Synthetic genomics: a new venture to dissect genome fundamentals and engineer new functions

Daniel Schindler[1], Junbiao Dai[2] and Yizhi Cai[1,2]

Check for
updates

Since the first synthetic gene was synthesized in 1970s, the efficiency and the capacity of made-to-order DNA sequence synthesis has increased by several orders of magnitude. Advances in DNA synthesis and assembly over the past years has resulted in a steep drop in price for custom made DNA. Similar effects were observed in DNA sequencing technologies which underpin DNA-reading projects. Today, synthetic DNA sequences with more than 10 000 bps and turn-around times of a few weeks are commercially available. This enables researchers to perform large-scale projects to write synthetic chromosomes and characterize their functionalities *in vivo*. Synthetic genomics opens up new paradigms to study the genome fundamentals and engineer novel biological functions.

Addresses
[1] Manchester Institute of Biotechnology, University of Manchester, 131 Princess Street, M1 7DN Manchester, UK
[2] Centre for Synthetic Genomics, Institute of Synthetic Biology, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, China

Corresponding author: Cai, Yizhi (yizhi.cai@manchester.ac.uk)

Introduction

One of the major challenges in biological sciences was the determination of DNA sequences. In the beginning, only single DNA fragments were sequenced using the chain termination sequencing technique [1]. However, the Human Genome Project (GP-Read) accelerated the evolution of new sequencing techniques by having the ambitious goal to sequence the human genome within 15 years. The development of Next Generation Sequencing techniques today allows sequencing of a human genome within days. However, most eukaryotic genomes are not fully sequenced and new sequencing techniques are still being developed. As exemplary achievement of this development, in 2017 sequencing of one of the highly repetitive human centromeres was achieved [2•]. Scientists are now performing well in reading genomes, a measurable output being the growing number of genome sequences in public databases. However, reading a book alone does not make a good writer, instead it requires one to start writing extensively and creatively to master the art and ultimately it leads to a better understanding of grammar and expression. In this case, one needs to write synthetic DNA sequences in order to better understand the grammar of life.

Writing DNA starts with short single-stranded fragments: the oligonucleotides. Since the development of the Polymerase Chain Reaction and the first complete synthesis of a gene, writing DNA *in vitro* has progressed impressively (Figure 1) [3,4]. Recent drops in DNA synthesis costs and the improved capability of synthesizing longer stretches of DNA allow the design and construction of whole synthetic chromosomes in the mega-base range. Recent publications report the construction of viral and microbial synthetic genomes, and the Sc2.0 project aims to generate the first synthetic eukaryotic genome. It is an open discussion how to define whether a chromosome or genome is synthetic. In this review, chromosomes and genomes are defined as synthetic when all building blocks of the final DNA molecule are generated by chemical synthesis. Chromosomes and genomes which are not completely synthesized are considered 'engineered' or 'modified' and are outside the scope of this review. We define synthetic genomics to be a new field where biology is being engineered at the genome level, and it is an intersection of synthetic biology and systems biology. This review neither aims to discuss assembly methods nor the dual-use character of synthetic genomics. The authors are fully aware of the potential dual use character, especially for the synthesis of viral genomes. However, these issues are discussed and reviewed extensively elsewhere [5–7].

Design concepts and assembly strategies for synthetic chromosomes

Computer-assisted design software (CADs) have been developed to ensure efficient and consistent design of synthetic DNA sequences at the genome scale [8••,9•]. The design space of synthetic DNA is enormous and many (if not infinite) design blueprints are possible, as long as they can result in the viability of the cell, to achieve the design intention [10]. Initial projects aiming to synthesize a whole genome were conservative in

Figure 1
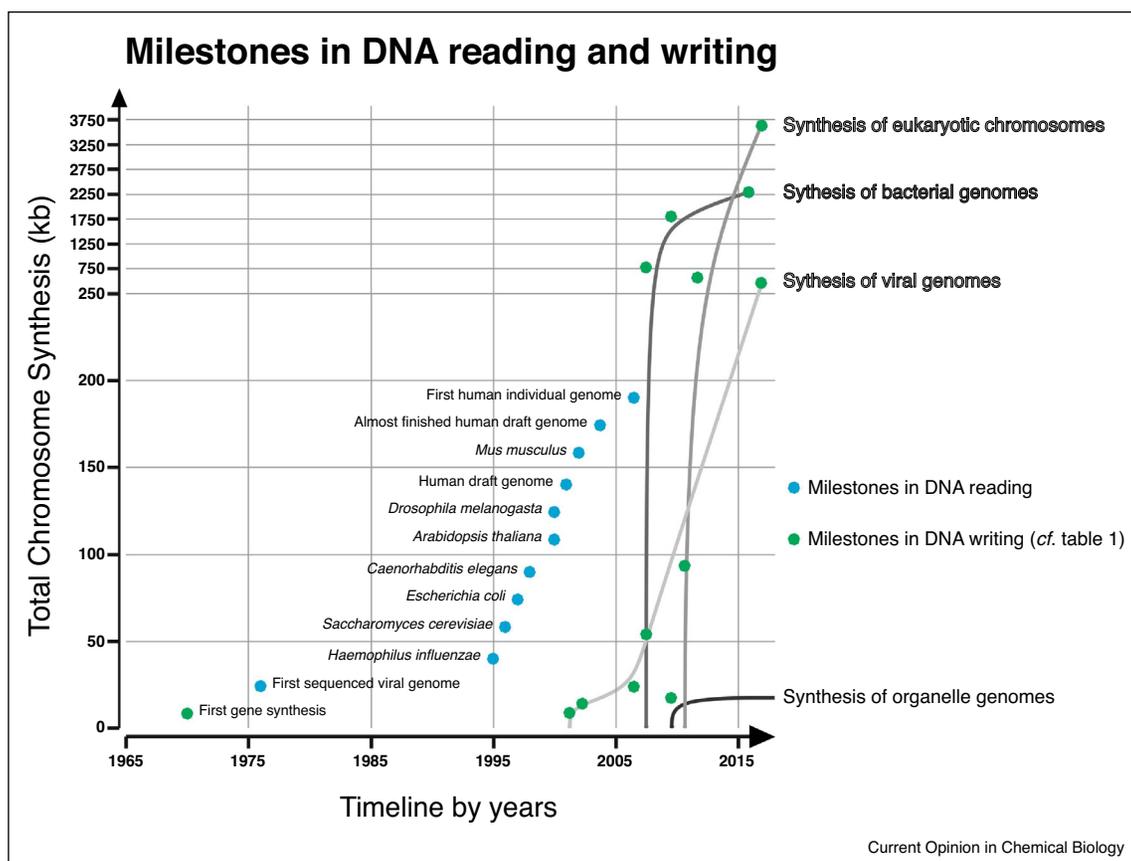
Milestones in DNA reading and writing. Reading and writing technologies depend on technological breakthroughs. Large scale DNA-reading projects (examples in blue) were accomplished after development of Sanger Sequencing, PCR and Next Generation Sequencing. The number of studies utilizing new technologies grows quickly after a developmental lag phase. The number of genome sequences uploaded to databases is exploding and it is impossible to give a number which would be accurate and valid for some time. The knowledge gained by genome sequencing and advantages in gene synthesis is the dawn of writing chromosomes. By now the number of bases incorporated into completely synthetic chromosomes is: 6.1 mb. The cost by today would be roughly $425 000 assuming the current competitive price rate of 7 cents per base for non-clonal 1.8 kb DNA-fragments. The synthesis cost for a haploid human genome would by today be roughly $45 000 000. However, lowering DNA synthesis costs is one of the major goals of GP-write. In the future, the DNA synthesis cost of a human genome in will be less than the price of the *Mycoplasma mycoides* JCVI-syn1.0 project (estimated $40 000 000 [33]).

changes to the genetic content, but nonetheless resulted in the breakthrough in synthesizing, assembling and ultimately transplanting chromosome-scale synthetic DNA [11,12•]. With increasing knowledge and progress in chromosome-scale DNA synthesis, the designs of synthetic sequences are becoming more complex and ambitious [8••,13••]. Many genome synthesis projects utilize a hierarchical genome assembly strategy starting with small building blocks which are assembled, by the technique of choice, to larger building blocks of around 50–100 kb. These fragments are used to further assemble the synthetic chromosome in a heterologous host or to replace the corresponding wild-type sequence in a stepwise manner. Each of the techniques have advantages and disadvantages (Box 1), and should be chosen carefully based on the use cases.

Synthesizing DNA goes viral

Although viruses and phages are not considered to be 'alive' they have a genome. They can reproduce themselves by leveraging the resources from a host. Viral genomes are rather small, with sizes between 1759 bps (*Porcine circovirus* [14]) and 1259 kb (*Megavirus chilensis* [15]) and can consist of DNA or RNA. The first complete synthesis of a viral RNA genome, the polio virus, was accomplished in 2002 [16••]. The 7.5 kb synthesized cDNA genome was *in vitro* transcribed by RNA polymerase and can generate infectious virus particles after transfer into a cell free extract. Further viral RNA and DNA genomes were synthesized up to a size of 212 kb in recent years (Table 1). Synthesizing, as well as engineering variations of viral genomes to produce genome libraries, has an enormous potential for therapeutic applications. Vaccines and drugs could be quickly generated in

---

Box 1 Advantages of heterologous host or native host for chromosome assembly

Advantages of chromosome assembly in heterologous host:
- Assembly in a heterologous host potentially prevents cellular burden to the targeted host.
- Well-established assembly methods are readily available (e.g. homologous recombination in *S. cerevisiae*).

Disadvantages of chromosome assembly in heterologous host:
- The size of synthetic chromosomes might be limited by the capacity of the heterologous host.
- The synthetic chromosome needs to be transferred to the final host.
- Difficult to debug design defects until the synthesized chromosome is transplanted to the target host.

Advantages of step-by-step chromosome replacement:
- Replacing the wildtype sequence gives a real-time fitness monitoring.
- No need to transfer the final chromosome to a different organism.
- Multiplex replacement can be carried out simultaneously for more efficient assembly.

Disadvantages of step-by-step chromosome replacement:
- Some organisms are not amenable to homologous recombination.
- Might be very time consuming if the host is a slow growing organism.
- It might be necessary to develop new tools to manipulate the target organism.

---

response to the emergency of a certain virus variant, which may help to prevent wider outbreaks [7].

Synthesizing genomes of organelles

Mitochondria in general, and the plastids of plants, contain a genome. Their sizes are rather small but show a huge variation in size and content. Studying these organelles is very interesting but challenging. Transformation of organelles must be done by bio-ballistic transformation [17]. The efficiency of synthetic DNA transformation is rather low. Mitochondria are the only organelles for which a complete organelle genome has been synthesized so far. The synthesis of the 16.3 kb mouse mtDNA genome was achieved by using 600 60mer oligonucleotides in four consecutive assembly rounds [18]. This step was predominantly the proof of principle for a DNA assembly method. However, it is an intriguing question why organelles still contain genetic content and have not migrated all necessary genes to the nucleus. There are exceptions in nature where the mitochondria do not contain any DNA [19].

Synthesizing microbial genomes

The genome of *Mycoplasma genitalium* was the first completely chemically synthesized genome [11]. However, the genome could not be transferred to produce a viable synthetic *M. genitalium* strain, presumably because of an interruption of *rnpB*, a subunit of RNaseP. In 2010, the same group demonstrated, in a remarkable work, the complete chemical synthesis and transfer of the 1.08 Mb *Mycoplasma mycoides* genome into the close relative

Mycoplasma capricolum [12•]. This is the first organism which is controlled by a synthetic genome, and is referred as *M. mycoides* JCVI-syn1.0. The genomic differences to *M. mycoides* are marginal and consist of designed 'watermark' sequences, 14 genes are deleted or disrupted and nineteen harmless polymorphisms were acquired during the building process.

This successful project was the starting point to generate a minimal *Mycoplasma* organism based on JCVI-syn1.0. Briefly, two independent teams failed to generate a viable cell, based on knowledge and genome synthesis, from scratch. However, multiple rounds of transposon mutagenesis and genome reduction finally generated *M. mycoides* JCVI-syn3.0 a minimal genome with a genome reduction of 50.8% in a design, build and test cycle manner [20,13••]. The 901 genes of *M. mycoides* JCVI-syn1.0 were reduced to 473 genes of which 149 are of unknown function and will give deeper insights into essentiality of genes [21].

An interesting ongoing project is the generation of a synthetic *Escherichia coli* genome. The genome of a previously diminished *E. coli* strain is redesigned in 87 *ca.* 50 kb segments to eliminate 7 codons in the coding sequence in a stepwise manner [9•]. The 62 214 (5.4%) excluded codons are replaced by synonymous codons to maintain viability. The freed-up codons may be used to incorporate non-natural amino acids into proteins in the future. Absence of seven codons and corresponding tRNAs will, in addition, provide sufficient resistance to phages, rendering this strain of great general interest. Currently 55 of the 87 segments have been tested experimentally but the incorporation into a fully synthetic *E. coli* genome still needs to be proven functional.

Synthesizing eukaryotic genomes

As of today, there is no complete synthetic eukaryotic genome. However, the synthetic yeast genome — Sc2.0 project (www.syntheticyeast.org) aims to generate the first eukaryotic cell operated by a synthetic genome. The 16 chromosomes are synthesized in individual strains by teams of scientists within the Sc2.0 Consortium. The chromosomes are re-designed in a higher order of magnitude compared to any other existing write project [8••]. The major changes include the removal of most introns, transposons and repetitive elements. One central element of Sc2.0 design is the relocation of all tRNA genes to an independent 17th *Saccharomyces cerevisiae* chromosome, designated as the tRNA neochromosome. tRNA genes are heavily transcribed and therefore are hotspots of genomic instability caused by replication stress and transposon insertions. In addition, all non-essential genes are flanked by loxPsym sites which allow inducible large-scale genomic re-arrangements mediated by Cre-recombinase. This implemented genome rearrangement technique is therefore referred to as Synthetic Chromosome

Table 1

Overview of finished synthesized chromosomes and genomes

Year	Species	Size in kb (% of wt genome)	Highlights of the study	Reference
Synthetic viral genomes				
2002	Polio virus	7.5	First infectious viral particles, based on a fully synthetic genome.	[16[••]]
2003	φX174 bacteriophage	5.4	Whole workflow: design, oligo synthesis, genome assembly and generation of infectious viral particles in fourteen days.	[34]
2007	Human endogenous retrovirus (HERV-K) consensus	9.5	Generation of a consensus genome of a human endogenous retrovirus which can replicate and is infectious.	[35]
2008	Bat severe acute respiratory syndrome (SARS)-like coronavirus (Bat-SCoV)	29.7	Until 2010 the largest synthetic replicating life form.	[36]
2017	Horsepox	212	First *de novo* synthesis of an in nature extinct orthopoxvirus of which close relatives causes smallpox in humans.	[7]
Synthetic bacterial genomes				
2008	*Mycoplasma genitalium*	583	Watermark sequences were inserted to identify synthetic DNA; disruption of one gene (MG408) to prevent pathogenicity and to do antibiotic selection.	[11]
2010	*Mycoplasma mycoides* JCVI-syn1.0	1079	Watermark sequences were inserted to identify synthetic DNA; deletion or disruption of fourteen genes (one accidentally mediated by an IS1 *E. coli* transposon).	[12[•]]
2016	*Mycoplasma mycoides* JCVI-syn3.0	531.6 (49.3%)	Reduction of *Mycoplasma mycoides* JCVI-syn1.0 in a design, build, test manner based on transposon mutagenesis to reduce the genome from 901 to 473 genes.	[13[••]]
Synthetic eukaryotic organelle genomes				
2010	*Mus musculus* mtDNA	16.3	Complete *in vitro* synthesis in four steps to produce a copy of the mouse mtDNA consensus genome.	[18]
Synthetic eukaryotic chromosomes				
2011	*Saccharomyces cerevisiae* synIXR	91.0 (101.9%)[a]	Applied design rules (if possible): • Synonymous replacement of TAG stop codons with TAA	[20]
2012	*Saccharomyces cerevisiae* synIII	272.2 (86.2%)	• Synonymous recoding of tandem repeats in CDSs	[24[••]]
2017	*Saccharomyces cerevisiae* synII	770.1 (94.7%)	• Deletion of genomic instability causing elements: introns,	[25]
2017	*Saccharomyces cerevisiae* synV	536.0 (92.9%)	transposons, subtelomeric repeats and tRNA genes	[28]
2017	*Saccharomyces cerevisiae* synVI	242.7 (89.9%)	• Insertion of loxPsym sites to 'SCRaMbLE' the genome	[26]
2017	*Saccharomyces cerevisiae* synX	707.5 (94.9%)	• Insertion of PCRTag sequences to identify synthetic DNA	[29]
2017	*Saccharomyces cerevisiae* synXII	999.4 (92.7%)[b]		[27]

[a] The size of synIXR is slightly longer due to the insertion of 43 loxPsym sites (34 bps each).
[b] Chromosome XII contains a cluster with >100 copies of the 9.1 kb rDNA operon which is not included in the shown size.

Rearrangement and Modification by LoxP-mediated Evolution (SCRaMbLE) and has already proven its functionality [20,22[•],23[•]].

Recent publications report the synthesis and characterization of six Sc2.0 chromosomes and the right arm of synthetic chromosome IX (Table 1) which collectively correspond to 32% of the yeast genome [20,24[••],25–29]. Strikingly, the individually synthesized chromosomes can be merged in a single cell by mating with a technique called endoreduplication intercross [8[••]]. Currently, the strain with the most synthetic chromosomes in one cell contains synIII, synVI and synIXR. With further progression of the Sc2.0 project more synthetic chromosomes will be finalised and ultimately merged to the final Sc2.0 strain.

GP-write: a sneak preview into the future of synthetic genomics

The successes in current genome synthesizing projects are leading to the next grand challenge in modern biological science: The Genome Project-write (GP-write). This project is a grand challenge using synthesis, gene editing and other technologies to understand, engineer and test living systems with the overarching goal to understand the blueprint for life provided by the Human Genome Project (HGP-read) [30[••],31,32]. Therefore, a new international consortium was formed and first meetings were held in 2016 and 2017. The consortium is an open, interdisciplinary and international research group to focus efforts to realize GP-write.

GP-write has several goals, one being the development of new techniques and to accelerate the evolution of existing techniques with an overall goal to reduce synthesis costs by 1000-fold within ten years. Similar effects were achieved by HGP-read: today the cost of sequencing a human genome are magnitudes lower than the initial human genome sequence. The open nature of GP-write allows everyone to submit project proposals which will be evaluated by the Scientific Executive Committee. As of

Table 2

Approved GP-write pilot projects (by January 2018)

Project title	Project goals[a]	Project lead(s)
UltraSafe Cell Line	The project aims to generate an Ultrasafe cell line by altering roughly 1% of the human genome. Some key goals are: Virus and prion resistance, removal of transposable elements, recoding of triplet repeats, recoding to a human consensus sequence in regard to SNPs and indels, implementing the bespoke SCRaMbLE system beside further alterations.	Jef Boeke & George Church
High-throughput HAC Design to Test Connections Between Gene Expression, Location and Conformation	There is still a lack of understanding of the regulation of gene expression. The project will build two 1 mb regions of the human genome. The regions will be constructed as combinatorial libraries with different promoters and insulators to investigate 'rules' for optimal gene expression.	Pamela Silver & Jeffrey Way
Safety and Containment; Chromatin and Chromosome Structure	This project tries to solve two questions based on Sc2.0 strains and implement the results into GP-write. Firstly: how can hybridization with natural occurring strains or organisms be excluded? Second: what shapes the contact maps of chromosomes? The latter will be answered by analysing multiple SCRaMbLed Sc2.0 strains to investigate chromatin structure based on re-arranged chromosomes	Jasper Rine
Empirically Designing Genomically Recoded Human Cell Lines	Codon alteration is an important part of GP-write. The project aims to develop: Firstly, a rapid method for multiplex targeted genome modification; secondly, a respective rapid and robust screening system for living cells in 96-well format; thirdly, a strategy for rapid evaluation of heterogenic cell populations; and finally, a software to design the synthetic DNA fragments and evaluate viability of codon replacements.	Marc Lajoie
The Seven Signals Toolbox: Leveraging Synthetic Biology to Define the Logic of Stem-Cell Programming	Cell differentiation is mainly driven by seven signal types. This project aims to generate a toolbox which allows the *in vitro* differentiation of GP-write cell lines. This is a crucial step for future applications in the field of cell therapies, tissue replacement or transplantation of organs.	Liam Holt
Precision Human Genome Engineering of Disease-Associated Noncoding Variants	Efficient and precise engineering of the human genome is still a challenge. This project aims to create a complete pipeline for rapid engineering of human cells with an enrichment for homologous recombination repair. The project will also provide bioinformatic tools to optimize CRISPR based engineering.	Neville Sanjana
Synthesizing a Prototrophic Human Genome	This project postulates to introduce pathways for the nine amino acids and a variety of vitamins which cannot be synthesized by humans. These molecules derive from the diet. It investigates whether the milieu in the cell makes a prototrophic cell line feasible. If the project succeeds, further engineering would be performed and the first achievement would be a drastic cost reduction of cell line cultivation media.	Harris Wang
Through the Looking Glass: Anticipating and Understanding Governance Systems and the Public's Views on HGP-write	Including the publics view and governance systems into GP-write is an important step. This project will generate a dialogue between scientists and the public. Incorporation of the society will enable acceptance and support for GP-write.	Todd Kuiken & Gigi Gronvall
Synthetic Screening for Essential Introns and Retroelements in Human Cell and Animals	This project aims to perform systematic screenings of intron and retroelements in the genome. Combinatorial variants of chosen genes will be investigated in a diploid background. The outcome will indicate if the removal of these elements, like in Sc2.0, is feasible in GP-write.	Yasunori Aizawa
Isothermal Amplification Array	This project proposes a new method to synthesize DNA. It depends on two steps. Firstly: generation of short oligonucleotides by an isothermal amplification on an array. Secondly: the amplified oligonucleotides can anneal according to their design and nicks are sealed by a ligase.	Max Berry
Recombinase-Mediated Assembly	This project proposes a new method to assemble DNA fragments by utilizing a RecA-like recombinase (UvsX). The method should allow, with the Isothermal Amplification Array assembly from short oligos to chromosome-sized DNA, with a significant labour reduction.	Max Berry
Synthetic Regulatory Genomics	The project aims to study regulatory variations of non-coding regions. The project will use multi-edited regulatory DNA sequences and analyse their function with multiple techniques. This project will give deeper insights into non-coding regions of complex genomes.	Matt Maurano
Concepts & Ethics in GP-write: Understand, Question, Advance	This project aims to build a model for deep analysis of concepts and ethics in GP-write, and aims include the dynamics of science and society. It aims to expand collaborations between sciences and the humanities and provide proper education and training.	Jeantine Lunshof

[a] More detailed information can be found at http://www.engineeringbiologycenter.org/.

January 2018 there are 13 pilot projects approved (Table 2). The projects cover many aspects of synthetic genomics, two highlights are the projects dedicated to the Concepts and Ethics in GP-write as well as Anticipating and Understanding Governance Systems and the Publics Views on HGP-write, which shows the importance to consider ethics and the publics views within GP-write.

One major remaining question is: What can we learn from GP-write? On one hand, there will be the *ad hoc* advances in enabling technologies, on the other hand there will be an immense gain of knowledge in biological sciences. Our knowledge of complex genomes is still limited. For instance, roughly 1% of the genome is responsible for all proteins in the cell. The remaining 99% are often referred as the 'dark matter' of the genome. Stepwise replacement of these elements, like in the Sc2.0 project, will potentially help us decipher the functions of the dark matter in the genome. On the application front, the pilot project to engineer a stable and safe cell line, has a profound implication for biomanufacturing and bioproduction (Table 2). GP-write still has a long way to go. However, the scientific community is curious about the outcome of the first pilot projects in the GP-write framework.

Conclusion
The initial genome writing projects summarized here show that individual native chromosomes and whole genomes can be replaced by chemically synthesized genomes. So far, the changes to DNA sequences are relatively modest but with growing knowledge of biological systems, the design will become more aggressive and adventurous which will lead us into previously unexplored territories. The exciting field of synthetic genomics will give new insights in basic research and will open new possibilities in applied science.

Funding

Conflict of interest statement
Nothing declared.

Acknowledgements

References and recommended reading
Papers of particular interest, published within the period of review, have been highlighted as:

• of special interest
•• of outstanding interest

1. Sanger F, Nicklen S, Coulson AR: **DNA sequencing with chain-terminating inhibitors**. *Proc Natl Acad Sci U S A* 1977, **74**:5463-5467.

2. Jain M, Olsen HE, Turner DJ, Stoddart D, Bulazel KV, Paten B,
• Haussler D, Willard H, Akeson M, Miga KH: **Linear assembly of a human Y centromere using nanopore long reads**. *bioRxiv* 2017.
The human genome reference still contain gaps, especially in highly repetitive regions. This study shows how highly repetitive regions can be analysed using a targeted aproach and a new generation of sequencing devices.

3. Agarwal KL, Buchi H, Caruthers MH, Gupta N, Khorana HG, Kleppe K, Kumar A, Ohtsuka E, Rajbhandary UL, Van de Sande JH *et al.*: **Total synthesis of the gene for an alanine transfer ribonucleic acid from yeast**. *Nature* 1970, **227**:27-34.

4. Saiki RK, Scharf S, Faloona F, Mullis KB, Horn GT, Erlich HA, Arnheim N: **Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia**. *Science* 1985, **230**:1350-1354.

5. Oldfield LM, Grzesik P, Voorhies AA, Alperovich N, MacMath D, Najera CD, Chandra DS, Prasad S, Noskov VN, Montague MG *et al.*: **Genome-wide engineering of an infectious clone of herpes simplex virus type 1 using synthetic genomics assembly methods**. *Proc Natl Acad Sci U S A* 2017, **114**:E8885-E8894.

6. Mitchell LA, Ellis T: **Synthetic genome engineering gets infectious**. *Proc Natl Acad Sci U S A* 2017, **114**:11006-11008.

7. Koblentz GD: **The de novo synthesis of horsepox virus: implications for biosecurity and recommendations for preventing the reemergence of smallpox**. *Health Secur* 2017, **15**:620-628.

8. Richardson SM, Mitchell LA, Stracquadanio G, Yang K,
•• Dymond JS, DiCarlo JE, Lee D, Huang CL, Chandrasegaran S, Cai Y *et al.*: **Design of a synthetic yeast genome**. *Science* 2017, **355**:1040-1044.
Summary of the design and principle of the Sc2.0 project including how to construct the final Sc2.0 strain with all synthetic chromosomes merged in one cell.

9. Ostrov N, Landon M, Guell M, Kuznetsov G, Teramoto J,
• Cervantes N, Zhou M, Singh K, Napolitano MG, Moosburner M *et al.*: **Design, synthesis, and testing toward a 57-codon genome**. *Science* 2016, **353**:819-822.
Ongoing study to generate a re-coded *E. coli* genome which has the greatest degree of re-coding of coding sequence so far. One result of this project will be a phage resistant *E. coli* strain.

10. Schindler D, Waldminghaus T: **Synthetic chromosomes**. *FEMS Microbiol Rev* 2015, **39**:871-891.

11. Gibson DG, Benders GA, Andrews-Pfannkoch C, Denisova EA, Baden-Tillson H, Zaveri J, Stockwell TB, Brownley A, Thomas DW, Algire MA *et al.*: **Complete chemical synthesis, assembly, and cloning of a *Mycoplasma genitalium* genome**. *Science* 2008, **319**:1215-1220.

12. Gibson DG, Glass JI, Lartigue C, Noskov VN, Chuang RY,
• Algire MA, Benders GA, Montague MG, Ma L, Moodie MM *et al.*: **Creation of a bacterial cell controlled by a chemically synthesized genome**. *Science* 2010, **329**:52-56.
The first chemically synthesized genome which can control a cell. This study is the dawn of synthetic genomics, it shows feasability of synthetic genomes.

13. Hutchison CA 3rd, Chuang RY, Noskov VN, Assad-Garcia N,
•• Deerinck TJ, Ellisman MH, Gill J, Kannan K, Karas BJ, Ma L *et al.*: **Design and synthesis of a minimal bacterial genome**. *Science* 2016, **351** aad6253.
Reduction of the first synthetic bacterial genome by 49.3% in multiple rounds with classification of the remaining genes into essential and quasi-essential (necessary for growth) genes.

14. Finsterbusch T, Mankertz A: **Porcine circoviruses — small but powerful**. *Virus Res* 2009, **143**:177-183.

15. Arslan D, Legendre M, Seltzer V, Abergel C, Claverie JM: **Distant Mimivirus relative with a larger genome highlights the fundamental features of Megaviridae**. *Proc Natl Acad Sci U S A* 2011, **108**:17486-17491.

16. Cello J, Paul AV, Wimmer E: **Chemical synthesis of poliovirus**
•• **cDNA: generation of infectious virus in the absence of natural template**. *Science* 2002, **297**:1016-1018.
The first chemically synthesized and assembled genome which is able to produce infectious viral particles.

17. Klein RM, Wolf ED, Wu R, Sanford JC: **High-velocity microprojectiles for delivering nucleic acids into living cells**. *Biotechnology* 1992, **24**:384-386.

18. Gibson DG, Smith HO, Hutchison CA 3rd, Venter JC, Merryman C: **Chemical synthesis of the mouse mitochondrial genome**. *Nat Methods* 2010, **7**:901-903.

19. Henriquez FL, Richards TA, Roberts F, McLeod R, Roberts CW: **The unusual mitochondrial compartment of *Cryptosporidium parvum***. *Trends Parasitol* 2005, **21**:68-74.

20. Dymond JS, Richardson SM, Coombes CE, Babatz T, Muller H, Annaluru N, Blake WJ, Schwerzmann JW, Dai J, Lindstrom DL *et al.*: **Synthetic chromosome arms function in yeast and generate phenotypic diversity by design**. *Nature* 2011, **477**:471-476.

21. Danchin A, Fang G: **Unknown unknowns: essential genes in quest for function**. *Microb Biotechnol* 2016, **9**:530-540.

22. Shen Y, Stracquadanio G, Wang Y, Yang K, Mitchell LA, Xue Y,
• Cai Y, Chen T, Dymond JS, Kang K *et al.*: **SCRaMbLE generates designed combinatorial stochastic diversity in synthetic chromosomes**. *Genome Res* 2016, **26**:36-49.
The first study which shows a great diversification in synthetic chromosomes by SCRaMbLE. It highlights the need for better techniques to resolve the structure of the re-arranged chromosomes as some SCRaMbLEd chromosomes could not be resolved by short sequencing reads.

23. Mercy G, Mozziconacci J, Scolari VF, Yang K, Zhao G, Thierry A,
• Luo Y, Mitchell LA, Shen M, Shen Y *et al.*: **3D organization of synthetic and scrambled chromosomes**. *Science* 2017, **355**.
This study indicates that the re-design of the analyzed synthetic yeast chromosomes has no or very little effect on the 3D organization within the nucleus.

24. Annaluru N, Muller H, Mitchell LA, Ramalingam S,
•• Stracquadanio G, Richardson SM, Dymond JS, Kuang Z, Scheifele LZ, Cooper EM *et al.*: **Total synthesis of a functional designer eukaryotic chromosome**. *Science* 2014, **344**:55-58.

The first synthetic eukaryotic designer chromosome which produces viable cells. The study proves synthetic chromosomes are not restricted to viral and bacterial genomes.

25. Shen Y, Wang Y, Chen T, Gao F, Gong J, Abramczyk D, Walker R, Zhao H, Chen S, Liu W *et al.*: **Deep functional analysis of synII, a 770-kilobase synthetic yeast chromosome**. *Science* 2017, **355**.

26. Mitchell LA, Wang A, Stracquadanio G, Kuang Z, Wang X, Yang K, Richardson S, Martin JA, Zhao Y, Walker R *et al.*: **Synthesis, debugging, and effects of synthetic chromosome consolidation: synVI and beyond**. *Science* 2017, **355**.

27. Zhang W, Zhao G, Luo Z, Lin Y, Wang L, Guo Y, Wang A, Jiang S, Jiang Q, Gong J *et al.*: **Engineering the ribosomal DNA in a megabase synthetic chromosome**. *Science* 2017, **355**.

28. Xie ZX, Li BZ, Mitchell LA, Wu Y, Qi X, Jin Z, Jia B, Wang X, Zeng BX, Liu HM *et al.*: **"Perfect" designer chromosome V and behavior of a ring derivative**. *Science* 2017, **355**.

29. Wu Y, Li BZ, Zhao M, Mitchell LA, Xie ZX, Lin QH, Wang X, Xiao WH, Wang Y, Zhou X *et al.*: **Bug mapping and fitness testing of chemically synthesized chromosome X**. *Science* 2017, **355**.

30. Boeke JD, Church G, Hessel A, Kelley NJ, Arkin A, Cai Y,
•• Carlson R, Chakravarti A, Cornish VW, Holt L *et al.*: **GENOME ENGINEERING. The Genome Project-Write**. *Science* 2016, **353**:126-127.
Publication which launch the GP-write and opens the discussion for the grand challenge to synthesize complex eukaryotic genomes.

31. Boeke JD, Church G, Hessel A, Kelley NC: *The GP-write Consortium: Genome Project-write: A Grand Challenge. Using Synthesis, Gene Editing and Other Technologies to Understand*. Engineer and Test Living Systems. White Paper; 2016.

32. http://www.engineeringbiologycenter.org/.

33. Pennisi E: **Genomics. Synthetic genome brings new life to bacterium**. *Science* 2010, **328**:958-959.

34. Smith HO, Hutchison CA 3rd, Pfannkoch C, Venter JC: **Generating a synthetic genome by whole genome assembly: phiX174 bacteriophage from synthetic oligonucleotides**. *Proc Natl Acad Sci U S A* 2003, **100**:15440-15445.

35. Lee YN, Bieniasz PD: **Reconstitution of an infectious human endogenous retrovirus**. *PLoS Pathog* 2007, **3**:e10.

36. Becker MM, Graham RL, Donaldson EF, Rockx B, Sims AC, Sheahan T, Pickles RJ, Corti D, Johnston RE, Baric RS *et al.*: **Synthetic recombinant bat SARS-like coronavirus is infectious in cultured cells and in mice**. *Proc Natl Acad Sci U S A* 2008, **105**:19944-19949.