Horizon 2020
European Union Funding
for Research & Innovation

D 2.1

Title: Establishment of SB chassis of reference

Authors: Víctor de Lorenzo

Poject Acronym	**BioRoboost**
Project Title	Fostering Synthetic Biology standardisation through international collaboration
Grant Agreement	820699
Funding Scheme	Coordination Support Actin (CSA)
Call	H2020-NMBP-BIO-CSA-2018
Project duration	36 months: October 2018-September 2021
Website	www.standardsynbio.eu

DOCUMENT SHEET	
WP	2
DELIVERABLE 2.1	Establishment of SB chassis of reference
Due date	September 2019
Version 1	September 2019
Version 2	July 2020
Authors	Victor de Lorenzo
Contributors	Victor de Lorenzo
Dissemination level	Public
Type of Deliverable	Report

EXECUTIVE SUMMARY

The definition of chassis is central to contemporary Synthetic Biology. In this report—which results from multiple discussions and documents from participants of WP2— we propose sound criteria for identifying and consolidating designations of given bacterial strains as synthetic biology platforms, the requirements that they need to fulfil to become certified chassis and the necessity to develop specific methodologies for ecological risk assessment (ERA) in the context of their use in industrial and environmental settings. Much of this analysis was made on the background of ongoing discussions and consultations within the European Food Safety Authority (EFSA, http://www.efsa.europa.eu/) for regulating the use of Synthetic Biology agents within the EU. Key concepts include [i] the need to limit the number of acceptable chassis, [ii] the separation of issues connected to the chassis from those related to genetic implants in order to simplify regulatory procedures, [iii] the demand of complete traceability of agents destined for industrial or environmental utilization, [iii] the necessity to develop suitable, standardized *metrics* for performance and safety and [v] the urgency to update and upgrade traditional ERA methods to the fast development of Synthetic Biology and the ensuing generation of more and more deeply engineered agents.

Introduction: towards a general definition of *chassis*

The notion of biological chassis is central to contemporary Synthetic Biology. The term chassis comes from classical engineering as the frame of an artificial object that supports plug-able physical components for construction and reuse. In SB, the concept of biological chassis refers to an organism capable of harbouring and supporting synthetic genetic components through its natural or modified molecular machinery, such as transcriptional and translational systems. *E. coli* is the most commonly used cell host, due to the existence of a large number of well-established genetic engineering protocols and tools. Over the years, complex genetic circuitry has been implemented not only in *E. coli* but also in many other microorganisms for generating novel biological functions, thus asking for a more precise definitions e.g.

A SynBio chassis is an engineerable and reusable biological platform[1] with a genome encoding a number of basic functions for stable self-maintenance, growth and optimal operation but with tasks and signal processing components growingly edited for strengthening performance under pre-specified environmental conditions[2]

or

A naturally-derived or highly engineered organism repurposed to build, maintain and amplify the components necessary for the deployment of synthetic biological systems/applications

or

A well-understood organism that has been engineered to simplify genetic/metabolic interventions and to reduce their adverse effects

or

A chassis is an evolutionarily robust, well characterized strain for which ample and practical engineering tools/resources exist, that can serve as an ideal platform to deploy engineered circuits with minimal surprise interactions with host functions, thus equipped with both natural and engineered features suited for facilitating optimal performance for a specific setting, and requiring the least engineering steps to achieve design goal.

The quest for the optimal chassis has been addressed from various perspectives (de Lorenzo, 2011; Beites and Mendes, 2015; Nikel et al., 2016; de Lorenzo and Schmidt, 2018; Nora et al., 2019). In one case, the idea is to start with a well characterized bacterium (e.g. *Escherichia coli*) and then delete the bits and parts of the genome that are not necessary for growth in a given environmental context.

The roadmap from being a rDNA host to a fully-fledged certified SynBio chassis

Fig. 1 below summarizes the itinerary proposed in terms of information and modifications needed for upgrading a promising environmental isolate to a fully-fledged standardized SynBio chassis. This roadmap recapitulates and expands earlier proposals in the same direction. Any (preferably non-pathogenic) environmental isolate able to capture (through transformation or conjugation) and stably maintain exogenous DNA and for which a minimum of genetic tools is available, can be

[1] In general, a *chassis* will be a given strain derived from a naturally-occurring (micro)organism.
[2] Note that the key here is optimal *performance*, not minimized genome size (although deletion of innecessary functions will cause a degree of genome reduction)

tagged in principle as a recombinant DNA (rDNA) host. The historical example of this category is *Escherichia coli,* but now there are dozens of species amenable to a suite of genetic manipulations, including pathogens that are handled under controlled laboratory conditions. But to become a true chassis, the biological host should be agreeable to and optimized for accommodating complex genetic devices and deploying their encoded properties under specified operational conditions.

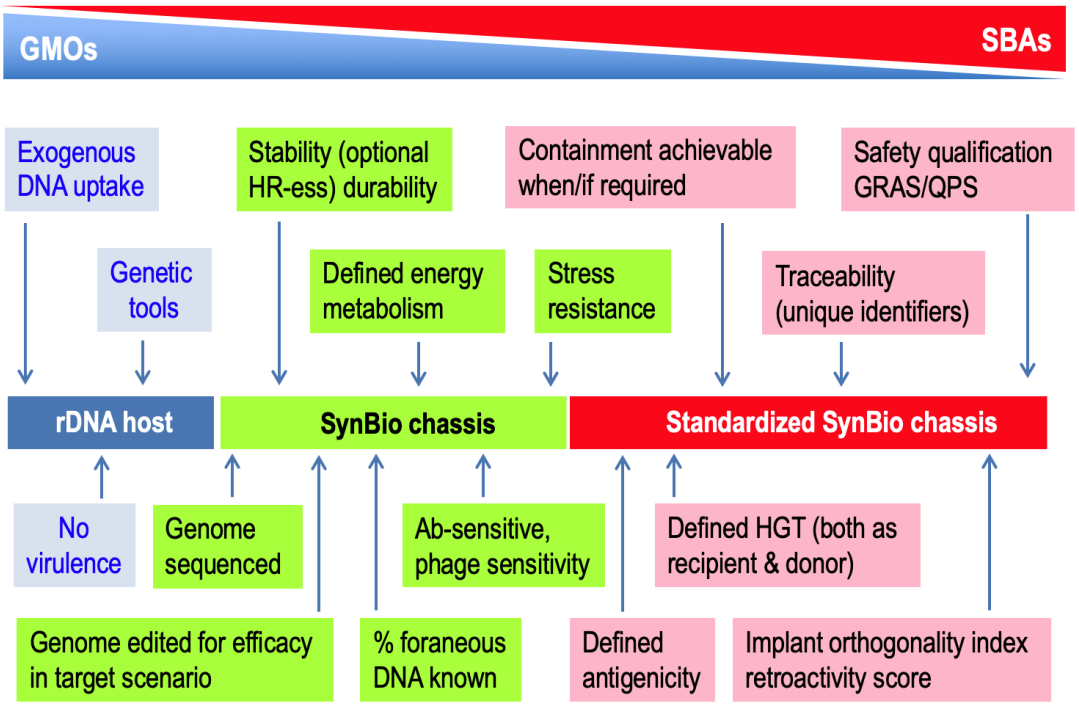


FIG. 1. The roadmap from environmental isolates to fully-fledged standardized SynBio chassis and from GMO (genetically modified organism) to SBA (SynBio agent). The scheme indicates the nature of the information that should be available for each category of strains. Note that there is not a defined boundary between GMOs and SBAs (see proposed definitions in Table 3). One important aspect of standardized chassis is their digital twinning that can be implemented through DNA barcoding as explained in the text. The final product of the process should be an effective and ERA-acceptable host of SynBio devices—or in general rDNA constructs.

For this, additional requirements are needed: the complete genetic complement should be known and advanced genetic tools for deep editing be at hand. This should result in a profound knowledge of the energy metabolism (typically through reliable metabolic models), stress resistance and sensitivity to antibiotics and phages. Knowing the ratio synthetic/engineered DNA vs. natural genetic complement is straightforward in these cases. Furthermore, genetic and evolutionary stability of the resulting constructs is a most desirable trait. This could be enhanced by engineering circuits that somehow punish mutations in the genetic implants or by making cells deficient on endogenous recombination systems. This, in turn, asks for specific genome editing methods that do not rely on recombination, such as targetrons or base editors. Up to that point, one can consider a large number of species and strains that can qualify as or become SynBio chassis (see e.g. Table below). Things get more restrictive, however, when strains are destined for actual, large-scale biotechnological applications, as they must meet additional specifications that are not that important in the laboratory or in academic settings. Most of them deal with safety and efficacy issues, which need to be addressed for overcoming environmental risk assessment (ERA) criteria

and gain a green light by regulatory agencies. Properties of interest to this end include antigenicity and horizontal gene transfer (HGT) —either as donors or recipients of DNA. For some specific applications, containment of the strains themselves or at least barriers to HGT to/from them are necessary while in others propagation of beneficial traits to the surrounding natural community might be desirable depending on the goal.

The extant genomes of microorganisms are populated with a large number of DNA sequences that, on a first sight, are dispensable and even deleterious for the final application of the bacterium. For the time being, some of these minimized *E. coli* strains are the best available chassis for the implantation of new genetic circuits. Note, however, that genomic reduction is not only accompanied by a growing dependence on the external milieu for survival, but also by the loss of antigens that may render the synthetic agent invisible to the immune system. This creates unexpected risks that are difficult to determine upfront. Note also that these definitions largely deal with optimal chassis per specific target environment and tasks therein. However, the concept implies that there may be not ultimate versions of them, but growingly upgraded variants (like improved operative systems in computers or improved iPhones with better/more utilities), an issue that intersects with the question of *barcoding* discussed below.

A provisional list of strains and species amenable to be developed as chassis is shown below (others will undoubtedly follow as the field progresses)

Genus / species	Qualities of interest	References
Mycoplasma sp.	Small genome, vehicle for delivering therapeutic activities to the lung	(Hutchison et al., 2016)
Escherichia coli	Laboratory work horse, recombinant DNA host, abundant genetic tools, the best known living organism	(Lee and Kim, 2015)
Pseudomonas putida	Tolerance to environmental insults (solvents, redox stress), platform for metabolic engineering	(Poblete-Castro et al., 2012; Nikel et al., 2016)
Bacillus subtillis	Laboratory workhorse, easy recombineering, efficient secretion systems	(van Dijl and Hecker, 2013)
Corynebacterium sp.	Long time applications in industrial biotechnology, large-scale production of amino acids	(Becker and Wittmann, 2012; Heider and Wendisch, 2015)
Saccharomyces cerevisiae	Laboratory workhorse, easy genetic manipulations, optimal eukaryotic metabolic engineering platform	(Heider and Wendisch, 2015)

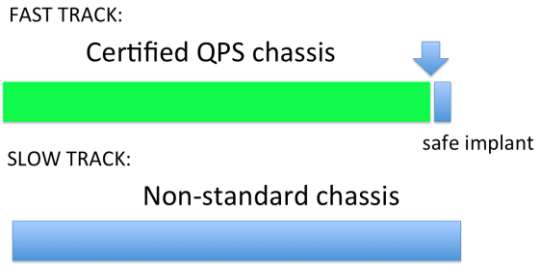
Synechocystis/Synechococcus	Photosynthetic organisms, CO2 fixation, emerging metabolic engineering	(Pinto et al., 2012; Yu et al., 2015)
Streptomyces sp.	Diverse secondary metabolism, production of antibiotics, efficient secretion systems	(Komatsu et al., 2010)
Vibrio natriegens	Super-rapid growth, easy to engineer, host of recombinant DNA constructs.	(Lee et al., 2016; Weinstock et al., 2016)
Halomonas sp	Growth in non-sterile seawater, appealing platform for large-scale processes	(Tan et al., 2011)

Criteria for definition of safe and reliable chassis

It is not realistic to have an infinite number of chassis. Instead, it will be more practical from an ERA and an application point of view to define a number of them (e.g. about 20 max) that are thoroughly characterized and given a certain safety score. This would then limit ERA of specific agents (i.e. chassis + implants) to the effects of the new genes and their *retroactivity* with the genetic and physiological network of the host (see below). ERA issues of the different chassis must first gather information on [i] Genomic sequence and its resilience/stability over time, [ii] Efficacy in the target scenario, [iii] Genetic/genomic stability (e.g. homologous recombination, insertion sequences) and durability, [iv] Sensitivity to phages and antibodies as preventive tool for emergency clearance, [v] Availability of advanced genetic tools, [vi] Traceability, preferably engineered through genomic barcoding (see below), [vii] Antigenicity, [viii] Energy metabolism, [ix] Stress resistance and [x] Horizontal gene transfer capability (both as a donor and as a recipient). All this should result in a chassis-specific safety ranking, similar to qualified presumption of safety (QPS: https://www.efsa.europa.eu/en/topics/topic/qualified-presumption-safety-qps) but associated to the specific platform, not to a whole given species.

Safe implementation of SynBio

Any SynBio **agent** is thus the result of combining a live **chassis** with one or more genetic **implants**. Emphasis on safety should thus be focused on the final agents and address the following questions: [i] Can SynBio agents colonize and eventually takeover natural microbial communities? [ii] Is there a chance that SynBio agents enter new niches that natural bacteria cannot? [iii] Might SynBio agents go into a stage of uncontrolled growth [iv] What are the chances of horizontal transfer of the synthetic genes to novel recipients? [v] Is there a tradeoff between safety and biotechnological efficacy of SynBio agents? [vi] Could traits engineered in SynBio agents evolve towards virulence or other deleterious behavior? [vii] Are there scenarios of SynBio agents capable of damaging life or property? [viii] What is the environmental fate of synthetic genes? [ix] Are there chances of malicious misuse of SynBio agents [x] Should SynBio agents be endowed with traits to increase their safety and predictability? Obviously, if we separate the questions in those applied to a preset chassis and those to the new implants we would considerably ease and accelerate the ERA process and the granting of permissions:



In that case, the questions would be limited to determining the retroactivity score/orthogonality index (i.e. the burden caused by the implant on the host physiology), genetic stability (e.g. plasmid *vs* chromosomal implant) and inspection of emergent properties. This is expected to make ERA studies much simpler than those currently undergone by GE agents.

The need of ERA standards and SOPs

While many ERA methodologies are appropriate for assessing potential risks of contemporary synthetic biology activities and products, it is necessary to ensure continued safety protection proportionate to risk, while at the same time enabling scientific, technological and socio-economic advances. In a scenario where risk assessment would rely on contemporary methods alone, and where synthetic biology is able to design and produce novel life forms that differ from wild types in a much deeper and substantial way, risk assessors are going to face considerable difficulties in trying to assess the cases in front of them, as they will have troubles understanding the level of change and potential impact of the engineered organisms. One specific problem is the necessity to develop good *metrics* for ERA studies on SynBio agents. Almost the only metric currently available for engineered systems is the evaluation of the escape frequency to a given genetic modification. Yet, *mutational escape frequency under laboratory growth conditions is a necessary but insufficient metric to evaluate risks.* Unfortunately, the detection limit to assess the escape frequency is about 10^{-11}. But in order to be considered for a release into the environment, a physically contained industrial fermenter or even into a human patient, the escape frequency will have to be significantly lower than that. Moreover, there are no standards in terms of the media to test the escape frequencies in different environmental contexts. This state of affairs asks for a new SOPs for quantifying risks and involving not only escape frequencies but also develop and incorporate benchmarks and best practices available in what can be called *biosafety engineering* that address the issues spelled out above (see *Criteria for definition of safe chassis*)

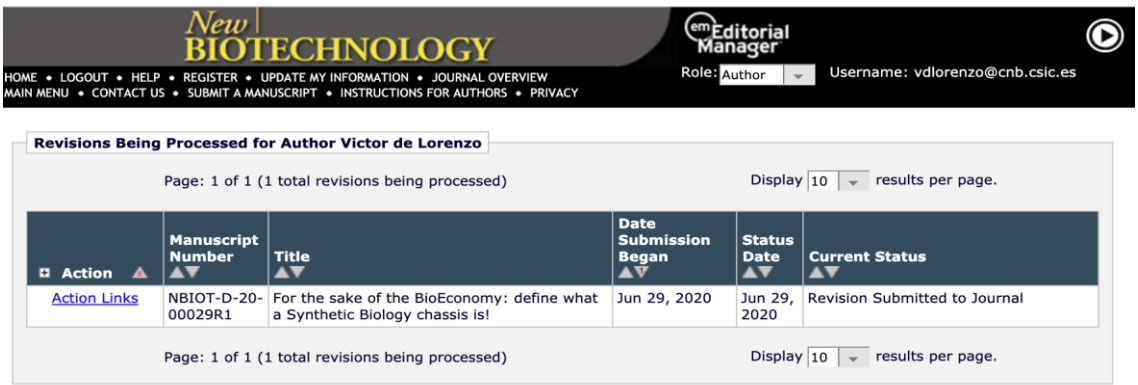
Barcoding as an avenue to ease traceability and manage containment uncertainties

Regardless of the many proposals of genetic firewalls for containing genetically engineered organisms and SynBio agents, reality if that current metrics (see above) never go beyond event occurring at frequencies of 10^{-11}, what is not enough for what has been called *certainty of containment* (CoC). There is widespread opinion in the Life Sciences community that no firewall, sophisticated as it might be, will stop engineered organisms to scape a given niche. While CoC is a fascinating scientific question most of the concerns on undesirable propagation of human-made constructs can be managed if chassis and agents were barcoded with specific and unique DNA sequences which—once decoded—could take users to the most detailed information available for this or that particular construct.

Barcoded clones would thus be equivalent to pets implanted subcutaneously with identification chips: in case they get lost or do some harm, their owner and their pedigree can be immediately identified, measures taken and even liabilities established. By the same token, *barcoded* SynBio strains would allow accessing all relevant information on its pedigree, safety and modifications implemented in them. This will be ultimately more useful that any containment circuit—which in all cases are bound to fail. It must be highlighted that the technologies to this end are not completely mature yet: DNA barcodes must be entirely alien to any biological significance, cannot have similar equivalents in extant genomes, must not cause any burden, should hold a minimum of mutations and should be designed for persisting in face of all types of stresses. This requires a considerable computational effort, as the challenge lies more in the side of information technology than purely biological methods. Once such technicalities are solved, the rules for barcoding SynBio agents will have to be promulgated following agreed standards. Barcodes will not only make traceability simple, but it will also assign a non-ambiguous cipher to the growingly improved versions of the same chassis (as is the case with OSs of computers or iPhone types, see above).

New methodologies (eg millifluidic setups) for SynBio ERA techniques

As mentioned above, there is a considerable issue re the need of ERA metrics beyond mere calculation of scape frequencies. New techniques are badly needed for measuring the parameters specified above and answering the questions on safety of SynBio chassis and agents. The key challenge is testing their behaviour in a very large number of environmental conditions not just those afforded by typical Biology laboratories with flasks, plates, reactors and perhaps a number of micro and mesocosms. Fortunately, there is at this time a growing number of technical options for high-throughput testing of environmental conditions and simulation of physicochemical scenarios at a mm^3-scale. With these technologies in hand one could combine in small droplets moving through a continuous tube and separated by gas and oil interfaces a large number of parameters (humidity, O2, textures, nutrients, chemical landscape, temperature, osmotic pressure etc) to produce thousands of different micro-environments were SynBio agents at stake could be tested. Existing mili-fluidic platforms such as Milli Drop (http://www.millidrop.com//) can be refactored as experimental setups able to provide quantitative evidence on specific ERA questions on given constructs supported with very solid statistics as compared to existing methods. Operating such platforms over time—which could be seen as environmental equivalents to clinical tests for new drugs— could deliver standards of reference for a wide range of situations, including long-term persistence, biological status of the cells at end points, worst case scenarios and testing of possible countermeasures (eg antibiotics, phages).



Note that much of the content of is in the process of publication as an invited manuscript for the journal *New Biotechnology*, the official publication of the European Federation of Biotechnology. The paper was submitted by end of February. 3 month later, the ms was returned for revision and uploaded on June 29 (see inserted screenshot above. At this date we are still waiting for an editorial decision, which the Journal has told us may take time because of all the delays connected to the COVID19 impasse.

REFERENCES

Becker, J., and Wittmann, C. (2012) Bio-based production of chemicals, materials and fuels–Corynebacterium glutamicum as versatile cell factory. *Current opinion in biotechnology* **23**: 631-640.

Beites, T., and Mendes, M.V. (2015) Chassis optimization as a cornerstone for the application of synthetic biology based strategies in microbial secondary metabolism. *Frontiers in microbiology* **6**: 906.

de Lorenzo, V. (2011) Beware of metaphors: chasses and orthogonality in synthetic biology. *Bioengineered bugs* **2**: 3-7.

de Lorenzo, V., and Schmidt, M. (2018) Biological standards for the Knowledge-Based BioEconomy: what is at stake. *New biotechnology* **40**: 170-180.

Heider, S.A., and Wendisch, V.F. (2015) Engineering microbial cell factories: Metabolic engineering of Corynebacterium glutamicum with a focus on non-natural products. *Biotechnology journal* **10**: 1170-1184.

Hutchison, C.A., 3rd, Chuang, R.Y., Noskov, V.N., Assad-Garcia, N., Deerinck, T.J., Ellisman, M.H. et al. (2016) Design and synthesis of a minimal bacterial genome. *Science* **351**: aad6253.

Komatsu, M., Uchiyama, T., Ōmura, S., Cane, D.E., and Ikeda, H. (2010) Genome-minimized Streptomyces host for the heterologous expression of secondary metabolism. *Proceedings of the National Academy of Sciences* **107**: 2646-2651.

Lee, H.H., Ostrov, N., Wong, B.G., Gold, M.A., Khalil, A., and Church, G.M. (2016) Vibrio natriegens, a new genomic powerhouse. *bioRxiv*.

Lee, S.Y., and Kim, H.U. (2015) Systems strategies for developing industrial microbial strains. *Nature biotechnology* **33**: 1061-1072.

Nikel, P.I., Chavarria, M., Danchin, A., and de Lorenzo, V. (2016) From dirt to industrial applications: Pseudomonas putida as a Synthetic Biology chassis for hosting harsh biochemical reactions. *Curr Opin Chem Biol* **34**: 20-29.

Nora, L.C., Westmann, C.A., Guazzaroni, M.-E., Siddaiah, C., Gupta, V.K., and Silva-Rocha, R. (2019) Recent advances in plasmid-based tools for establishing novel microbial chassis. *Biotechnology Advances*: 107433.

Pinto, F., van Elburg, K.A., Pacheco, C.C., Lopo, M., Noirel, J., Montagud, A. et al. (2012) Construction of a chassis for hydrogen production: physiological and molecular characterization of a Synechocystis sp. PCC 6803 mutant lacking a functional bidirectional hydrogenase. *Microbiology* **158**: 448-464.

Poblete-Castro, I., Becker, J., Dohnt, K., dos Santos, V.M., and Wittmann, C. (2012) Industrial biotechnology of Pseudomonas putida and related species. *Appl Microbiol Biotechnol* **93**: 2279-2290.

Tan, D., Xue, Y.-S., Aibaidula, G., and Chen, G.-Q. (2011) Unsterile and continuous production of polyhydroxybutyrate by Halomonas TD01. *Bioresource technology* **102**: 8130-8136.

van Dijl, J.M., and Hecker, M. (2013) Bacillus subtilis: from soil bacterium to super-secreting cell factory. *Microbial cell factories* **12**: 1.

Weinstock, M.T., Hesek, E.D., Wilson, C.M., and Gibson, D.G. (2016) Vibrio natriegens as a fast-growing host for molecular biology. *Nat Methods*.

Yu, J., Liberton, M., Cliften, P.F., Head, R.D., Jacobs, J.M., Smith, R.D. et al. (2015) Synechococcus elongatus UTEX 2973, a fast growing cyanobacterial chassis for biosynthesis using light and CO2. *Scientific reports* **5**.