RESEARCH REVIEW

INTERNATIONAL MICROBIOLOGY (2015) 18:71-84
**doi:10.2436/20.1501.01.236.** ISSN (print): 1139-6709. e-ISSN: 1618-1095
**www.im.microbios.org**

INTERNATIONAL MICROBIOLOGY
OPEN ACCESS

# Synthetic biology: Novel approaches for microbiology

**Felipe Padilla-Vaca, Fernando Anaya-Velázquez, Bernardo Franco***

Department of Biology, Division of Natural and Exact Sciences, University of Guanajuato, Guanajuato, Mexico

Received 21 April 2015 · Accepted 9 June 2015

**Summary.** In the past twenty years, molecular genetics has created powerful tools for genetic manipulation of living organisms. Whole genome sequencing has provided necessary information to assess knowledge on gene function and protein networks. In addition, new tools permit to modify organisms to perform desired tasks. Gene function analysis is speed up by novel approaches that couple both high throughput data generation and mining. Synthetic biology is an emerging field that uses tools for generating novel gene networks, whole genome synthesis and engineering. New applications in biotechnological, pharmaceutical and biomedical research are envisioned for synthetic biology. In recent years these new strategies have opened up the possibilities to study gene and genome editing, creation of novel tools for functional studies in virus, parasites and pathogenic bacteria. There is also the possibility to re-design organisms to generate vaccine subunits or produce new pharmaceuticals to combat multi-drug resistant pathogens. In this review we provide our opinion on the applicability of synthetic biology strategies for functional studies of pathogenic organisms and some applications such as genome editing and gene network studies to further comprehend virulence factors and determinants in pathogenic organisms. We also discuss what we consider important ethical issues for this field of molecular biology, especially for potential misuse of the new technologies. [**Int Microbiol** 2015; 18(2):71-84]

**Keywords:** synthetic biology · genetic engineering · genomics · pathogenesis · bioethics · artificial cells · astrobiology

## "Defining" life

What is life? This fundamental question has intrigued scientists for centuries and a strong definition is needed in order to understand life and all its manifestations. What are the universal constituents that make a living organism alive? This question is also complicated to give a compelling answer since not all the components, dynamics and interactions inside a living organism are well understood and the relationship among them are also difficult to assess [6]. At first, life has been defined as an entity that is capable of passing on genetic information and is subjected to diverse environmental selective pressures that ensure diversity, but a more appropriate definition considers the following: "Life (a living individual) is a self-sustaining object belonging to a set of elements capable of undergoing Darwinian evolution" [10].

In the case of pathogenic organisms, this definition also involves the host and the selective pressures that affect both organisms. With the "omic" approach we are one step closer to answer this fundamental biological, philosophical question since we can tackle the limitation of the population diversity found in all living organisms and also gain knowledge on the constituents and the interactions they undergo. However, or-

*Corresponding author: B. Franco
Departamento de Biología
División de Ciencias Naturales y Exactas
Universidad de Guanajuato
Noria Alta, s/n
Guanajuato, Gto., 36050, México
Tel. +52-4737320006, ext. 8154
E-mail: bfranco@ugto.mx

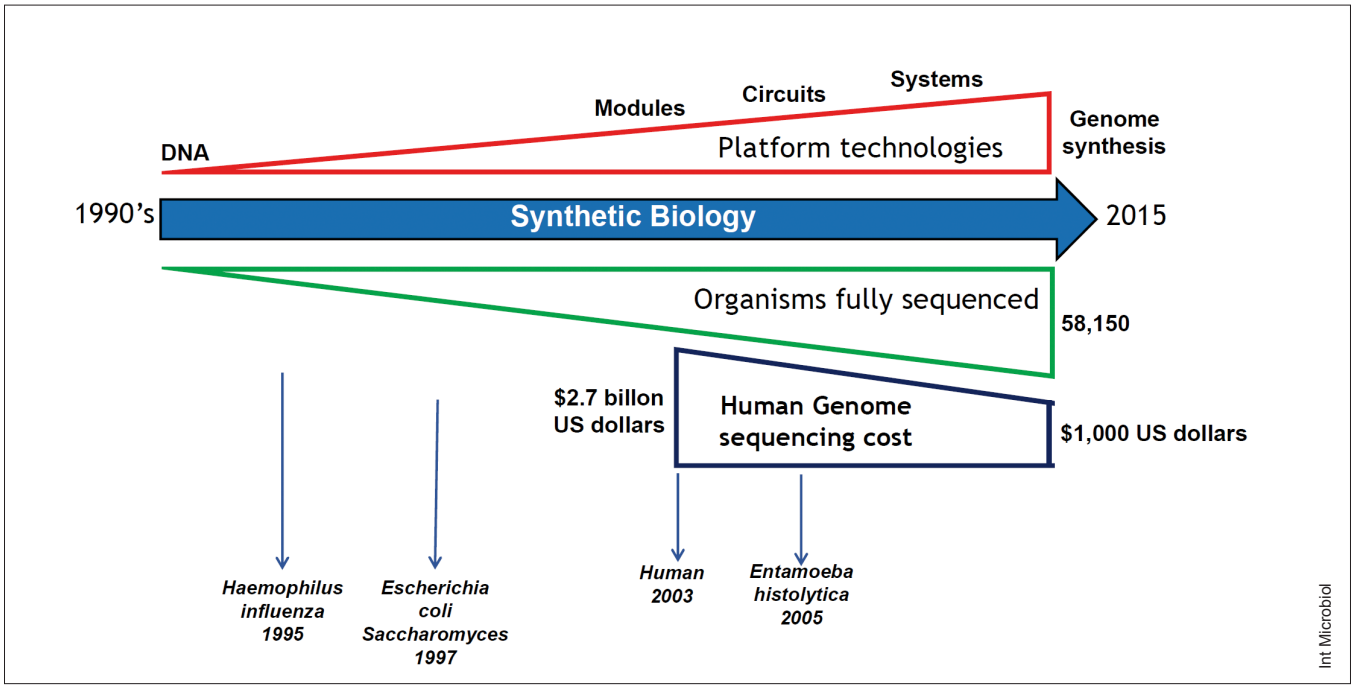


**Fig. 1.** Time line for the development of synthetic biology.

ganisms—even those with small genomes—are too complex to characterize and understand all the biological processes they carry out under different conditions. New technologies are bringing us closer to achieve the goal of understanding how living organisms work and evolve, which in turn may also lead us in the right path to fight pathogens. As presented in Fig. 1 some major technological advances and their applications may help basic and applied research to clarify the most important aspects of pathogenicity and bring about new strategies to fight pathogens. One important step towards understanding pathogenesis is to understand the genes and factors required from both pathogen and host.

## "Reading" DNA

The discovery of DNA structure has been a major advance to understand the molecular basis of heredity [13]. With that discovery, molecular genetics was born and made it possible to study both gene function and the molecular basis of development and disease, which ultimately led to modern biotechnology and genetic engineering. Understanding DNA structure allowed researchers to assess sequencing methods or the ability to "read" DNA that was important to uncover the diversity of life, evolution and taxonomy. Several DNA technologies including cloning, the polymerase chain reaction (PCR), sequencing, next-generation sequencing (NGS) and more recently synthetic biology have broadened our capabilities to study organisms [11,45].

The tools of genomic biology are the most effective analytical tools to determine the gene structure and content of any given organism. Unfortunately, they tells us nothing about gene function and protein networks. However, this is the most fundamental basis for assessing knowledge on gene function and, without it, synthetic biology loses its most fundamental tool.

Up to June 2015, 58,150 organisms have been sequenced that comprise 1,037 Archaea, 44,576 Bacteria and 8,181 Eukarya. In addition, GenBank contains data from 300,000 formally described species in the form of expressed sequence tag (EST), genome survey sequence (GSS) and whole-genome shotgun sequence (WGS) [7]. All these data are the basis for for the study of all model organisms. Having its genome sequenced, research on a given organism can be boosted to find out all gene functions and to develop new tools for manipulating it. This information can open up discovery of gene networks that provide us with comprehensive data for mining essential genes for particular biological functions.

The exponential growth of genome projects is explained in the reduction of costs (the first human genome project costs were around 2,700 millon US dollars, while currently it costs less than 1,000 US dollars) (Fig. 1) [11]. NGS equipment can

be afforded even by medium and small-sized laboratories, where data are required for antibacterial molecule research, vaccine development, diagnostics and epidemiology research [33].

The first genomes that were sequenced were confined to large consortia and a considerable amount of resources to achieve the full genome sequence and annotation was required. The first organism fully sequenced and annotated was *Haemophilus influenza*, a highly relevant human pathogen [22]. The genomes of *Escherichia coli* and *Saccharomyces cerevisiae*, which are the most extensively studied organisms, were both published in 1997 (Fig. 1). Refinement of sequencing and analytical tools to generate faster and more accurate sequences have been developed since then. With NGS technologies, whole genome data can be quickly generated and applied to different biological questions such as total RNA sequencing, which gives more information than microarray data on abundance, half-life and processing of RNA, genome sequencing and enriched pools of environmentally adquired DNA samples [39].

As it is discussed latter, microbial diversity has been uncover using these powerful techniques and more questions than answers have arisen from all the data generated. Genomics has been modified extensively, from Sanger sequencing methods to micro-reactor sequencing and single cell genomics, and it has become of great importance to study pathogenic organisms. Even though most of the information that we can gather on microbial populations come from classical microbiological studies, now it is possible to determine at broad scale the microbial populations present in a sample and their dynamics. One such example is the city-scale metagenomics survey undertaken by Afshinnekoo and colleagues by which several pathogen distributions, antibiotic resistance and large scale new organism identification was assessed and correlated to environmental, geographical distribution and even human associated genomic data [2]. Full genome sequencing and annotation also speeds up the research for identifying essential and virulence-related genes present in a given population or sample (see [24] for some review examples that illustrate this matter).

## Uncovering microbes through genomics

Evidence on genomic dynamics and evolution provides part of the answer to life's diversity. Once an organism genome is fully sequenced, knowledge and research on the biological function of each gene is speed up and therefore more knowledge is gained from functional studies and genome data mining. Since the establishment of public genome databases and standard criteria for genomic data storage and mining, the amount of genomic, metagenomic and pangenomic data has increased rapidly [21]. The standards set for genomic data created by the Genomic Standards Consortium considered the following: standards for new genomic data, methods for storing and sharing the data generated and harmonization of the information so that the scientific community can easily access it [21]. The value of genomic data can be exemplified with the need to achieve rapid comparison between isolates (both by standard microbiological methods or whole sample sequencing), especially during epidemic conditions. Data accession is uniformed using sequence identifiers, and databases are linked together in order to keep information available, protected, curated and up to date [7].

The use of whole genome sequencing have opened up the field of evolutionary and population genomics, which allows to characterize population structure and dynamics, and to know what factors affect the population in certain environments [18]. In pathogenic organisms, this is of outmost importance. Recent efforts have shed light into the factors that modulate pathogenicity. Epidemiology can be studied in much depth since data from whole genome sequencing can inform the pattern of pathogenicity displayed by an outbreak, resistance to antibiotics, virulence and persistence factors. Also, information about mobile genetic elements, horizontal gene transfer and adaptation features can be uncovered. Such information can be obtained not only as a general or representative sample, but also at the "local" level, such as in the event of a recurrent or emergency outbreak, contaminated sources or even at the level of hospital related infections [15]. Other aspects of microbial life have been revealed by genomic biology.

We are facing the "dark matter" of microbial life. This considers all those organisms and even divisions that have eluded isolation and characterization under laboratory conditions. McLean and collaborators showed that a "mini-metagenome" could be generated from rare events. The authors sequenced whole genomes from single cells isolated by flow cytometry (again, converging many techniques for new purposes) from a sink in a hospital restroom [46]. This type of work leads research in a different direction: single cell genomics. In this particular study, tackling the heterogeneity of samples (sequence itself or G+C content) as well as contaminants is cumbersome. But with the implementation of novel computational toolsfaster and more accurate sequencing methods could be developed. Once the genome of an organ-

ism has been fully sequenced, the next step is to analyze it in order to obtain useful information. The study of pathogenesis and host-pathogen interactions has been favored by genomics.

In the case of virulence related genes it can be complicated to be certain of their specific role. Even defining virulence is not easy. It is often defined as the capability to damage or harm the host when invaded by a pathogen, but this definition does not consider the role of other factors that are important for survival within the host. Genomics provides the challenge when there is a dramatic increase in host damage or spreading or when pathogens shift to infect new host species (such as SARS or Ebola virus), resulting in new and devastating outbreaks and epidemics. The use of high throughput sequencing methods, can allow to define natural reservoirs of deadly pathogens, ecological interactions and perhaps to understand several of their pathogenesis mechanisms and how disease evolves. Sequencing environmental samples, such as pan enome sequencing projects, have revealed the true microbial diversity that is making evident also the real landscape of pathogen–epidemiology–persistence. Hand in hand, this also may lead to important health-related discoveries for antibiotic resistance and pathogenicity statistics near real-time.

## Microbiomes and health

Disease is also related to our own microorganisms. Recently new sequencing techniques have revealed that humans hold an immense universe inside, which modulates many biological functions. Research efforts concentrate to achieve the power of all the living organisms inside the body to promote immune boost, vaccine delivery systems and keeping a safe environment for beneficial microorganisms. We now have knowledge on some effects of the microbiota present in vertebrates; one such example is the effect on circadian clock disturbances as reported by Thaiss and collaborators [59]. But how farther can we go? From environmental genome sequencing we learned that emerging pathogens are lurking that can potentially infect humans. One example is the enigmatic chlorovirus ATCV that usually infects algae but also can cause changes in cognitive functions in humans and mice [69].

The human microbiome plays a major role in health and disease and its actual composition is variable for each individual. Gut microbes have been proposed to regulate behavior and social skills needed for the bacteria to colonize other individuals. The Human Microbiome Project has two important goals; study the composition of healthy individuals and to understand the effect of changes during disease. Also, this project bypassed the limitation of cultivating and characterizing the microbial communities in samples from patients or experimental animals. NGS have provided a glimpse on the composition profile of normal intestinal microbiome and opened the possibility of treating certain disorders in patients. The Human Microbiome Project is a good example of how modern science should thrive, as a collaborative and multidisciplinary effort. In a future, with the available data on the host's microbial DNA and transcriptome, we will be able to predict the metabolic capacity of the bacteria present in the host and measure its impact on health besides the host lifestyle and environmental stress [62].

With novel bioinformatic and statistical tools, identifying dysbioses (changes in the normal microbiota content) becomes easier with the possibility of developing new treatments for diseases such as chronic inflammatory bowel disease instead of using fecal transplant on patients [67]. Analyses that identify perturbations in molecules related to disease can be used to interpret their physiological role, interaction with truly pathogenic organisms or integrate sequence data with whole-community relationships.

New technologies are not problem-free and require troubleshooting. One good example is the amount of false positive organisms present in samples and library preparation, for example in ancient pathogen identification (which is a problem in any given sample), can give false diagnostics and population content biases. NGS requires stringent cleaning measurements and laboratory reagent preparations in order to avoid cross contamination and misleading results, especially with rare samples [53].

Technology and human wit have designed novel approaches to answer complex biological questions and problems, for instance, how many organisms are present in dust and airborne microbial communities [71]. This particular question poses a major problem of low biological sample content and avoids growth to eliminate biased conclusions. Integrating engineering for designing and applying specialized apparatus for sample collection, enrichment and analysis, and novel computer algorithms to assemble and analyze data, provides an image of microbes present in air samples. All this major technical advances are needed only to read DNA properly and then make some sense out of all that information. There are remaining questions about environmental samples, such as how many bacteria are metabolically active, capable of division and the cycles between population densities.

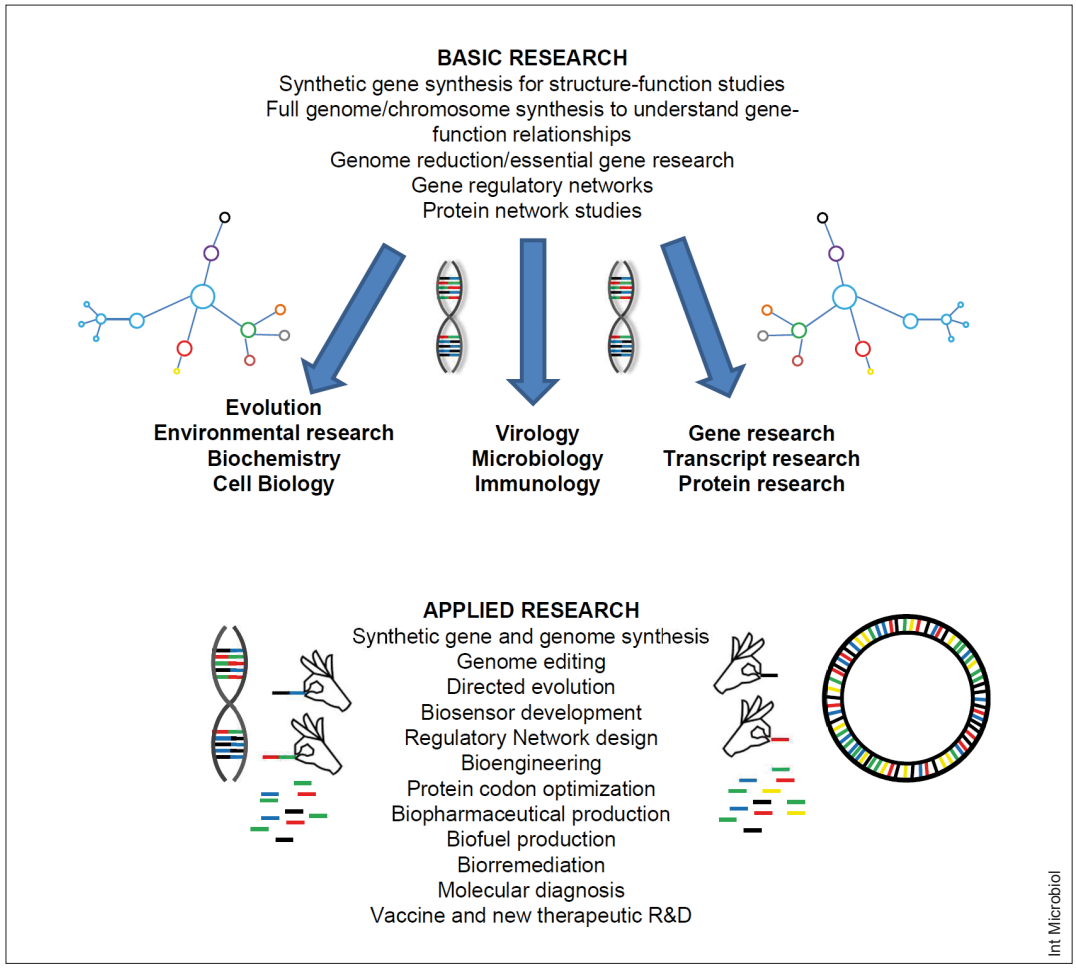


**Fig. 2.** Basic and applied research areas that synthetic biology has important influence.

## Understanding the genome

"*Data does not equal information; information does not equal knowledge; and, most importantly of all, knowledge does not equal wisdom. We have oceans of data, rivers of information, small puddles of knowledge, and the odd drop of wisdom.*" Henry Nix made this statement in his keynote address (A National Geographic Information System–An achievable objective?) to the Australasian Urban and Regional Information Systems Association in 1990) [32]. In fact, gathering terabytes of data does not mean givung them some sense or fully understand them. Figures 2 and 3 show some applications that data so far gathered can have in different fields. High-throughput analyses that have been carried out so far for genome, transcriptome and proteome assays, have provided more data that we can handle and make sense of it. Most of the research in 'omics' is now done by molecular biologists, biochemists, mathematicians, physicists and computer experts, and the main reason for pluridisciplinarity is to find useful information and search for patterns in all that data.

Genomic data tell us which and how many genes are present in any organism's genome, but cannot tell neither the function of all them nor the roles of protein networks, which requires specific experimental data. Moreover, gene structure and G+C contents are quite different between bacteria and eukaryotic organisms, so the algorithms for analysis must take into account different approaches for mining data, for example for generating comprehensive data for regulatory networks or genome analysis and annotation. With massive genome data, it is also possible to know some of the evolutionary relationships between organisms and the influence of the environment on genome structure, regulation and diversity (Fig. 2). In pathogenic organisms this is of outmost impor-

tance, since the lifestyle of each organism modulates the genomic landscape and protein effector machineries that give rise to the phenotypes of pathogenic organisms (Figs. 2 and 3). In addition, interactions between the microbiome and host cells can remodel the genomic landscape in pathogens. Even with the most powerful computers and algorithms available, genetic sequence only tells half the story behind the physiology of the organism and the possible roles of the protein networks that it can use at a given point of its life cycle,. Most useful data are represented by well-known genes, but it is not possible to make sense of unknown open reading frames.

Two examples lead us to reformulate life and the dynamics that render diversity. These two examples can further increase our knowledge on gene and protein networks. The first one was developed by Suzuki and collaborators [58]. Using a novel approach of genome assembly [27] the authors generated multiple deletions (clusters from 5 to 24 genes) on the genome of *Mycoplasma mycoides* synthetic genome JCVI-syn 1.0 and marked each with a reporter protein (green fluorescent protein, GFP) on each insertion. Endeavors like this can only be achieved with full genome sequences. In this particular case, the complete nucleotide sequence (580,070 base pairs) of *Mycoplasma genitalium* [23], the smallest known genome of any free-living organism, was used to start building up large-scale molecules. The technique of genome transplantation was used to replace the full genome of a recipient cell and changing one bacterial species into another and to generate a new strain controlled by a chemically synthetized genome [27,40]. In this genome, 470 predicted coding regions identified include genes required for DNA replication, transcription and translation, DNA repair, cellular transport, and energy metabolism.

Using the mating machinery of yeast, several deletions were generated at eight unique deletable regions of the genome, eliminating in a single step 91 genes and approximately 10% of the original synthetic JCVI-syn 1.0 genome. With a second round of deletion and selection, the new strain contained a deletion on seven of the eight targeted regions, representing 84 genes. This approach also considers the use of Tn insertions to characterize the content and targeting of several genes and determine their essentiality. The engineered genomes were analysed to look for growth defects, but most strains retained doubling times similar to the original strain, except for JCVI-syn1.0 ΔL strain and the multiple deletions strain JCVI-syn1.0 Δ1–6 ΔA ΔB ΔC ΔD ΔE ΔI ΔL ΔN, which had growth defects. Under the conditions tested, genetic interactions of 91 genes were found. Despite predicted effects of the deletions on sugar metabolism, cell envelope and DNA metabolism, growth rates were unaffected by gene deletions in seven strains. All the data generated gave rise to the possibility of generating deletion strains or minimal genomes using rapid strategies and designing cells with the desired characteristics. The data also pointed out to the minimal set of genes required for life. Essential genes are important for biotechnology applications. In fact, the limiting steps for most metabolic reactions can also be essential for cell physiology, and those genes, which are not present in the host's genome, can also render specific targets for new drugs against pathogenic organisms.

Essential genes for *Mycoplasma* had been previously assessed by Glass and collaborators [30]. They carried out an extensive mutagenesis study in *Mycoplasma genitalium* to determine the number of essential genes required for lifeand was the milestone for generating the first chemically synthesized bacterial genome [30,34]. Both studies showed that there is a minimal set of genes to make a completely functional bacterial cell and that in vitro synthesis of a genome from scratch can give rise to a functional genome with the desired characteristics. In our opinion, this approach is complementary to systems biology and all the 'omics'. Instead of handling extensive amounts of data, it creates and puts to the test whole engineered genomes. The ultimate goal is to simplify the created organisms with the minimal set of genes, given that, for *M. genitalium,* from 265 to 350 of the 480 protein-coding genes are essential under laboratory growth conditions, including about 100 genes of unknown function.

Now we can start characterizing at a large scale the world of genes with unknown function, something relevant and difficult to approach. But even with the minimal set of genes, understanding the function is still out of reach, especially when mutants of genes without evident function also lack any phenotype. These studies also made clear that even if "reading" DNA has been already achieved, understanding how to "write" DNA and make long functional "sentences" is still at an early stage of development.

Other studies in organisms such as in yeast with the same aim indicated that 12% of its genome was essential and that almost 70% of the genomic disruptions gave no new phenotype [31]. Now the question arises for other organisms with more complex genomes. Estimates on *Bacillus subtilis* indicate that only 9% of its genome is essential, rendering a nonessential 562-kb genetic material from its total genome [36]. With the birth of synthetic biology, microbes can be assembled for bits and render a desired phenotype or even behavior, thus providing a better understanding of how a cell works [20]. Other studies conducted in *Escherichia coli* have revealed that, from the total 4288 genes it has, 303 genes were unable to be deleted from those 37 of unknown function [70].

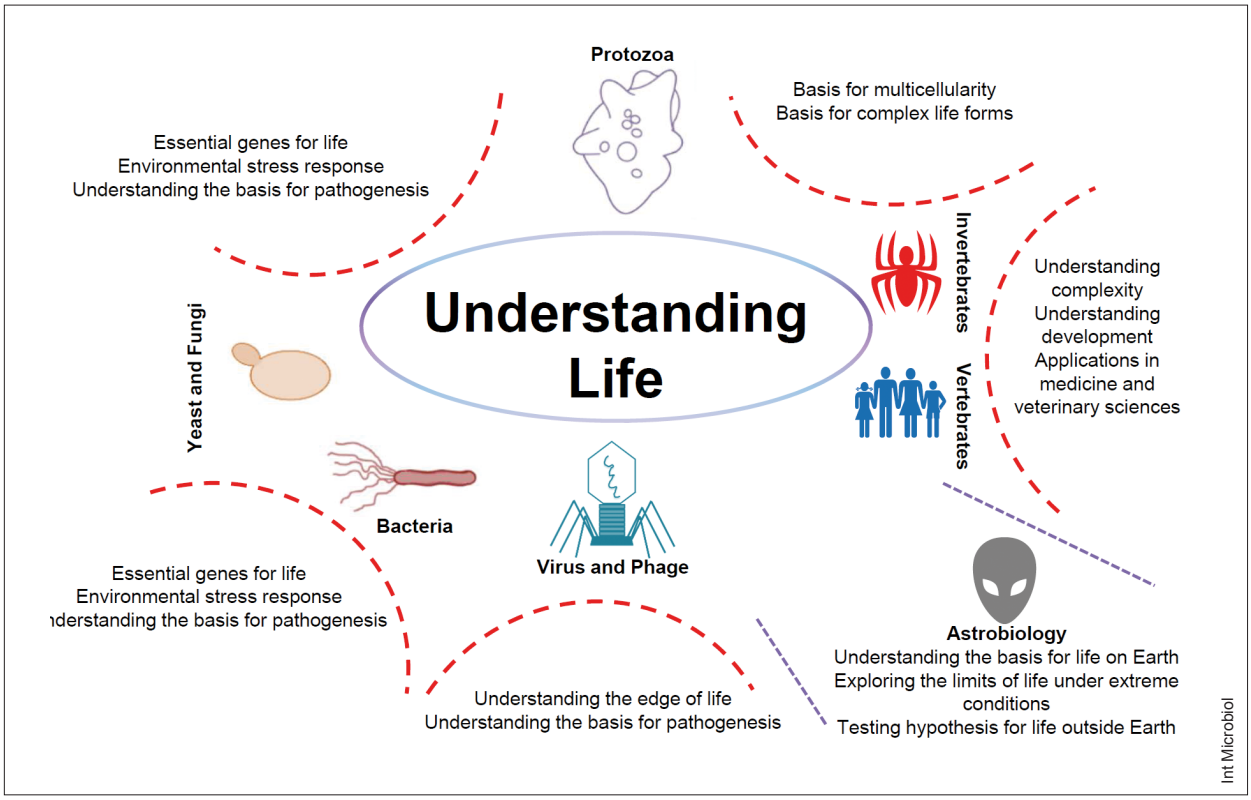


**Fig. 3.** General overview of possible applications of synthetic biology for microbial and other organisms. Our understanding of life from the single cell to the population and multicellular levels.

This study used a novel technique for interrupting single genes using PCR products containing a selective marker that had been used also in other organisms [14].

Research with microorganisms has the advantage of having molecular tools for generating diverse mutants and genetic engineering, a task that is more complicated on higher organisms. In humans, knockouts are off-limits, but we can learn from naturally occurring diseases. One recent example of gene mutation and deletion in humans that renders inactive genes can shed light on possible future research on the general population for dispensable genes. Recently Sulem et al. identified deletions of up to 1000 autosomal genes from the genomes of normal human individuals. Thus, at least 1000 genes can be eliminated without observing defects. As defined in the research paper, these are "healthy knockouts" [57]. From this concept, a novel field has arisen that looks for genes that, when they are missing, can confer a benefit, such as resistance to certain diseases, including AIDS. The effect of a disease-related gene can be mimicked by blocking the normal gene's protein the absence of that gene is known not to be a trheat for the host. Regarding health-associated organisms, scientists are trying to understand the role of diverse gene networks that are involved in virulence and pathogenesis. Using cumulative data and with the help of systems biology (i.e., an holistic approach to biology) the fundamental cellular processes in organisms can be simulated [60] (Fig. 2).

All the data collected thus far from the minimal gene set or essential genes in diverse organisms can lead to a better understanding of cell basic functions and in a near future a whole picture of what is required to generate a fully functional living organism or at least make more accurate predictions on protein networks and the interactions within an organism.

From the biotechnological point of view, this may be the future for all pharmaceutical advances and pipeline production of different molecules, which must be expanded from *E. coli*, *Mycoplasma* and yeast due to ecological niche restrictions or environmental restrains. Synthetic biology can broaden our knowledge on organism's basic functions and the regulation of life processes using a bottom-up approach of engineering (Figs. 2 and 3).

## "Writing" DNA: engineering biological systems

Our ability to 'read' DNA is far better than that to 'write' it. Although the development of software and databases have made this task easier, most of the biological processes are not yet well understood. The first modest attempt to generate a synthetic DNA fragment led to one of the most fundamental advances in molecular genetics, breaking the genetic code [1]. One impediment was related to costs and fidelity with DNA in vitro synthesis. The history of DNA synthesis has been brilliantly reviewed in [52], so we will focuse on its applications. This technology, which allowed sequencing and PCR, nowadays is used also to produce synthetic genes in a cost-effective manner or large-scale genome synthesis. There are major advances in large DNA synthesis with important improvements in fidelity, length and yield. The transition from column-based synthesis to array synthesis, allowed companies to offer more cost-effective, higher yields and fidelity on the desired sequence, also making less expensive to generate full open reading frames (ORF) and with the desired codon usage for the proper host [36].

Higher fidelity synthesis either by PCR or de novo DNA synthesis is a major achievement for microbial and pathogen research. Both techniques allow to tackle the restrictions of construct design (which for certain applications can be troublesome) and assembly, gene expression and regulation, codon usage bypasses, transgene and vector creation, mutagenesis, protein engineering, and reporter protein adaptation for a particular host, creating new sequences with the desired codon usage for specific hosts among others [38]. One excellent example is the synthesis of nearly 44 Mb for knocking out with iRNA most human and mouse genes [12]. For protein studies, antimicrobial peptides or proteins and vaccine subunit preparation and purification, codon usage has proved to be a major hindrance to obtain high yields of recombinant proteins in many hosts, in particular for *E. coli*, which is a cost-effective host [5].

The motivation to review this subject was especially to encourage other researchers to explore synthetic biology as a powerful approach to complex problems. It is of particular interest to our laboratories to improve existing tools for understanding the fundamental basis of pathogenesis in the parasitic protist *Entamoeba histolytica*. There are tools for manipulating this organism, but there are also limitations to generate mutants and versatile plasmid vectors or reporter genes. Synthetic biology, however, can help to overcome these problems. We are also interested in the improvement of genetically encoded biosensors to couple metabolite, metals, nanoparticles, toxins and other environmental or biomedical relevant molecules to several outputs and easier to detect [48].

But, what does synthetic biology do? Using a broad definition, synthetic biology is in the quest to simplify our understanding of biological entities (viral, bacterial and eukaryal organisms) by constructing biochemical or genetic pathways both in vivo and in silico and building up computational models to simulate the behavior of those pathways with the ultimate goal of testing them in the real world [65] (Fig. 3). Synthetic biology also attempts to generate genetically recoded organisms or biological entities by using a design process more systematic and predictable and by analyzing models that use all the data available for that particular engineered process, robustness in the output of the designed organism, scalable to different niches or conditions, and ideally, more efficient than the wild type counterpart [65]. These criteria cover basically the necessities for applied genetic engineering with the easy approach of designing modules or parts to do specific tasks, which in a complete organism is still unpredictable at a large scale due to problems of interacting protein networks or enzyme cascades that are self-regulated or that interact with other pathways.

A major motivation for massive genome engineering is the applicability of modified microorganism that generates a profit or can contribute to generate new biologically based processes [66]. With the birth of synthetic biology, microbes can be assembled for bits and render a desired phenotype or even behavior, thus providing a better understanding of how a cell works. This approach, which is complementary to systems biology and all the 'omics', instead of handling extensive amounts of data, creates and puts to the test whole organisms with truly engineered parts and not using the traditional random mutagenesis or directed evolution and selection approach [25]. Now we can build biological parts and genetic modules that are not found in nature or that are poorly characterized due to difficulties on growing the desired microorganism, with the purpose of generating fully functional genetic circuits that can modulate the behavior of cells and respond to specific stimuli. Synthetic biology shares many similarities with genetic engineering principles such as minimality to avoid complex interactions and futile cycles, modularity to improve interactions and sustain specificity, and controllability of a complete system [16] (Figs. 1 and 2). Controllability of any genetic circuit is important to achieve the desired results, but also to prevent modified microorganisms from escaping from the lab (see "Potential misuse of the new technologies" Section). When working with pathogens difficult to handle and transport, an

alternative option might be inactivation and sequencing at the deployment site and revitalization in a laboratory for further scientific assessment [72] (see Section on "Potential misuse of the new technologies").

## The toolbox

Based on engineering principles and specialized syntax, BioBricks can overcome the troubles found when designing large molecules composed of different modules comprising regulatory, coding and terminator sequences to generate circuits or recombinant plasmid vectors [51]. Using wild type or designed promoters, coding regions, terminators and reporter genes, we can integrate them to characterize cryptic genes or gene clusters, create new reporter plasmids and proteins, simplifying maintaining reading frame with the coding regions and the assembly itself. This particular strategy is an assembly method based on type II restriction enzymes, in which building blocks are assembled directionally by adding standardized flanking and complementary sequences that specify the orientation. The use of a reference syntax and access numbers, BioBrick parts sequence and properties are deposited in a specialized database [the Registry of Standars Biological Parts, http://parts.igem.org/Main_Page] available to the synthetic biology community and can speed up the research world-wide. Using traditional restriction and ligation reactions, assembly and joining together all the elements needed is easier when using in-house software to generate regulatory circuits with natural or artificially upgraded regulatory sequences according to the needs or aims.

Using modifications of the restriction enzymes for each BioBrick, the combinatorial capacity is sufficient for even constructing fragments of over 20 kb (such as the secondary metabolite actinorhodin gene cluster from *Streptomyces coelicolor*) [43]. However, this strategy is time consuming and troublesome for joining multiple genes or DNA fragments using different joining sequences and finding the proper endonucleases to achieve such constructs. Moreover, mutations can arise at the overlapping Section. Protocols using modified versions of BioBrick methods can produce faster results using PCR-fusion techniques were each BioBrick part contains overlapping ends and can be used to fuse up to four individual modules. However, there are still some limitations regarding fidelity during synthesis and assembly success, especially when designing the overlapping regions [55]. Improvements on this technique led to the Golden Gate assembly method where subcloning of up to 9 different modules (as undigested plasmids) can be cloned into a recipient vector with the usage of type II restriction enzymes and it is done in one step in a single tube with a 90% efficiency, rendering several clones that can be analyzed to verify sequence and orientation of the fragments [19]. This strategy is useful to shuffle modules of different molecules to generate combinatorial versions of it and improve its characteristics. It is especially useful to generate proteins with a desired activity or even for the production of vaccines in which epitopes can be screened for higher reactivity against anti-serum from patients. The generation of bigger constructs faces technical problems, including the lack of restriction enzymes to be used in a highly specific manner. Due to such problems, other techniques have been developed for both in vitro and in vivo assembly. The most prominent in vitro methods are PCR-based with overlapping 15 or more bases and require in vitro recombination such as In-Fusion, SLIC and Gibson, which are more efficient for generating kilobase-sized fragments which led to the assembly of a fully functional synthetic bacterial genome and more recently the complete and engineered chromosome III of yeast synthesized stepwise [4,28,54]. With the possibility of synthesizing larger DNA molecules in vitro, the next logical step was to start the assembly of genes and genomes. Now the assembly of viral, bacterial and yeast chromosomes is possible. Before attempting the synthesis and assembly of large DNA molecules and facing the ethical implications of generating a fully functional organism, Smith and colleagues first assembled the ϕX174 genome under strict ethical evaluation (see also "Potential misuse of the new technologies" Section) [56]. A serious limitation is the formation of multimer assembled molecules, for which a method has been recently proposed to avoid using linear fragment assembly and ligation, followed by in vivo cyclization after transforming recipient *E. coli* cells [32]. Methods used for generating synthetic chromosome III of yeast involved 750-bp modules and all the techniques previously described. In this experimnent, undergraduate students from the Build-A-Genome class at Johns Hopkins University were involved in generating the starting building blocks as part of a class project [4]. The cases of *Mycoplasma* and chromosome III from yeast synthetic molecules, in which the cells remained viable and functional, set the foundations for more aggressive engineering for gene function and protein network studies, as well for full organism engineering. The accurate assembly of large DNA fragments is still challenging and requires refinement in order to generate fully functioning genes. Gene clusters must be done in vivo as described for *Mycoplasma,* and yeast synthetic genome and chromosome III using overlapping fragments or pools of oligonucleotides, require yeast recombination machinery [26,32].

Comprehension on essential genes is one step further into knowing new molecular targets for new drugs against pathogens (Fig. 3). To achieve this, extensive knowledge on each particular genome is necessary. With the discovery of enzymes capable of cleaving DNA at specific sequences, now this can be performed.

TALEN (transcription activator-like effectors) or CRISPR (clustered regularly interspaced short palindromic repeats)/Cas both can be used nowadays to edit specific targets in any genome by designing either the zinc finger specificity in TALEN zinc finger nucleases ZFNs) or the RNA guide molecule that directs enzyme specificity [8,29]. TALEN technology relays on ZFNs that are fairly specific to triplet sequences and can cleave specific regions on any given genome and can be engineered (in combination with the techniques described so far) to target cleavage to a desired sequence [29]. The CRISPR/Cas system in bacteria is a prokaryotic adaptive defense barrier against foreign DNA and is the most basic form of adaptive defense mechanisms against foreign DNA inside the genome that can be active as a phage or inactive [29]. One spectacular feature is that this system learns to recognize the self from the non-self DNA and it can adapt to new molecules invading the host's genome. This novel system has been adapted in numerous ways to visualize (by only binding to target sequence), edit and control genes and gene expression in several organisms [29].

One of the limitations of this enzymatic system is that the isolated and characterized CRISPR/Cas enzymes have high molecular weight, making large constructs which are limited to transfection on eukaryotic cells. More recently, a smaller version of Cas9 has been isolated with the same genome editing properties and can be used more extensively than the previously isolated enzymes [50]. With all the current tools available we can envision that new tools can be generated. As already mentioned, many microorganisms have different G+C contents, codon usage and different cellular properties that can make certain studies more difficult. In such cases, synthetic biology offers techniques than can overcome such complications.

Plasmids for different purposes can be easily generated and adapted to each application, like expression plasmids designed with the regulatory elements necessary to render proper expression using limited information (for organisms which genome sequence is still underway or just partial sequences are available), codon usage bypass for gene function studies and reporter protein applications. Other examples are plasmids for the generation of genetic circuits with specific functions that can share light on gene function or protein network interactions, and codon usage expansion and incorporation of artificial amino acids for structural studies, or when protein expression is limited in alternate hosts due to specific post-translational modifications that are under study, such as glycosylation. In our laboratories we do research to discover novel virulence factors and determinants in parasitic protists. We now explore the generation of novel tools using approaches such as the one reported by Wegner and colleagues, who described the generation of a fully functional vector for *Plasmodium falciparum* by using the Gibson assembly method [64]. This kind of studies motivated us to expand our comfort zone and move into synthetic biology.

## Applications for microbiology

With the increasing antibiotic resistance in pathogens, the quest for new molecules that can attend the needs of patients requires novel approaches (Figs. 2 and 3). Nichols and colleagues have developed a high-throughput platform, the iChip, to grow and isolate bacteria [42]. This device allows to grow bacteria directly on soil avoiding the problems of growing microorganisms in the laboratory. Using this technology, a new antibiotic was isolated (teixobactin) without observing resistant mutants of the bacteria tested [42]. This technique offers the possibility of identifying and characterizing new antibiotics as well as new organisms for NGS. Coupling both strategies, synthetic cells can produce the desired secondary metabolites and open a new field for developing novel pharmaceuticals. A good example is reviewed by Nikel et al., who point out the applications that can be generated using synthetic biology applied to pseudomonads also taking the advantage of using the ecological niche of each organism isolated and characterized. If synthetic biology can be used to adapt organisms to grow in the lab, in its corresponding ecological niche and control them, the applications are endless for understanding the biology of of microorganisms, in particular molecular, cellular and environmental-ecological characteristics [47]. Molecular tools can be used also to study gene function and gene relationships. In the case of pathogenic protists, there are certain limitations regarding genetic manipulation. We have tools for transfecting protists and a limited set of vectors, but this is now not a limiting factor. King et al. have described several novel genes involved in phagocytosis in *Entamoeba histolytica* that were discovered by using a genome-wide overexpression screening or the use of CRISPR/Cas system to generate single or multiple mutants in *Trypanosoma cruzi* or *Toxoplasma gondii* [37,49].

Finally, epidemics and the role of emerging diseases with host shifts that can render more virulent strains (such as Ebola or SARS) are a growing concerned worldwide. In the previous sections we have provided some examples of the power of NGS to address major epidemiologic problems including the role of pathogen distribution even at city scale. One limitation for vaccine production is the expression of proteins for vaccine research and mass production. This powerful technique can solve this problem by rapidly synthesizing several epitopes from fully sequenced virus and testing for their effectiveness as vaccines. One example is the use of hemagglutinin and neuraminidase of influenza virus, optimized for their expression in MDCK cells. With such studies the response to epidemics can be improved and even optimized for local outbreaks with a particular mutant strain using sequence data from patients during an outbreak [17]. This kind of technology can also be used at large scale for other purposes, for example to produce full-length viral particles or several antigenic proteins from pathogens with genomic differential G+C content or that are difficult to grow in the laboratory. This strategy can be applied also in veterinary medicine. This kind of technology can speed up the trajectory from development, production and commercialization of both biopharmaceuticals and traditional antibiotics; the end molecules are the same, just the process of generation changes. Synthetic biology is useful not only to improve the yields of natural products of producing strains but also for chemical structure diversification to generate new active analogues.

As above mentioned, the microbiome plays a major role in health and disease. New sequencing techniques have revealed that humans are a holobiont, and their microorganisms can modulate many biological functions. Some recent research efforts have been made to achieve the power of all living organisms inside the body to promote immune boost, vaccine delivery systems, diagnostics, biosensors for different diseases and keeping a stable environment or promote the growth of beneficial organisms [68]. Genome editing strategies may have also therapeutic applications. One challenging task is to safely remove viral particles from any given genome, which under these terms seems science fiction. However, the genome editing capabilities of the CRISPR/Cas system theoretically might reduce the total viral genomes integrated to the genome in any integrating viral infection.

Using NGS on several infected cells we can design the specificity of either TALEN nucleases or Cas9 enzymes to edit and delete viral sequences, in the same manner that bacteria do it. This was achieved in human cells infected with HIV [42]. The application of the CRISPR/Cas technology to reduce the total viral pool in cells infected with HIV is still at an early stage, but the some results so far obtained are promising.

## Potential misuse of synthetic biology

With all the technology at hand and the resources to explore new boundaries, and despite of the oncoming benefits for science, and biotechnology in particular, many citizens including scientific community are wondering how dangerous this emerging field of biology can be. Definitely the understanding of virus, bacteria and protists that affect humans can increase with technologies such as the described above. However, how ethical is to synthesize a fully functional and potentially dangerous organism in order to study it?

With the announcement of the synthetic *Mycoplasma* genome, during a press conference, Dr. Hamilton Smith was asked about the potential application of this technology as a bioweapon. His answer was: "We could make the small pox genome". At that moment, Dr. Venter tried to soften his colleague's statement by saying that DNA is not infective by itself, and Dr. Smith insisted that "But you [Dr. Venter] and I [Dr. Smith] have discussed ways to get around that". Dr. Smith finally said "I probably shouldn't have said that, huh?" [63]. The only limits to this technology are set by imagination. We can sinthesize many things, but science is strictly regulated and requires highly trained personnel to achieve the desired goals. For example, when Smith and colleagues attempte to create the synthetic ϕX174 phage [56], it took almost a year to set up a bioethics committee to review the proposal and deliberate that "[researchers are] taking a reasonable scientific approach to an important biological question" [9]. The biological question at hand can lead to important findings regarding the basis for life and perhaps one day also to understand what is necessary for life to flourish elsewhere. The tools generated with the ϕX174 lead Venter's group to a 10-year quest to achieve a fully functional chemically synthesized genome [27]. The "bottom-up" approach to generate genomes can lead to several unethical applications as envisioned by many since databases contain the genomic sequence of all kinds of organisms including pathogenic and potentially turned into weapons, but can this be real?

A major concern is the introduction of synthetic organisms that may be harmful to others or to the ecological niche. Synthetic organisms should be contained to laboratory or controlled conditions. Mandell et al. [44] reported the generation of two different modified strains of *E. coli* so that they exhib-

ited metabolic dependence on non-standard amino acids and demonstrated that they were impaired of bypassing this biocontrol by horizontal gene transfer or mutagenesis [44].

This kind of research brings us closer to generate fully synthetic organisms that could be used on controlled environments and to rewire organisms to synthesize biomolecules with the desired function without the obvious dangers of releasing an uncontrolled organism into the environment. The synthetic biology community is aware of those dangers and works to provide more safety features to genetically modified organisms. It is true that infective viral particles, bacteria or other pathogens could be synthezided as bioweapons. However, the technical challenge is too great to consider it a constant treat. Synthetic biology techniques are more beneficial to understand the pathogenesis process than to do harm. One excellent example is the ability to obtain and characterize viruses that are not cultivable or can share light on the structure and infectiveness of virus no longer present and only their genomic sequence is available and renders it fully functional [61].

## Concluding remarks

Synthetic biology offers powerful, elegant techniquea to reduce genome sizes, generate new molecules or tools and expand our knowledge on the essential genes of bacteria and yeast and maybe in a near future for any given organism. Technology based on synthetic biology can provide not only tools, but extensive new knowledge of gene function and in particular of essential genes. In our opinion, the fields that can take the most advantage from this technology are: biotechnology, medicine, cellular and molecular biology, and astrobiology. In the case of astrobiology (Fig. 3), we envision that, once the basic rules for life are understood here on Earth, life in other conditions can be tested on varying conditions and even at extreme environments such as the International Space Station. Synthetic organisms can be a powerful tool to study the requirements for alien conditions and hypotheses can be tested based using this experimental approach.

Genome engineering can lead to several and highly significant biotechnological advances. In other areas we expect that knowledge on protein machinery assembly and interactions can one day answer the most fundamental aspects of life and in particular what is required to generate functional gene and protein networks (Fig. 2). Thereof we can approach not only microbes in general, but pathogens or hosts, providing basis to advance our understanding on the molecular and cellular basis for disease.

All these achievements require computational and molecular biology skills that our youth should develop and get involved as with the yeast chromosome III synthesis. With standardized methods of assembly and the design of particular modules that can be interchanged, the International Genetically Engineered Machine (iGEM) Foundation has established an international competition to propel youth into the field of synthetic biology [http://igem.org/Main_Page]. Soon genetic circuits and synthetic organisms will be part of basic molecular genetics courses at colleges and high schools.

Can this new technology be a potential hazard? As with any new technology, it is possible to be used for warfare and biological or organism-based terrorism, but the technical and scientific infrastructure needed renders this possibility to the minimum as for many other technologies. For that reason it is important to emphasize that scientists must avoid the modification of microorganisms for bioweapons production [3]. We envision more benefits than treats to synthetic biology and we encourage the scientific community dedicated to study microorganisms and pathogens to turn their eyesight to this new and exciting new field of molecular biology. It can expand the knowledge on microorganisms, neither harmful for the environment nor for living beings. Studies on pathogenicity mechanisms, for instance, will benefit from the responsible manipulation and ethical use of modified microorganisms.

**Acknowledgements.** Authors acknowledge support for the research conducted in our labs from CONACyT (CB-2012-01 182671), Promep-SEP (Grant number F-PROMEP-38/Rev-03 SEP-23-005) and University of Guanajuato (Grant number FO-DAI-05).

**Competing interests.** None declared.

## References

1. Abramova T (2013) Frontiers and approaches to chemical synthesis of oligodeoxyribonucleotides. Molecules 18:1063-1075
2. Afshinnekoo E, Meydan C, Chowdhury S, et al. (2015) Geospatial resolution of human and bacterial diversity with city-scale metagenomics. Cell Systems 1:1-15
3. Anaya-Velázquez F (2002) Bioethics, bioweapons and the microbiologist. Rev Latinoam Microbiol 44:38-45
4. Annaluru N1, Muller H, Mitchell LA, et al. (2014) Total synthesis of a functional designer eukaryotic chromosome Science 344:55-58
5. Baeshen MN, Al-Hejin A, Bora RS, et al. (2015) Production of Biopharmaceuticals in *Escherichia coli*: Current scenario and future perspectives. J Microbiol Biotechnol 28:953-962
6. Benner SA (2010) Defining life. Astrobiology 10:1021-1030
7. Benson DA, Clark K, Karsch-Mizrachi I, Lipman DJ, Ostell J, Sayers EW (2015) GenBank. Nucleic Acids Res 43:D30-35

8. Cermak T, Doyle EL, Christian M, et al. (2011) Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. Nucleic Acids Res 39(12):e82
9. Cho MK, Magnus D, Caplan AL, McGee D (1999) Policy forum: genetics. Ethical considerations in synthesizing a minimal genome. Science 286:2087-2090
10. Chodasewicz K (2014) Evolution, reproduction and definition of life. Theory Biosci 133:39-45
11. Church GM (2013) Improving genome understanding. Nature 502:143
12. Cleary, MA Kilian K, Wang Y, et al. (2004) Production of complex nucleic acid libraries using highly parallel in situ oligonucleotide synthesis. Nat. Methods 1:241-248
13. Commoner B (1964) Deoxyribonucleic acid and the molecular basis of self-duplication. Nature 203:486-491
14. Datsenko KA, Wanner BL (2000) One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. Proc Natl Acad Sci USA 97:6640-6645
15. D'Auria G, Schneider MV, Moya A (2014) Live genomics for pathogen monitoring in public health. Pathogens 3:93-108
16. Ding Y, Wu F, Tan C (2014) Synthetic biology: A bridge between artificial and natural cells. Life (Basel) 19:1092-1116
17. Dormitzer PR, Suphaphiphat P, Gibson DG, et al (2013) Synthetic generation of influenza vaccine viruses for rapid response to pandemics. Sci Transl Med 15:185ra68
18. Ekblom R, Wolf JB (2014) A field guide to whole-genome sequencing, assembly and annotation. Evol Appl. 7:1026-1042
19. Engler C, Gruetzner R, Kandzia R, Marillonnet S (2013) Golden gate shuffling: a one-pot DNA shuffling method based on type IIs restriction enzymes. PLoS One 4:e5553
20. Ferber D (2004) Synthetic biology. Microbes made to order. Science 9:158-161
21. Field D, Amaral-Zettler L, Cochrane G, et al. (2011) The Genomic Standards Consortium. PLoS Biol 9:e1001088
22. Fleischmann RD, Adams MD, White O, et al. (1995) Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. Science 269:496-512
23. Fraser CM, Gocayne JD, White O, et al. (1995) The minimal gene complement of *Mycoplasma genitalium*. Science 270:397-403
24. Georgiades K (2012) Genomics of epidemic pathogens. Clin Microbiol Infect 18:213-217
25. Gibson DG (2014) Programming biological operating systems: genome design, assembly and activation. Nat Methods 11:521-526
26. Gibson DG, Benders GA, Axelrod KC, et al. (2008) One-step assembly in yeast of 25 overlapping DNA fragments to form a complete synthetic *Mycoplasma genitalium* genome. Proc Natl Acad Sci USA 105:20404-20409
27. Gibson DG, Glass JI, Lartigue C, et al. (2010) Creation of a bacterial cell controlled by a chemically synthesized genome. Science 329:52-56
28. Gibson DG, Young L, Chuang RY, Venter JC, Hutchison 3rd CA, Smith HO (2009) Enzymatic assembly of DNA molecules up to several hundred kilobases. Nat Methods 6:343-345
29. Gilles AF, Averof M. (2014) Functional genetics for all: engineered nucleases, CRISPR and the gene editing revolution. Evodevo 18;5:43
30. Glass JI, Assad-Garcia N, Alperovich N, et al. (2006) Essential genes of a minimal bacterium. Proc Natl Acad Sci USA 103:425-430
31. Goebl MG, Petes TD (1986) Most of the yeast genomic sequences are not essential for cell growth and division. Cell 46:983-992
32. Goodacre R, Vaidyanathan S, Dunn WB, Harrigan GG, Kell DB (2004) Metabolomics by numbers: acquiring and understanding global metabolite data. Trends Biotech 22:245-252
32. Guo YY, Shi ZY, Fu XZ, Chen JC, Wu Q, Chen GQ (2015) A strategy for enhanced circular DNA construction efficiency based on DNA cyclization after microbial transformation. Microb Cell Fact 14:18
33. Haft DH (2015) Using comparative genomics to drive new discoveries in microbiology. Curr Opin Microbiol 23C:189-196
34. Hutchison CA, Peterson SN, Gill SR, Cline RT, White O, Fraser CM, Smith HO, Venter JC (1999) Global transposon mutagenesis and a minimal *Mycoplasma* genome. Science 286:2165-2169
35. Itaya M (1995) An estimation of minimal genome size required for life. FEBS Lett 362:257-260
36. Kang Z, Zhang J, Jin P, Yang S (2015) Directed evolution combined with synthetic biology strategies expedite semi-rational engineering of genes and genomes. Bioengineered 26:1-5
37. King AV, Welter BH, Koushik AB, Gordon LN, Temesvari LA (2012) A genome-wide over-expression screen identifies genes involved in phagocytosis in the human protozoan parasite, *Entamoeba histolytica*. PLoS One 7:e43025
38. Kosuri S, Church GM (2014) Large-scale de novo DNA synthesis: technologies and applications. Nat Methods 11:499-507
39. Kukurba KR, Montgomery SB (2015) RNA sequencing and analysis. Cold Spring Harb Protoc Apr 13. doi: 10.1101/pdb.top084970
40. Lartigue C, Glass JI, Alperovich N, Pieper R, Parmar PP, Hutchison CA 3rd, Smith HO, Venter JC (2007) Genome transplantation in bacteria: changing one species to another. Science 317:632-638
41. Liao HK, Gu Y, Diaz A, et al. (2015) Use of the CRISPR/Cas9 system as an intracellular defense against HIV-1 infection in human cells. Nat Commun 6:6413
42. Ling LL, Schneider T, Peoples AJ, et al. (2015) A new antibiotic kills pathogens without detectable resistance. Nature 517:455-459
43. Liu JK, Chen WH, Ren SX, Zhao GP, Wang J (2014) iBrick: a new standard for iterative assembly of biological parts with homing endonucleases. PLoS One 9:e110852
44. Mandell DJ, Lajoie MJ, Mee MT, Takeuchi R, Kuznetsov G, Norville JE, Gregg CJ, Stoddard BL, Church GM (2015) Biocontainment of genetically modified organisms by synthetic protein design. Nature 518:55-60
45. Margulies M, Egholm M, Altman WE, et al. (2005) Genome sequencing in microfabricated high-density picolitre reactors. Nature 437:376-380
46. McLean JS, Lombardo MJ, Badger JH, et al. (2013) Candidate phylum TM6 genome recovered from a hospital sink biofilm provides genomic insights into this uncultivated phylum. Proc Natl Acad Sci USA 110:E2390-2399
47. Nikel PI, Martínez-García E, de Lorenzo V (2014) Biotechnological domestication of pseudomonads using synthetic biology. Nat Rev Microbiol 12:368-379
48. Padilla-Martínez F, Carrizosa-Villegas LA, Rangel-Serrano Á, Paramo-Pérez I, Mondragón-Jaimes V, Anaya-Velázquez F, Padilla-Vaca F, Franco B (2015) Cell damage detection using *Escherichia coli* reporter plasmids: fluorescent and colorimetric assays. Arch Microbiol 197:815-821
49. Peng D, Kurup SP, Yao PY, Minning TA, Tarleton RL (2014) CRISPR-Cas9-mediated single-gene and gene family disruption in *Trypanosoma cruzi*. MBio 6:e02097-14
50. Ran FA, Cong L, Yan WX, et al. (2015) *In vivo* genome editing using *Staphylococcus aureus* Cas9. Nature 520:186-191
51. Røkke G, Korvald E, Pahr J, Oyås O, Lale R (2014) BioBrick assembly standards and techniques and associated software tools. Methods Mol Biol 1116:1-24
52. Roy S, Caruthers M (2013) Synthesis of DNA/RNA and their analogs via phosphoramidite and H-phosphonate chemistries. Molecules 18:14268-14284
53. Salter SJ, Cox MJ, Turek EM, Calus ST, Cookson WO, Moffatt MF, Turner P, Parkhill J, Loman NJ, Walker AW (2014) Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. BMC Biol 12;12:87
54. Sleight SC, Bartley BA, Lieviant JA, Sauro HM (2010) In-Fusion BioBrick assembly and reengineering. Nucleic Acids Res 38:2624-2636

55. Sleight SC, Sauro HM (2013) BioBrick™ assembly using the In-Fusion PCR Cloning Kit. Methods Mol Biol 1073:19-30
56. Smith HO, Hutchison CA 3rd, Pfannkoch C, Venter JC (2003) Generating a synthetic genome by whole genome assembly: phiX174 bacteriophage from synthetic oligonucleotides. Proc Natl Acad Sci USA 100:15440-15445
57. Sulem P, Helgason H, Oddson A, Stefansson H, Gudjonsson SA, Zink F, Hjartarson E, Sigurdsson GT, Jonasdottir A, Jonasdottir A, Sigurdsson A, Magnusson OT, Kong A, Helgason A, Holm H, Thorsteinsdottir U, Masson G, Gudbjartsson DF, Stefansson K (2015) Identification of a large set of rare complete human knockouts. Nat Genet 47:448-52
58. Suzuki Y, Assad-Garcia N, Kostylev M, et al. (2015) Bacterial genome reduction using the progressive clustering of deletions via yeast sexual cycling. Genome Res 25:435-444
59. Thaiss CA, Zeevi D, Levy M, et al. (2014) Transkingdom control of microbiota diurnal oscillations promotes metabolic homeostasis. Cell 159:514-529
60. Tomita M, Hashimoto K, Takahashi K, et al. (1999) E-CELL: software environment for whole-cell simulation. Bioinformatics 15:72-84
61. Tumpey TM, Basler CF, Aguilar PV, et al. (2005) Characterization of the reconstructed 1918 Spanish influenza pandemic virus. Science 310:77-80
62. Turnbaugh PJ, Ley RE, Hamady M, Fraser-Liggett CM, Knight R, Gordon JI (2007) The human microbiome project. Nature 449:804-810
63. Venter JC (2007) A life decoded: my genome, my life. Viking, New York
64. Wagner JC, Goldfless SJ, Ganesan SM, Lee MC, Fidock DA, Niles JC. (2013) An integrated strategy for efficient vector construction and multigene expression in *Plasmodium falciparum*. Malar J 12:373
65. Wang YH, Wei KY, Smolke CD (2013) Synthetic biology: advancing the design of diverse genetic systems. Annu Rev Chem Biomol Eng 4:69-102
66. Weiss R (2014) Realizing the potential of synthetic biology. Nat Rev Mol Cell Biol 15:289-294
67. West CE, Renz H, Jenmalm MC, Kozyrskyj AL, Allen KJ, Vuillermin P, Prescott SL; in-FLAME Microbiome Interest Group (2015) The gut microbiota and inflammatory noncommunicable diseases: associations and potentials for gut microbiota therapies. J Allergy Clin Immunol 2015 135:3-13
68. Yaung SJ, Church GM, Wang HH (2014) Recent progress in engineering human-associated microbiomes. Methods Mol Biol 1151:3-25
69. Yolken RH, Jones-Brando L, Dunigan DD, et al. (2014) Chlorovirus ATCV-1 is part of the human oropharyngeal virome and is associated with changes in cognitive functions in humans and mice. Proc Natl Acad Sci USA 111:16106-16111
70. Yong HT1, Yamamoto N, Takeuchi R, Hsieh YJ, Conrad TM, Datsenko KA, Nakayashiki T, Wanner BL, Mori H (2013) Development of a system for discovery of genetic interactions for essential genes in *Escherichia coli* K-12. Genes Genet Syst 88:233-240
71. Yooseph S, Andrews-Pfannkoch C, Tenney A, et al. (2013) A metagenomic framework for the study of airborne microbial communities. PLoS One 8:e81862
72. Zautner AE, Masanta WO, Hinz R, Hagen RM, Frickmann H (2015) Artificially designed pathogens–A diagnostic option for future military deployments. Mil Med Res 2:17