Syst Biol. 2013;9:702.
46. Temme K, Hill R, Segall-Shapiro TH, Moser F, Voigt CA. Modular control of multiple pathways using engineered orthogonal T7 polymerases. Nucleic Acids Res. 2012;40(17):8773-81.
47. Greber D, Fussenegger M. Mammalian synthetic biology: engineering of sophisticated gene networks. J Biotechnol. 2007;130(4):329-45.
48. Jacob F, Monod J. Genetic regulatory mechanisms in the synthesis of proteins. J Mol Biol. 1961;3:318-56.
49. Hu MC, Davidson N. The inducible lac operator-repressor system is functional in mammalian cells. Cell. 1987;48(4):555-66.
50. Brown M, Figge J, Hansen U, Wright C, Jeang KT, Khoury G, et al. lac repressor can regulate expression from a hybrid SV40 early promoter containing a lac operator in animal cells. Cell. 1987;49(5):603-12.
51. Wyborski DL, Short JM. Analysis of inducers of the E.coli lac repressor system in mammalian cells and whole animals. Nucleic Acids Res. 1991;19(17):4647-53.
52. Kramer BP, Viretta AU, Daoud-El-Baba M, Aubel D, Weber W, Fussenegger M. An engineered epigenetic transgene switch in mammalian cells. Nat Biotechnol. 2004;22(7):867-70.
53. Kramer BP, Fischer C, Fussenegger M. BioLogic gates enable logical transcription control in mammalian cells. Biotechnol Bioeng. 2004;87(4):478-84.
54. Kramer BP, Fussenegger M. Hysteresis in a synthetic mammalian gene network. Proc Natl Acad Sci U S A. 2005;102(27):9517-22.

55. Weber W, Stelling J, Rimann M, Keller B, Daoud-El Baba M, Weber CC, et al. A synthetic time-delay circuit in mammalian cells and mice. Proc Natl Acad Sci U S A. 2007;104(8):2643-8.
56. Tigges M, Marquez-Lago TT, Stelling J, Fussenegger M. A tunable synthetic mammalian oscillator. Nature. 2009;457(7227):309-12.
57. Tigges M, Denervaud N, Greber D, Stelling J, Fussenegger M. A synthetic low-frequency mammalian oscillator. Nucleic Acids Res. 2010;38(8):2702-11.
58. Weber W, Luzi S, Karlsson M, Sanchez-Bustamante CD, Frey U, Hierlemann A, et al. A synthetic mammalian electro-genetic transcription circuit. Nucleic Acids Res. 2009;37(4):e33.
59. Yazawa M, Sadaghiani AM, Hsueh B, Dolmetsch RE. Induction of protein-protein interactions in live cells using light. Nat Biotechnol. 2009;27(10):941-5.
60. Wang X, Chen X, Yang Y. Spatiotemporal control of gene expression by a light-switchable transgene system. Nature methods. 2012;9(3):266-9.
61. Crefcoeur RP, Yin R, Ulm R, Halazonetis TD. Ultraviolet-B-mediated induction of protein-protein interactions in mammalian cells. Nat Commun. 2013;4:1779.
62. Muller K, Weber W. Optogenetic tools for mammalian systems. Mol Biosyst. 2013;9(4):596-608.
63. Kis Z, Pereira HS, Homma T, Pedrigi RM, Krams R. Mammalian synthetic biology: emerging medical applications. J R Soc Interface. 2015;12(106).
64. Lienert F, Lohmueller JJ, Garg A, Silver PA. Synthetic biology in mammalian cells: next generation research tools and therapeutics. Nat Rev Mol Cell Biol. 2014;15(2):95-107.
65. Carvalho A, Menendez DB, Senthivel VR, Zimmermann T, Diambra L, Isalan M. Genetically encoded sender-receiver system in 3D mammalian cell culture. ACS Synth Biol. 2014;3(5):264-72.
66. Grob A, McStay B. Construction of synthetic nucleoli and what it tells us about propagation of sub-nuclear domains through cell division. Cell Cycle. 2014;13(16):2501-8.
67. Bolzer A, Kreth G, Solovei I, Koehler D, Saracoglu K, Fauth C, et al. Three-dimensional maps of all chromosomes in human male fibroblast nuclei and prometaphase rosettes. PLoS Biol. 2005;3(5):e157.
68. Lieberman-Aiden E, van Berkum NL, Williams L, Imakaev M, Ragoczy T, Telling A, et al. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. Science. 2009;326(5950):289-93.
69. Noordermeer D, de Wit E, Klous P, van de Werken H, Simonis M, Lopez-Jones M, et al. Variegated gene expression caused by cell-specific long-range DNA interactions. Nat Cell Biol. 2011;13(8):944-51.
70. Phair RD, Misteli T. High mobility of proteins in the mammalian cell nucleus. Nature. 2000;404(6778):604-9.
71. Thomson I, Gilchrist S, Bickmore WA, Chubb JR. The radial positioning of chromatin is not inherited through mitosis but is established de novo in early G1. Curr Biol. 2004;14(2):166-72.
72. Dundr M, Misteli T. Biogenesis of nuclear bodies. Cold Spring Harb Perspect Biol. 2010;2(12):1-15.

73. Chubb JR, Boyle S, Perry P, Bickmore WA. Chromatin motion is constrained by association with nuclear compartments in human cells. Curr Biol. 2002;12(6):439-45.
74. Kind J, Pagie L, Ortabozkoyun H, Boyle S, de Vries SS, Janssen H, et al. Single-cell dynamics of genome-nuclear lamina interactions. Cell. 2013;153(1):178-92.
75. Lanctot C, Cheutin T, Cremer M, Cavalli G, Cremer T. Dynamic genome architecture in the nuclear space: regulation of gene expression in three dimensions. Nat Rev Genet. 2007;8(2):104-15.
76. Kaiser TE, Intine RV, Dundr M. De novo formation of a subnuclear body. Science. 2008;322(5908):1713-7.
77. Mais C, Wright JE, Prieto JL, Raggett SL, McStay B. UBF-binding site arrays form pseudo-NORs and sequester the RNA polymerase I transcription machinery. Genes Dev. 2005;19(1):50-64.
78. Grob A, Colleran C, McStay B. Construction of synthetic nucleoli in human cells reveals how a major functional nuclear domain is formed and propagated through cell division. Genes Dev. 2014;28(3):220-30.
79. Raska I, Shaw PJ, Cmarko D. Structure and function of the nucleolus in the spotlight. Curr Opin Cell Biol. 2006;18(3):325-34.
80. Hernandez-Verdun D. The nucleolus: a model for the organization of nuclear functions. Histochem Cell Biol. 2006;126(2):135-48.
81. Thiry M, Lafontaine DL. Birth of a nucleolus: the evolution of nucleolar compartments. Trends Cell Biol. 2005;15(4):194-9.
82. Floutsakou I, Agrawal S, Nguyen TT, Seoighe C, Ganley AR, McStay B. The shared genomic architecture of human nucleolar organizer regions. Genome research. 2013;23(12):2003-12.
83. Shevtsov SP, Dundr M. Nucleation of nuclear bodies by RNA. Nat Cell Biol. 2011;13(2):167-73.
84. Mao YS, Sunwoo H, Zhang B, Spector DL. Direct visualization of the co-transcriptional assembly of a nuclear body by noncoding RNAs. Nat Cell Biol. 2011;13(1):95-101.
85. Salzler HR, Tatomer DC, Malek PY, McDaniel SL, Orlando AN, Marzluff WF, et al. A sequence in the Drosophila H3-H4 Promoter triggers histone locus body assembly and biosynthesis of replication-coupled histone mRNAs. Dev Cell. 2013;24(6):623-34.
86. Sarge KD, Park-Sarge OK. Gene bookmarking: keeping the pages open. Trends Biochem Sci. 2005;30(11):605-10.
87. Keung AJ, Bashor CJ, Kiriakov S, Collins JJ, Khalil AS. Using targeted chromatin regulators to engineer combinatorial and spatial transcriptional regulation. Cell. 2014;158(1):110-20.
88. Keung AJ, Joung JK, Khalil AS, Collins JJ. Chromatin regulation at the frontier of synthetic biology. Nat Rev Genet. 2015;16(3):159-71.
89. Hathaway NA, Bell O, Hodges C, Miller EL, Neel DS, Crabtree GR. Dynamics and memory of heterochromatin in living cells. Cell. 2012;149(7):1447-60.
90. Ragunathan K, Jih G, Moazed D. Epigenetics. Epigenetic inheritance uncoupled from sequence-specific recruitment. Science. 2015;348(6230):1258699.

91. Sadelain M, Papapetrou EP, Bushman FD. Safe harbours for the integration of new DNA in the human genome. Nat Rev Cancer. 2012;12(1):51-8.
92. Harrington JJ, Van Bokkelen G, Mays RW, Gustashaw K, Willard HF. Formation of de novo centromeres and construction of first-generation human artificial microchromosomes. Nat Genet. 1997;15(4):345-55.
93. Kouprina N, Earnshaw WC, Masumoto H, Larionov V. A new generation of human artificial chromosomes for functional genomics and gene therapy. Cell Mol Life Sci. 2013;70(7):1135-48.
94. Weber W, Malphettes L, de Jesus M, Schoenmakers R, El-Baba MD, Spielmann M, et al. Engineered Streptomyces quorum-sensing components enable inducible siRNA-mediated translation control in mammalian cells and adjustable transcription control in mice. J Gene Med. 2005;7(4):518-25.
95. Williams SC, Patterson EK, Carty NL, Griswold JA, Hamood AN, Rumbaugh KP. Pseudomonas aeruginosa autoinducer enters and functions in mammalian cells. J Bacteriol. 2004;186(8):2281-7.
96. Neddermann P, Gargioli C, Muraglia E, Sambucini S, Bonelli F, De Francesco R, et al. A novel, inducible, eukaryotic gene expression system based on the quorum-sensing transcription factor TraR. EMBO Rep. 2003;4(2):159-65.
97. Basu S, Mehreja R, Thiberge S, Chen MT, Weiss R. Spatiotemporal control of gene expression with pulse-generating networks. Proc Natl Acad Sci U S A. 2004;101(17):6355-60.
98. Wolpert L, Gingell D. Cell surface membrane and amoeboid movement. Symp Soc Exp Biol. 1968;22:169-98.
99. Greber D, Fussenegger M. An engineered mammalian band-pass network. Nucleic Acids Res. 2010;38(18):e174.
100. Barcena Menendez D, Senthivel VR, Isalan M. Sender-receiver systems and applying information theory for quantitative synthetic biology. Curr Opin Biotechnol. 2015;31:101-7.
101. Xie Z, Wroblewska L, Prochazka L, Weiss R, Benenson Y. Multi-input RNAi-based logic circuit for identification of specific cancer cells. Science. 2011;333(6047):1307-11.
102. Wroblewska L, Kitada T, Endo K, Siciliano V, Stillo B, Saito H, et al. Mammalian synthetic circuits with RNA binding proteins for RNA-only delivery. Nat Biotechnol. 2015;33(8):839-41.
103. Weber W, Daoud-El Baba M, Fussenegger M. Synthetic ecosystems based on airborne inter- and intrakingdom communication. Proc Natl Acad Sci U S A. 2007;104(25):10435-40.
104. Ye H, Aubel D, Fussenegger M. Synthetic mammalian gene circuits for biomedical applications. Curr Opin Chem Biol. 2013;17(6):910-7.
105. Grupp SA, Kalos M, Barrett D, Aplenc R, Porter DL, Rheingold SR, et al. Chimeric antigen receptor-modified T cells for acute lymphoid leukemia. N Engl J Med. 2013;368(16):1509-18.
106. Mali P, Esvelt KM, Church GM. Cas9 as a versatile tool for engineering biology. Nature methods. 2013;10(10):957-63.
107. Chavez A, Scheiman J, Vora S, Pruitt BW, Tuttle M, E PRI, et al. Highly efficient Cas9-mediated transcriptional programming. Nature methods. 2015;12(4):326-8.

108. Zalatan JG, Lee ME, Almeida R, Gilbert LA, Whitehead EH, La Russa M, et al. Engineering complex synthetic transcriptional programs with CRISPR RNA scaffolds. Cell. 2015;160(1-2):339-50.
109. Tebas P, Stein D, Tang WW, Frank I, Wang SQ, Lee G, et al. Gene editing of CCR5 in autologous CD4 T cells of persons infected with HIV. N Engl J Med. 2014;370(10):901-10.
110. Agustin-Pavon C, Isalan M. Synthetic biology and therapeutic strategies for the degenerating brain: Synthetic biology approaches can transform classical cell and gene therapies, to provide new cures for neurodegenerative diseases. Bioessays. 2014;36(10):979-90.
111. Mimee M, Tucker AT, Voigt CA, Lu TK. Programming a human commensal bacterium, *Bacteroides thetaiotaomicron*, to sense and respond to stimuli in the murine gut microbiota. Cell Systems. 2015;1(1):21-3.
112. Ando H, Lemire S, Pires DP, Lu TK. Engineering modular viral scaffoelds for targeted bacterial population editing. Cell Systems. 2015;1(3):187-96.

## Figure legends

**Figure 1.** Synthetic and natural bacterial gene circuits.

(A) One of the first synthetic circuits, the repressilator, is composed of three repressors that are arranged in a feedback loop, and results in oscillatory behaviour. (B) An improved synthetic oscillator, the relaxation oscillator, including both positive and negative feedback loops, results in stable and robust oscillations. (C) *Bacillus subtilis* energy stress response frequency modulation (FM) pulse control. A noise-activated rise in the fluctuation of the phosphatase RsbQP levels results in de-phosphorylation of RsbV, activating the alternative sigma factor $\sigma^B$, leading to upregulation of expression of the operon genes and resulting in operon feedback. When RsbW kinase activity overcomes phosphatase activity, the pulse is terminated. Increased stress results in higher phosphatase levels and results in more frequent pulsing. (D) Transient differentiation to competence is controlled by noise-related fluctuation of the master regulator ComK in *Bacillus subtilis*. Expression noise controls competence initiation, with a rise in the fluctuating levels of ComK resulting in positive autoregulation of transcription and negative feedback through control of the ComK inhibitor ComS. (E) Multicellular patterning: a bacterial sender-receiver system generating two-dimensional pattern formation. The circuit contains sender (yellow) and receiver (green and red) cells. Adapted from the *Vibrio fischieri* motility network, sender cells secrete the enzyme LuxI, that generates the messenger acyl homoserine lactone (AHL). AHL diffuses through the medium away from the sender population (yellow triangle), is internalized, and binds to

LuxR in receiver cells, thus activating the expression of cI and $LacI_{MI}$ in a circuit engineered to act as a band detection filter (band detection high and low thresholds are indicated by blue stars). Two populations of receiver cells are plated together, expressing green or red fluorescent protein expression (GFP or RFP). Repression of GFP or RFP results from differences in LacI stability, which controls the distance of the ring from the sender population. (F) Multicellular patterning: a dual-module density-sensing and motility-directing circuit generates radiating ring patterns away from the centrally placed AHL signal. AHL is generated via expression of LuxI. LuxR, activated by AHL, turns on the expression of cI, which represses the motility regulator cheZ.

**Figure 2.** Exploring the design space of possible dynamical mechanisms of stripe-forming 3-node gene regulatory networks (GRNs).

GRN complexity decreases vertically, with all combinations derived computationally and thus converged onto "stalactites" representing the simplest minimal networks that can execute a function. The design space of functional stripe-forming networks divides into four minimal networks of incoherent feed-forward (IFF) loops (third tier, I-1-4), which converge on a two-node archetype I-zero. Nodes I3 and I4 are novel arrangements previously unreported both in nature and in synthetic biology studies (30).

**Figure 3.** Functional compartmentalization of human nucleoli

Functional nucleoli are tripartite. Their inner FCs (green box) contain rDNA and RNA polymerase I (pol I) machinery unengaged in transcription. Pol I machinery (UBF, SL1, Rrn3 and pol I) is dedicated to rDNA transcription and is active at the interface between FCs and DFCs (white box). Primary transcripts, 47S pre-rRNAs, are produced into the DFCs. Processing of these transcripts into mature 18S, 5.8S and 28S rRNAs requires a succession of coordinated cleavages (green arrows) and modifications. While early steps of this process occur within the DFCs (yellow box), the later steps and assembly of the ribosomal subunits 40S and 60S occur within the GC (red box). FCs are at the core of nucleoli, surrounded by DFCs and further embedded into the GC.

**Figure 4.** Lessons from synthetic neonucleoli

(A) Nucleoli are formed around nucleolar organizer regions or NORs, located, in humans, on the short arms of acrocentric chromosomes. NORs contain rDNA arrays in a specific conserved chromosomal context formed of the telomeric Distal Junction (DJ) and centromeric Proximal Junction (PJ). rDNA are constituted of an intergenic spacer (IGS) and a transcription unit.

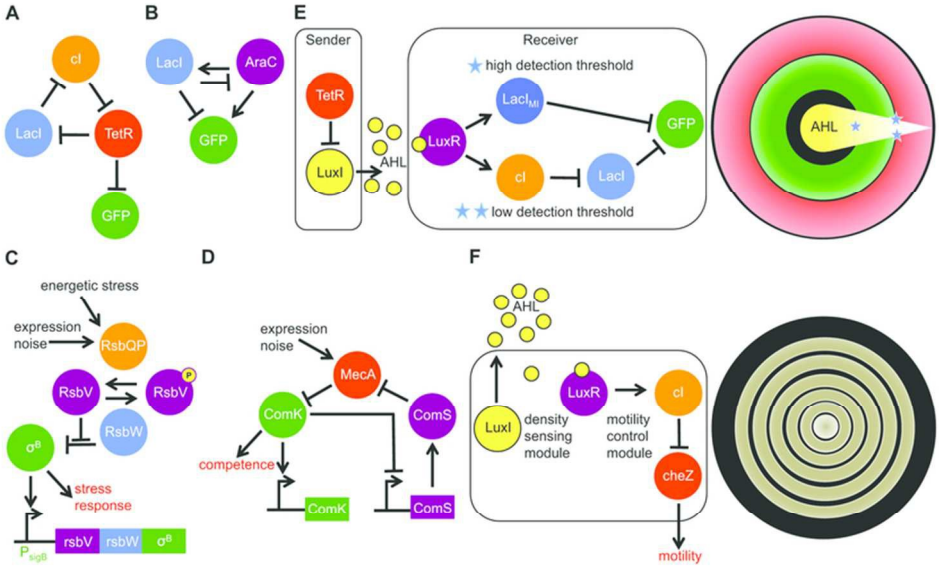
(B) Synthetic neo-NORs revealed the existence of NOR territories within nucleoli, reminiscent of the chromosome territories (white dashed outline). These NOR territories are overlay with functional compartmentalization into FC, DFC and GC; in a similar way that chromosome territories are overlay with nuclear bodies, such as nucleoli or HLBs. In humans, nucleoli are typically surrounded by perinucleolar

heterochromatin (dark blue) derived from the DJs (white spots). While nucleoli disassemble during mitosis, NORs competent in nucleolar formation remain undercondensed and retain components of the pol I machinery. This results in a secondary (2°) constriction, a prominent feature of metaphase acrocentric chromosomes. Synthetic pseudo-NORs, arrays of XEn UBF binding sites (green boxes), resulted in the formation of novel mitotic 2° constrictions and interphase bodies reminiscent of nucleolar FCs, arguing that 2° constrictions are the mitotic counterpart of interphase FCs. Synthetic neo-NORs, formed of XEn arrays interspersed with rDNA transcription units, drove the formation of mitotic 2° constrictions and fully functional compartmentalized neonucleoli. Upon UBF depletion, synthetic pseudo-NORs and neonucleoli are lost.

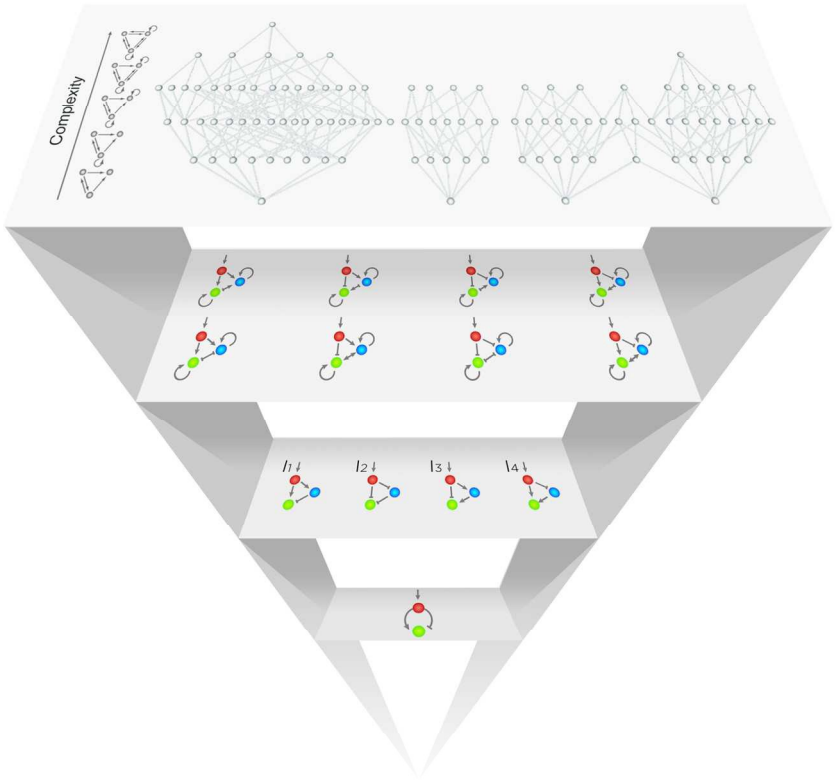
(C) During “open” mitosis, even though pol I transcription is repressed, nucleolar FC components (green spots on the blue metaphase plate) remain associated with competent NORs, while DFC and GC components (red) relocate at the chromosomes' periphery. At the exit from mitosis, rDNA transcription resumes and DFC/GC processing factors are recruited to the newly synthesized pre-rRNAs, thus reforming functional nucleoli. Construction of pseudo-NORs, neo-NORs and neonucleoli revealed that UBF induces the formation of a specific chromatin state ensuring a mitotic bookmarking that is essential but not sufficient for nucleolar formation. Thus, cell cycle inheritance and biogenesis of nucleoli is a staged process, where UBF-dependent bookmarking of competent NORs ensure an early resumption of rDNA transcription and pre-rRNA-dependent reassembly of nucleoli.

**Figure 5.** Epigenome editing

Introduction of ectopic arrays of UBF binding sites (green) ensures the recruitment of UBF, resulting in epigenetic bookmarking. Ectopic binding sites of other natural DNA-binding factors (light blue), like LacI or TetR, can also be used for epigenome editing following the fusion of these factors with a chromatin-modifying domain (grey). Fusion with a zinc finger protein (yellow), a TALE protein (dark blue) or a nuclease-deficient dCas9 protein (red) can also be used to target chromatin-modifying factors to a specific locus. These synthetic DNA-targeted factors can either activate gene expression (e.g. VP16) or repress it (e.g. KRAB).

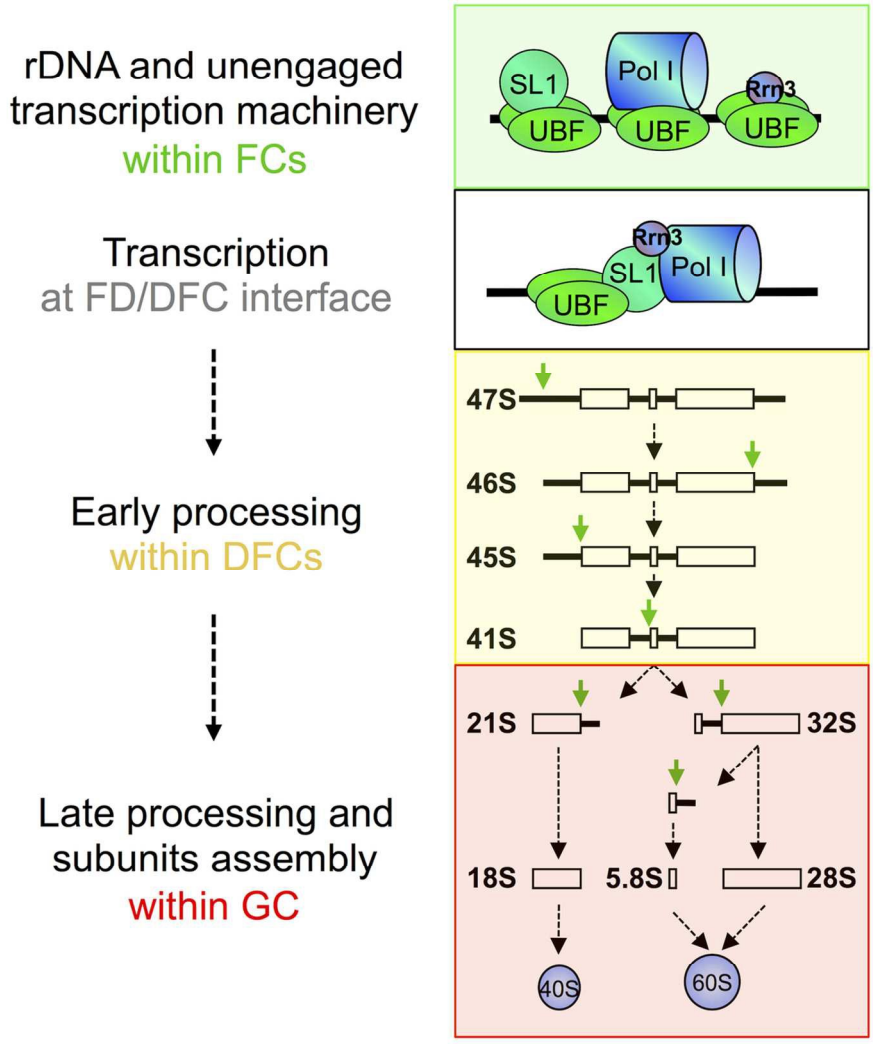



75x56mm (300 x 300 DPI)

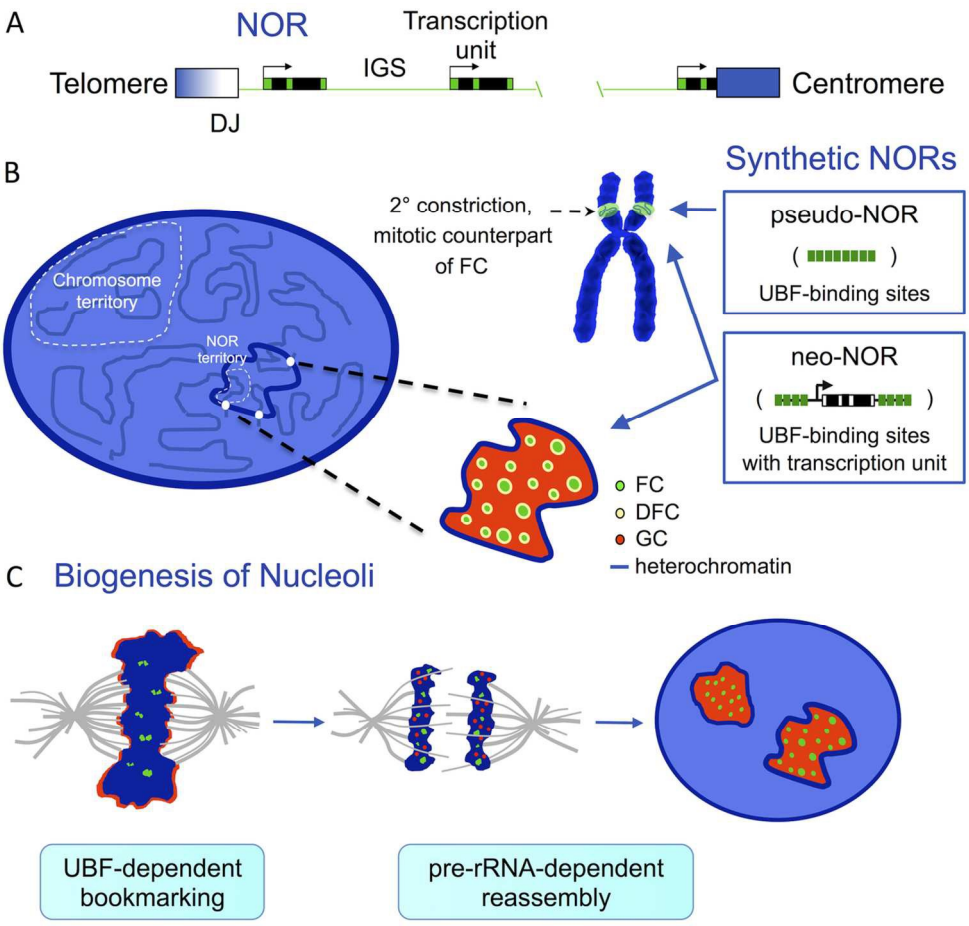



152x143mm (300 x 300 DPI)

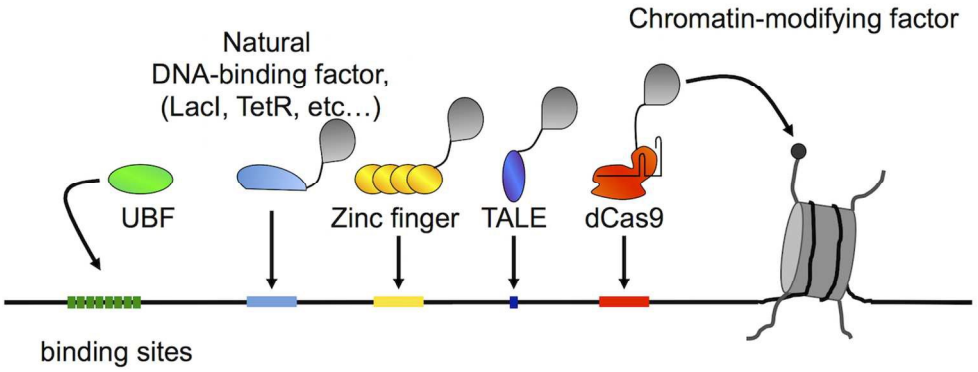
# Functional compartmentalization of nucleoli



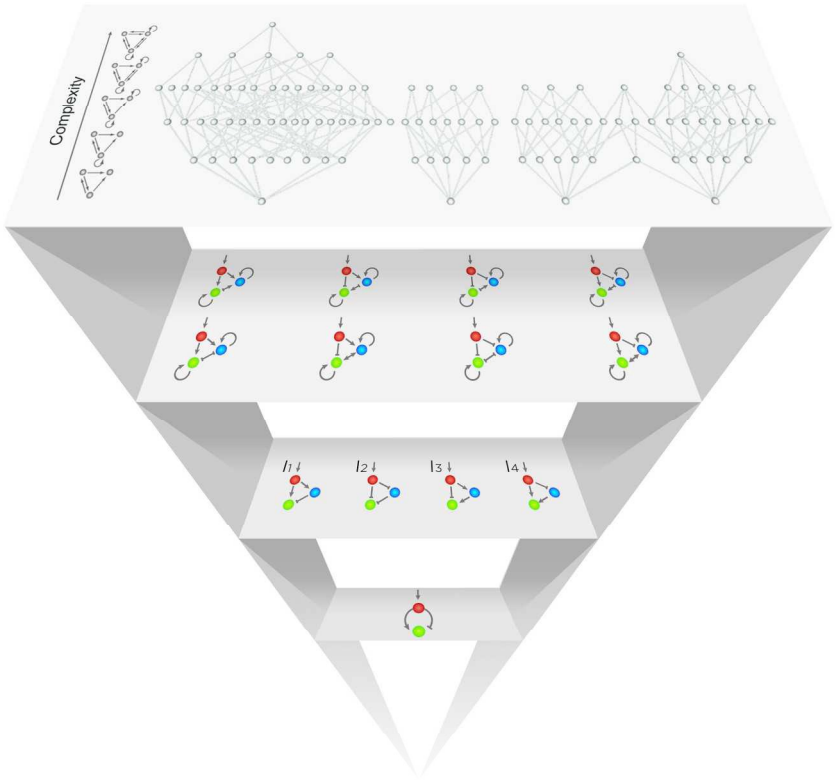

118x139mm (300 x 300 DPI)




119x111mm (300 x 300 DPI)




119x46mm (300 x 300 DPI)




215x203mm (300 x 300 DPI)