# On Programming Bacteria and the Rise of Synthetic Biology

Since the invention of the battery over 200 years ago, an exponentially increasing number of electronic circuits — trillions upon trillions — have quietly become interwoven into society as integral passageways of information transfer, many of which now tend to hover only millimeters from our bodies for the majority of our waking lives, and could taken individually or collectively as a multi-scaled supra-biological nervous system of electromagnetic data transfer. Yet for several billion years, DNA has been honing its own types of circuits upon which the biological nervous systems responsible for this digital influx were steadily assembled. The earliest known vessels used by DNA for assembling and propagating biological circuits at a level of complexity sufficient enough to be considered "life" are believed to be ancient "extremophile" microorganisms that lived near hydrothermal vents, similar to those of today. The Last Universal Common Ancestor, or LUCA, comprises a small population of organisms that via fossilized evidence and comparative genomics are believed to be, as of now, the most recent common ancestors of life on Earth. The LUCA are currently thought to have resembled modern-day bacteria and, like those of today, possessed DNA that looped into a circle. Due to the staggering complexity of biology, it may come as little surprise, then, that as human knowledge of circuit design has begun to be spilled into the processes of genes and of life, the circular genomes of bacteria, some of the simplest and most ancient forms of life, have been at the forefront as our living circuit boards. Bootstrapped from bacteria thus emerges the relatively new discipline of Synthetic Biology, which involves the precise construction and implementation of biological circuits for applications spanning medicine to biofuels to sheer entertainment.

The most obvious manner in which circuit design in Synthetic Biology differs from other engineering fields lies in the circuits themselves. Traditional electronic circuits, stated extremely briefly, pull a voltage between opposite regions of geometrically linear lanes of structurally static, doped silicon atoms comprising transistors, providing atomic "holes" that offer up enough space to slingshot electrons through loosely-bound atomic shells from one side of the circuit to the other, which can be monitored in the ON-OFF fashions of boolean logic. Biological circuits also involve the input, bonding or export of electrons on, in and off atomic shells, but the structural platforms upon which these atoms work are entirely different. The atoms of biological circuits are bundled up by the tens of thousands or more into molecules of, most often, DNA, RNA and proteins, and all three cases quickly exit the realm of linear operation. There is also very little directional operation, as there is no up, down, left or right inside a cell, and molecules tend to only localize when necessary or "called upon" to do so. Because the omnidirectional, constant movement of not only electrons, but also of the structures shuttling them between their atomic-shell destinations inside a single bacterium is exponentially more complex than a human-designed circuit board, a great deal is left to be spatiotemporally understood about these processes, and this is applicable not only to general populations but to scientific communities as well.

In Harvard professor George Church's sweeping synthetic biology tome *Regenesis*, the words "DNA", "genome", "replication", and "E. Coli", taken in combination, are mentioned over 800 times, and for good reason. E. Coli are well studied, easy to culture and duplicate very quickly, making them a model organism for understanding and prototyping biological circuits. Despite this, the book still finds itself run-through with textbook descriptions of monomer assembly into polymers inside a "dense three-dimensional [cytoplasmic] throbbing blob" that "careens relentlessly toward self-replication" via the evolutionary buildup of "replicated complexity". Therefore, a description that does justice to the E. Coli genome and what its DNA is doing (or of DNA in general) is never provided, even though the DNA inside E. Coli tends to be considered the molecular foundation upon which the entire discipline of synthetic biology was built.

The DNA inside an E. Coli bacterium might resemble something like an unwoven but highly tangled ball of molecular yarn composed of two strings woven together in a helical form. The long, double-stranded thread is connected at both ends to create a closed loop. If the entire length of this tangled molecular yarn was taken to be the full E. Coli genome, then taking each of its approximately 4,000-5,000 of genes into account would involve functionally dividing the thread into tiny pieces, each only about 0.02% of the length of the whole. These pieces — genes — are typically not separated from the primary structure (with some exceptions) and everywhere around this long, circular, tangled molecular thread, tiny little segments (genes) are constantly being opened and closed (zipped and unzipped) in response to external signals. Each time a tiny segment of the molecular "yarn"/genome is unzipped, a structural representation of this tiny segment (mRNA) is threaded together and released by a separate molecule (mRNA polymerase), and then downstream a second structural representation (a protein) of the first released tiny segment (mRNA) is threaded together by a massive ribosome and subsequently self-assembles into a 3D functional unit, respectively representing transcription (DNA-mRNA) and translation (mRNA to proteins). mRNA polymerase molecules and ribosomes share the commonality of each acting as electrochemical siphons, respectively sucking in nearby nucleic acids to string together mRNA strands or amino acids to string together proteins.

The capability of a long thread of the DNA molecule to constantly release tiny, fragmented representations of its geometry (0.02% here, 0.015% there, 0.02% there, many times and in many locations at once) as packaged responses to environmental stimuli makes the molecule so astounding. Its structure essentially walks through time by flipping out and surrounding itself with its own structural fragments (genes bundled up into proteins). With the help of the lipid membrane that forms the overall structural boundary of an E. Coli, its DNA is able to cocoon itself in a complex interlocking web of its own functional components. Complexity is increased further as many functions inside an E. Coli are carried out by molecules of which are not proteins but are produced by the proteins themselves— such as the signaling molecule AHL. In other words, many molecules performing important functions for the bacteria are not fragmented structural descendants of DNA itself, but are produced by said fragments. DNA is therefore capable of not only task distribution via acting as the substrate for catalyzing the release of functionally-partitioned segments of its structure, but can further use these segments to externalize the juggling of additional structures/functions.

Even stranger, taken as a whole, if the entirety of this tangled, molecular/DNA-yarn ball and all of its ceaseless unzipping and zipping were observed as-is with no other surrounding molecules — e.g. if every part of an E. Coli was stripped away except its DNA, and its DNA was observed doing what it does naturally through time, nearly all locations throughout the thread would appear to be constantly riddled with the appearance and disappearance of tiny, dynamic holes (created by polymerase and helicase molecules) that open up one part of the thread, travel slightly and close, essentially appearing as an orchestra of thousands of small structural ripples of unzipping and zipping. Taken collectively, the appearance of this molecular orchestra might be imagined to fire in a form of constantly reorienting, rhythmic chaos, almost like the light shows of bioluminescent jellyfish or the swarming behavior of starling [bird] flocks, mimicking how DNA is locked in a continuum of omnidirectional environmental stimuli and responds accordingly with functionally-partitioned structural fragments. Furthermore, approximately every 20-30 minutes, one of the same tiny holes would appear in the circular DNA thread but then travel outwards in both directions instead of only one, and both of these holes would continue moving forward (this time *not* closing up / zipping up behind them) until they had traveled the entire length of the DNA thread/circle. In this case, it is not a small structural/functional fragment (mRNA) that is being released, but it is a copy of the DNA itself that is let loose — each of the two single, unwoven strings ends up with a new DNA strand threaded around it in the process of DNA replication. The myriad tiny structural "micro-ripples" of constant gene expression are therefore occasionally interrupted every half hour or so by one large "macro-ripple" of DNA replication.

The structural ripples to DNA caused by gene expression (micro) and DNA replication (macro) have an inherent, *tunable* waveform that be be visualized by DNA or RNA polymerase binding [“excitation"] and unbinding [“relaxation”], abstracted in Figures 1-3 below.

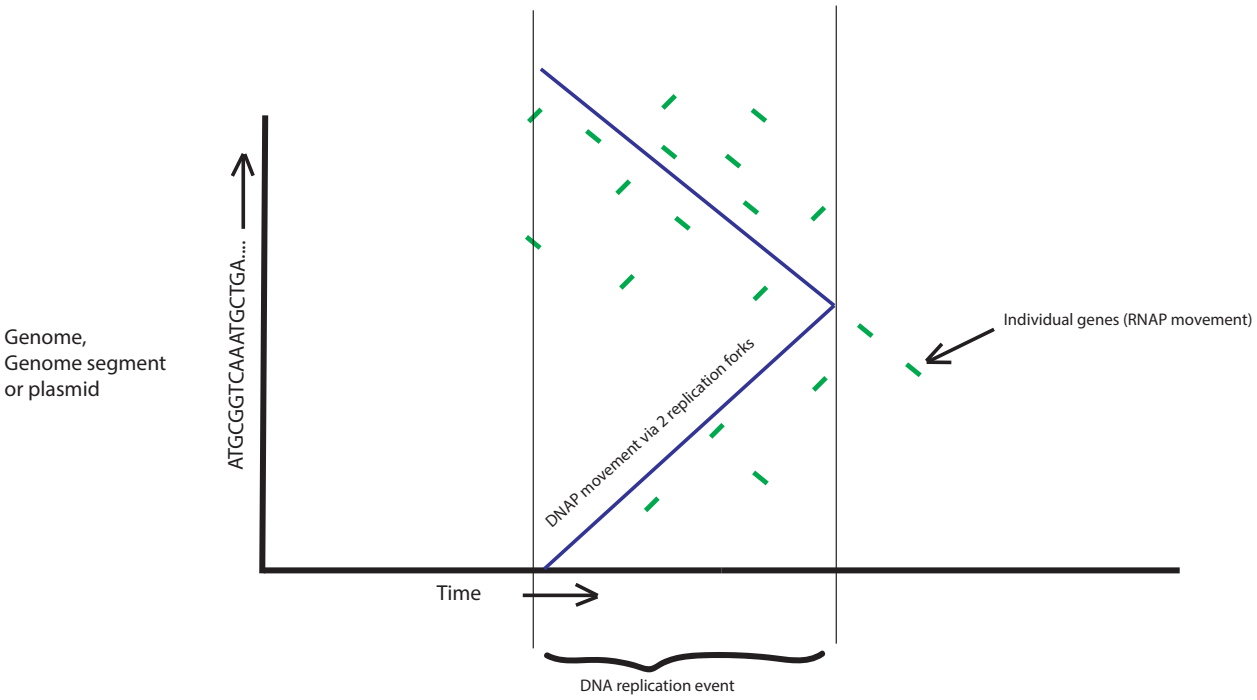


Figure 1: A single, hypothetical, structural “macro-ripple” event of two DNA polymerase molecules moving through the genome through time via helicase unzipping, orchestrated by a swarm of associated “micro-ripples” (gene expression).

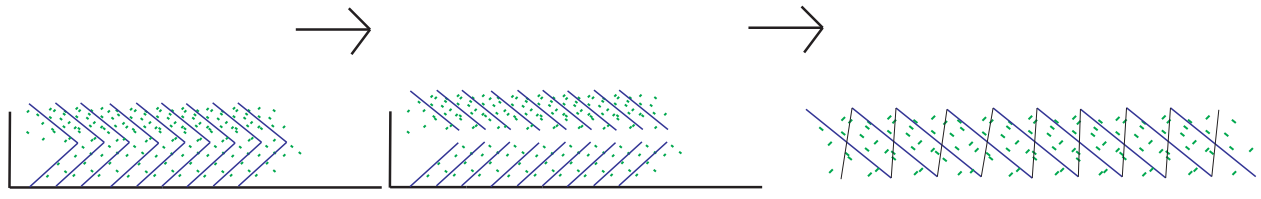

Figure 2: Since the bacterial genome often starts replicating a second time before the first has completed, replication events are shown to be overlapping. The involvement of two DNAP molecules moving towards opposite ends of the genome results in the replication event being divided in two (middle). Periods of non-binding (relaxation - thin black lines), are used to connect periods of binding (excitation, purple lines), to visualize the structural waveform of DNA replication.

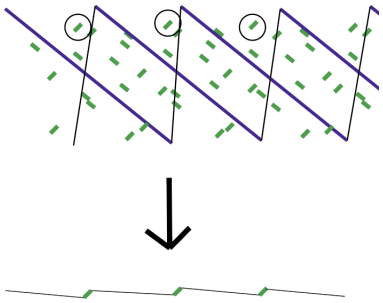

Figure 3: Structural waveform of “micro-ripples” using the same tactic as above (repeated expression of a single gene over time involved in DNA replication).

The smooth integration of human-constructed biological circuits into the near-continuum of replication events in E. Coli began by studying important sets of naturally occurring gene circuits, in particular those involved in fueling said replication events (e.g. metabolically regenerating the propensity for structural ripples in DNA). Upon studying the *lac* operon genetic circuit in E. Coli, which allows for sugar metabolism to switch to lactose upon the disappearance of glucose, a synthetic, molecular “toggle switch” was constructed that could be turned on and off at will via heating and binding of the IPTG molecule. Various forms of ON/OFF bistability have been subsequently integrated into a slew of gene circuits, in addition to more complex representations of boolean logic gates (AND, NAND, NOR, etc.), as well as genetic counting devices, clocks, sensors, pattern detectors and memory mechanisms. Circuits in probiotic E. Coli can be designed to release diagnostic or therapeutic molecules in predetermined situations and environments, or produce antimalarial drugs and other medicines, biofuels and bioplastics, and even extract electricity from our own sewage. Additionally, genetic circuits allow us to peer into the processes of biology at levels never seen before to reveal startling and striking behaviors and patterns. In E. Coli, for example, fine-tuning the molecular coupling of the AHL signaling molecule to proteins on both ends of its life cycle, and then linking this with fluorescent protein expression resulted in ability to visualize propagating waves of bacterial communication. In [Supernova figure], a circular wave of genetic “micro-ripples” correlating to the synthesis of AHL and fluorescent proteins expands outward from the colony center, while right behind it, an “invisible” wave of inhibitory protein synthesis cancels this effect to create a striking appearance.

The genetic circuits of Synthetic Biology, sprung into being through the ease of genetically manipulating simultaneously ancient yet consistently novel and relentlessly remixed strings of bacterial DNA may seem to have progressed rapidly, but there is a great deal more they could actuate, or show us about biology and about ourselves. Biology can be notoriously difficult to teach since there is still so much to be understood, and likewise, the implications of Synthetic Biology can often quickly become painfully abstract to those working far outside its domain. This can make it difficult to see what is under one’s own nose, or hands, or feet, or anywhere else in/on the body. In addition to human bacteria outnumbering our own cells by trillions in number, it is important to not forget that their circular genomes are composed of the same molecular building blocks as ours, and indeed genetic circuits have already made their way into mammalian cells for applications such as immunotherapy. Even after closing one’s own eyes, manipulatable DNA is found to be inescapable, as staring into the darkness is in reality *staring into a cell culture* of over half a million flattened epithelial cells per square centimeter behind the eyelids, each carrying a functionally-modified copy of your genome. There is thus great room for both creativity and caution in Synthetic Biology, and how it will be used to improve the functions of the continuum of metabolically regenerative circuits of each of our genomes will be in the hands of future generations, such as the scores of students competing at the intersection of science and of fun at the annual International Genetically Engineered Machine (iGem) competitions.