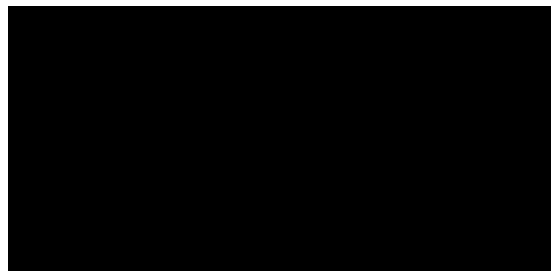# Modular Cell-free systems and their adaptable logic as core processing machinery

Felipe A. Millacura

Thesis submitted in fulfillment of
the requirements for the degree of
Doctor of Philosophy
University of Edinburgh
2020

Chapter 1

Declaration

I declare that this thesis has been composed solely by myself and that it has not been submitted, in whole or in part, in any previous application for a degree. Except where states otherwise by reference or acknowledgment, the work presented is entirely my own.

Felipe Aguilera Millacura

January 2021

Chapter 2

Abstract

While Synthetic Biology represents a promising approach to solve real-world problems, the use of genetically modified organisms (GMO) is still a cause of legal and environmental concerns. Cell-free systems (CFS) are an emerging technology where cell extracts are used instead of genetically-modified cells, thus, not presenting a "living prospect" applicable to current legal regulations.

Since there is a need for development of novel systems using the cell-free approach, and considering that most attempts have been focused on mimicking normal cell behaviour. This work has as principal aim to generate modular cell-free and cell-based systems capable of not only detect variables present in the environment, such as heavy metals or antibiotics, but also analyse them via the usage of unnatural behaviours, such as logic gates.

In vivo logic gates, for instance, have proven difficult to combine into larger devices. Here it is explained a cell-based logic system, ParAlleL, which decomposes a large genetic circuit into a collection of small subcircuits working in parallel, responding each subcircuit to a different combination of inputs. A final global output is then generated by a combination of the responses. Using ParAlleL, for

the first time a completely functional 3-bit full adder and full subtractor were generated using *Escherichia coli* cells, as well as a calculator-like display that shows a numeric result, from 0 to 7, when the proper 3 bit binary input is introduced into the system. This parallel approach facilitates the design of cell-based logic gates by the decomposition of complex processes into their main components, avoiding the need for complex genetic engineering and solving some legal limitations.

Cell-free systems, on the other hand, have emerged as a possible solution but much work is needed to optimize their functionality and simplify their usage for Synthetic Biology. Here it is presented a transcription-only genetic circuit (TXO), which is independent of translation or post-translation maturation. RNA aptamers are used as reaction output allowing the generation of fast, reliable and simple-to-design transcriptional units. TXO cell-free reactions and their possible applications become a promising new tool for fast and simple bench-to-market genetic circuit and biosensor applications.

Additionally, it's shown a versatile cell-free system based on the master survivalist bacteria *Cupriavidus metallidurans* CH34, able of not only sensing environmental variables, such as heavy metals, but also synthesize proteins and produce bioplastics. This novel cell-free chassis comes after the discovery of the unstable genome that this bacteria carries, which is also explained here, offering novel possibilities of development considering the cell free approach.

Chapter 3

Lay Summary

In an era where Genetically Modified Organisms (GMOs) are a mainstream topic within society, there is a need for different approaches without the usage of living organisms. One of the most promising technologies is the use of Cell-free systems, which don't imply a living prospect applicable to current legal or moral society restrictions.

During Cell-free preparation the cell wall is ripped off, the insides collected, and the cell catalysts used to produce new kinds of molecules and biological processes without the evolutionary constraints of using intact living cells.

In this work novel cell-based and cell-free system approach were developed. These allow the detection of pollutants, such as heavy metals or antibiotics, which are used as input signals and later analysed in binary using logic gates, just as it happen in computers.

Simple biological molecules, such as RNA, are used to accelerate the detection process, allowing us to avoid imitating ordinary cell conduct, which is the most common scientific rationale to solve these kind of problems. RNA aptamers are

used as outputs allowing the generation of fast, reliable and simple-to-design detection units. These cell-free reactions and their possible applications become a promising new tool for fast and simple bench-to-market genetic circuit and biosensor applications.

Working with computer-like behaviour, a system that decomposes large operations into smaller ones is here explained. This system allows faster analysis by processing information in a collection of small subcircuits instead of a large one, just like it happens in a Graphical Processing unit (GPU), also known as parallel computing. This was further demonstrated by creating a calculator-like display that shows a numeric result, from 0 to 7, when the proper 3 bit binary input is introduced into the system. This parallel approach facilitates the analysis avoiding the need for complex genetic engineering, solving some legal and ethical implications.

Additionally, a versatile cell-free system based on the extremely tolerant bacteria *Cupriavidus metallidurans* CH34 is shown here. This is able of not only sensing environmental variables, such as heavy metals, but also synthesize proteins and produce bioplastic. This novel cell-free chassis comes after the discovery of the unstable genome that this bacteria carries, which is also explained in this work, offering novel possibilities of development considering the cell free approach.

Chapter 4

Acknowledgements

Soyoung, Ainur, Kuanish, Bakyt, Aigerim, Philip, Peter and Barbara. And also to the CodeClan mates: Dhileas, Jennie, Rhiannon, Ric, Kostas, Jonathan, Kuba, Miles, David, Stephanie, Del, Aileen, and Mhairi. Most importantly, I would like to thank my family for the unconditional support. To Griselda and Francisco who are my major motivation to succeed in life. To my dad and cousins who convinced me to visit Chile while doing my PhD. Specially to Camila and Cassandra. To my aunts and uncles for the love given. To all my friends who even though the distance always tried to keep in touch, an honorific mention to Valeska, Ariel and Cristian for their incredible support. Thanks also to Macarena, Leyla and Merari who kept me rational and sane during the harsh times. Finally, thanks to Maria Victoria who always believed in me even more than what I did myself. Thanks to all of you I am able to finish this PhD thesis and, hence, this is also part of your success. Your support and patience contributed to make the impossible possible.

Chapter 5

Agradecimientos

A mis compañeros e instructores del ELE-EAP: Yermek, Nurlan, Daniyar, Korlan, Mohammed, Dayoung, Dauren, Soyoung, Ainur, Kuanish, Bakyt, Aigerim, Philip, Peter y Barbara. También a mis compañeros de CodeClan: Dhileas, Jennie, Rhiannon, Ric, Kostas, Jonathan, Kuba, Miles, David, Stephanie, Del, Aileen, y Mhairi. Me gustaría agradecer a mi familia por su apoyo incondicional. A Griselda y Francisco por ser mi mayor motivación para salir adelante en la vida. A mi padre y primos quienes me convencieron en visitar Chile mientras realizaba mi PhD. Especialmente a Camila y Cassandra. A mis tías y tíos por el cariño entregado. A todos mis amigos, que aun a la distancia trataron siempre de mantener el contacto, una mención honorifica a Valeska, Ariel y Cristian por su increíble apoyo. Gracias también a Macarena, Leyla y Merari quienes me mantuvieron racional y cuerdo durante los tiempos difíciles. Finalmente, gracias a María Victoria quien creyó siempre en mi incluso más de lo que yo mismo hacía. Gracias a todos ustedes hoy soy capaz de terminar esta tesis, y por lo tanto, esta victoria también es suya. Vuestro apoyo y paciencia contribuyó a hacer lo imposible posible. Esta investigación no seria factible sin el apoyo de la 'Agencia Nacional de Investigación y Desarrollo (ANID, ex CONICYT)' y su programa 'Formación de Capital Humano Avanzado' quienes me hicieron benefactor de la 'Beca Chile-PhD 72170403'.

Contents

List of Tables

List of Figures

Chapter 6

General Overview

Since the begining of time life has made his way through multiple adverse conditions. Organisms have evolved from single cells interacting directly with their environment, to specialised multi cellular entities that respond specifically to different stimuli. Even though cells can specialise their functions by differentiating into diverse cell-types, by aggregating into tissues, generating organs and/or forming a variety of organisms, they all respond to the same paradigm. DNA is used to save and protect the hereditary information, which is later transcribed into functional orders in the form of RNA, being most of the time translated into protein products that execute those final orders.

This process was broadly studied on the twentieth century in order to explain how life works. DNA was discovered to be the main source of information, in a double helix format, carrying and saving all necessary instructions to secure life over generations (**?**). It was discovered a few years later that the incorporation of the deoxyribonucleotides, from the triphosphates of deoxyadenosine, deoxyguanosine, deoxycytidine and thymidine, into the deoxyribonucleic acid was catalyzed by an enzyme, the DNA polymerase, which was purified at that time directly from *Escherichia coli* cell-free extracts (**?**).

Proteins were further studied as the final product of DNA's information translatation, helped by the discovery of the green fluorescent protein (GFP). A naturally fluorescent protein present in the *Aequorea Victoria* jellyfish (**?**). This allowed the generation of proteinic chimeras, detection and real-time tracking of the products. Another type of enzyme that would revolutionise Molecular Biology were the Endonucleases. Discovered in the early 60s by **?** and recognised as "molecular scissors" protecting bacteria from bacteriophages.These restriction enzymes work in conjuction with methylases protecting the DNA of the bacterium. The first Type IIS enzyme described was HindII isolated for the first time from *Haemophilus*

influenzae Rd **?**, which also described the 6 bp site-specific cleavage. The understanding of how restriction enzymes cut DNA, and how host DNA tries to protect itself are the basis of modern genetic engineering.

Thanks to the advances in DNA cleavage, recombinant DNA molecules were created for the first time by the fusion of genetic material from more than one species. Here the DNA from two viruses was cut creating "sticky ends" and the ends annealed, finishing with the addition of DNA ligase for sealing it. **?**. The aquired knowledge was further expanded by the corroboration of the previous experiments via gel electrophoresis and testing them in SV40 viral genome **?**, work that was vital for the generation of genomic mapping.

The introduction of recombinant DNA into living organisms was achieved by the work **?**, proving that DNA can replicate naturally, despite being artificially introduced in another organism. Establishing the principles of modern genetical engineering.

Most previous achievements would not have been posible without the discovery of *in vitro* translation via the usage of cell-extracts. **?** demontstrated that RNA and protein synthesis can occur almost simultaneously in the cell, and that this process can be completly replicated in a test tube. Even though the coupling of transcription and translation accellerate the product production, proteins still need to acquire their secondary, tertiary or even quaternary conformation in order to become functional. Even sometimes needing post-translational modifications to work.

As it can be observed, the usage of cell-free technology is not a new technic neither Synthetic Biology. For instance, one of the first mentions of Synthetic Biology as a term can be tracked way back to **?**, who sought to synthesise life "by directing the physical forces which are its cause".

Cell-free expression systems are a powerful tool in applications where cell based expression systems are too variable, slow, difficult to store or prohibitive due to legislation on the release of genetically modified organisms. Cell-free systems (CFS) are not without their limitations; since their very nature is to be a defined, precisely controlled system that lacks self-replicating functionality, they depend on a limited quantity of supplies, such as ATP, amino acids and nucleotides **?**. Consequently, in-vitro synthesis of protein is limited by two dimensions: the overall maximum protein amount that can be synthesized with the supplied components, and the time during which the CFS stays active before background

processes and chemical deterioration have used up supplies or inactivated the system **???** One of the most promising applications of CFS is the possibility to design cell-free biosensors, which do not present a living prospect subject to the current GMO legal regulations and offer an alternative to standard analytical techniques, such as ICP-MS and ICP-OES. However, cell-free biosensors that rely on protein expression also present the common limitations of CFS, such as a long response time and a lack of precursor regeneration. Sensors that could function without the need for protein expression would drastically reduce the requirements for such a system. Paige et al., (2011) designed synthetic fluorescent RNA aptamers that mimic fluorescent proteins without the need for translation. The secondary structure of these aptamers presents a loop in which a fluorophore is trapped, conferring the same intramolecular immobilization that confers fluorescence in fluorescent proteins (Paige et al. 2011; Strack et al. 2013). These RNA aptamers also suffer from several limitations as incorrect folding is observed under non optimal thermal or ionic conditions (Autour et al. 2016). Improved aptamers with a stronger secondary structure, such as Broccoli and iSpinach, show increased stability under non-optimal conditions (Filonov et al. 2014, Autour et al. 2016). Both aptamers are based on the 95 base core sequence of Spinach2 and can be flanked by a tRNA scaffold, increasing the stability of their secondary structures (Filonov et al. 2014, Autour et al. 2016, Strack et al. 2013). However, these were designed for in vivo use (e.g. analysing RNA quantities) and have never been tested as outputs for a cell-free sensor. Such a sensor could require only transcription, not translation, hence could be much faster and simpler. On the other hand, the simplicity of CFS can become a disadvantage when complex processes need to beperformed. For instance, the high dependence on E. coli as a model organism generates a problem, since genes or genetic pathways from non-model organisms or obtained by metagenomic studies, may fail to be expressed in standard organisms. Moore et al., (2016 and 2017) have attempted to solve this issue by using non-model bacteria, such as Bacillus or Streptomyces, for the generation of new cell-free TX/TL systems. This approach can be applied to other organisms, such as extremophiles, in order to use their unique capacities. For example, Cupriavidus metallidurans possesses the capability to tolerate, remove and degrade diverse environmental pollutants(Millacura et al., 2017). A cell-free system based on Cupriavidus metallidurans extract might show increased tolerance for heavy metals and might be superior for heavy metal sensors.

Another challenge for CFS is to analyze multiple variables simultaneously. As they rely on the normal cell processing/synthesis machinery (interaction between transcription factors, polymerases, ribosomes, and other diverse macromolecules), they suffer the same issues as a living organism. Furthermore, when complex interactions are carried out, problems may arise due to kinetic mismatches, lack of oscillation or Boolean/Reversible designs (Guz et al., 2016). The generation of cell-free systems that respond to variables using totally synthetic processing machinery seems to be one of the most promising approaches. There have been some attempts to generate logic gates without mimicking normal cell machinery (Bordoy et al., 2016, Chatterje et al., 2016, Kim et al., 2014 and Zhang et al., 2016), however, due to their dependence on translation and/or the need for manually addition of foreign oligonucleotides (Kim et al., 2014) they still show slow response. Additionally, some of them rely on recombination processes that make their implementation even more complicated than under in vivo conditions (Zhang et al., 2016). In electronics, a single circuit accepts one or more binary inputs to generate one or more binary outputs. There have been many attempts to replicate such circuits using in vivo genetic networks. A typical biological logic gate consists of an output macromolecule that is produced only if the corresponding pattern of inputs is present, inputs that are commonly associated with the presence of transcriptional regulators, transcription factors, polymerases or other macromolecules **?** . Here we propose an alternative approach, BioLogic, which decomposes a large circuit into a collection of small subcircuits, solved in parallel. Rather than having a single type of cell (or genetic material) doing the computation, we have separate versions (subcircuits) each reacting to a different combination of inputs and generating the desired response by combination of each final output. For instance, if each input bit is considered in two forms, ZERO and ONE, each of which is essential to certain output agents, any arbitrary pattern of outputsfor any pattern of inputs could be generated, making all kinds of binary operation possible A further proposal is to achieve this goal by using just a transcriptional genetic network, without need for translation. Each agent in this case, rather than a cell, is a plasmid or double stranded DNA that does not possess functional transcriptional networks (Figure 1). As cell-free systems can assemble DNA parts in vitro at very low cost by using cell extracts instead of enzyme cocktails (Casini et al., 2015), and Golden Gate Assembly is an excellent tool to assemble multiple DNA fragments in a defined linear order, by using type

IIS restriction enzymes (Engler et al., 2008), these techniques may be combined to generate a new and inexpensive in vitro system, which through DNA assembly and processing (hereafter DNALogic) would be able to sidestep common problems associated with the use of live cell based systems. As a "triggering" step is vital for the DNALogic processing machinery, the novel CRISPR/Cpf1 technology (or similar) will be used to generate the respective recognition and cutting steps. CRISPR/Cpf1 generates 5' overhanging cuts, producing 8 nt sticky ends and cleaving at the 14th base from the PAM site TTN (Lei et al., 2017 and Li et al., 2016). Due to its similarities to Type IIS restriction enzymes, Cpf1 can also be used to generate a new kind of Golden Gate Assembly, which would not rely on the presence or absence of cutting sites, but on gRNA. The gRNA expression system can be combined with the recognition step, working as a transduction system between the detection and processing steps (Figure 1a). This CRISPR triggered sequence recognition, followed by cutting and ligation steps, would result not only in a fully functional transcriptional unit, but also in a different double stranded genetic sequence, accomplishing the parameters for synthetic memory generation explained and demonstrated in vivo by Siuti et al. (2013). This creates DNA memory that can not only be later amplified and processed by sequencing and/or PCR (later response), but that can also generate RNA expression of a fluorescent aptamer or a further gRNA if there is formation of a proper transcription unit (immediate response).Figure 1. Proposed DNALogic system. A) crRNA expression is regulated by inducible promoters B) Ribonucleoprotein formation is followed by recognition and cutting of the target DNA C) Ligation of the overhangs produced by the Ribonucleoprotein generates a transcriptional unit D) Expression of the transcriptional unit generates a functional fluorescent RNA aptamer.Aims: This work has as its general aim the generation of a modular cell-free system capable of not only sensing variables present in the surrounding environment, such as heavy metals, but that is also able of analyzing them by using a totally synthetic in vitro logic gate. Specific aims are: 1. Testing fluorescent RNA aptamers as output signals for cell-free sensors. 2. Developing Cpf1-based signal processing technology for use in cell-free systems. 3. Creation of an adaptable DNALogic gate, which allows memory and more complex analysis in cell free systems. 4. Testing Cupriavidus metallidurans extracts as basis for metal-sensing cell free systems using endogenous and exogenous metal sensing

promoters. 5. Integrating these to develop a modular cell-free sensing system with versatile in-vitro signal processing.

Chapter 7

ParAlleL: A Novel Population-Based Approach to Biological Logic Gates

7.1 Abstract

In vivo logic gates have proven difficult to combine into larger devices. Our cell-based logic system, ParAlleL, decomposes a large circuit into a collection of small subcircuits working in parallel, each subcircuit responding to a different combination of inputs. A final global output is then generated by a combination of the responses. Using ParAlleL, for the first time a completely functional 3-bit full adder and full subtractor were generated using *Escherichia coli* cells, as well as a calculator-style display that shows a numeric result, from 0 to 7, when the proper 3 bit binary inputs are introduced into the system. ParAlleL demonstrates the use of a parallel approach for the design of cell-based logic gates that facilitates the generation and analysis of complex processes, without the need for complex genetic engineering.

7.2 Introduction

A major challenge in the field of synthetic biology is the construction of complex logic circuits that analyze variables as in electronics; where a single circuit accepts one or more binary inputs to generate one or more binary outputs. A cell-based logic network consists of engineered cells producing an output macromolecule only if the corresponding pattern of inputs is present. The mechanism of analysis is commonly based on the use of transcriptional regulators, transcription factors, polymerases, receptors, or recombinases **?**. Some examples of genetic circuits

mimicking computational behavior are toggle switches, oscillators, boolean logic gates, feedback controllers, and multiplexers. Although there are genetic circuits that simulate computational behavior, the complex engineering of their biological chassis is affected by gene expression noise, mutation, cell death, undefined, and changing extracellular environments and improper interactions with the cellular context **?**. Furthermore, complex genetic engineering is necessary when multiple input variables are analyzed, limiting the processing capacity of the system.

Biological multiplexers analyze one or more signals over a common transmission line using interconnected transcription factors, recombinases, antisense RNA, or CRISPR-like technology **???**. However, complex genetic engineering is needed for wiring the basic computational units, becoming inefficient for moving beyond simple NOT or AND logic gates or for scaling to 3 bit logic circuits. The complexity of the genetic engineering required can be reduced by using distributed logic circuits, where the computation is distributed among several physically separated cellular consortia that each sense only one signal and respond by secreting a communication molecule **?**. As a circuit responds to one signal, but not another, due to spatial distribution, a change in the state of the system can be triggered as response, making synthetic learning possible **??**. Even though the consortium approach makes Boolean circuit design simpler, it still shows a slow response and considerable complexity since each cell needs to recognize, synthesize and secrete a wiring molecule **?**.

Here we propose an alternative logic architecture, which decomposes a large circuit into a collection of small subcircuits acting in parallel (hereafter ParAlleL). Rather than having a single type of agent (such as a genetically engineered cell) doing the computation, ParAlleL has separate types of agent that each react to a different combination of inputs. A final output is then generated by combination of the responses, making all kinds of binary operation possible. As an example, here we show the implementation of this concept using cells resistant to different combinations of antibiotics, with the response indicated by growth. This is used to demonstrate a completely functional 3 bit full adder and full subtractor, as well as a calculator-style display that shows digits from 0 to 7 based on three binary input bits.

7.3 Methodology

7.3.1 Reagents and Stock Solution Preparations

Antibiotic stock solutions were prepared as follows: 100 mg/ml carbenicilin disodium salt (Sigma-Aldrich #C1389), 50 mg/ml kanamycin sulfate (PanReac Applichem #A1493), 20 mg/ml chloramphenicol (Acros Organics #22792), 10 mg/ml tetracycline hydrochloryde (Duchefa Biochemie #T0150), 10 mg/ml gentamicin sulfate (Melford #G0124), and 50 mg/ml spectinomycin.HCl (LKT Labs #S6018). Developing solution contained 0.1 %w/v bromothymol blue (Sigma-Aldrich #114421) and 400 mM Trizma base pH7.5 (Sigma-Aldrich #T1503).

7.3.2 Generation of Subcircuit Cells

E. coli JM109 was transformed with 200–300 pg of plasmid pSB4A5 (AmpR) or pSB4C5 (ChlR) (Registry of Standard Biological Parts) and selected on 100 µg/ml carbenicilin (Am) or 20 µg/ml chloramphenicol (Ch), respectively. Cells carrying the first bit plasmid were made chemically competent ? and transformed with 200–300 pg of the 2nd bit plasmid, pSB1T3 (TetR) or pSB1K3 (KanR) (Registry of Standard Biological Parts). Selection was performed with the first antibiotic (Am or Ch) and the addition of 10 µg/ml Tetracycline (Tc) or 50 µg/ml kanamycin (Km), obtaining the two-bit combinations Km/Am (KA), Tc/Am (TA), Km/Ch (KC), and Tc/Ch (TC). This set of strains is sufficient to implement all two-bit binary operations.

The third bit layer was generated by transforming these four strains with pSEVA631 (GenR) ? (GenBank JX560348) or pMO9075 (SpmR) ?. Resulting strains were selected on the 2-bit antibiotic combinations plus 10 µg/ml gentamicin (Gm) or 50 µg/ml spectinomycin (Sm). This gave 8 strains, designated ATG (Am/Tc/Gm), AKG (Am/Km/Gm), ATS (Am/Tc/Sm), AKS (Am/Km/Sm), CTG (Ch/Tc/Gm), CTS (Ch/Tc/Sm), CKG (Ch/Km/Gm), CKS (Ch/Km/Sm) based on their resistance markers. This set of strains is sufficient to implement all three-bit binary operations. Plasmid specifications are listed in Figure S2, Tables S1, S3, with further information about these antibiotics in Table S2. Plasmid sequences are available in different formats at https://doi.org/10.7488/ds/2497.

7.3.3   Three-Bit Logic Operations

Tests were performed in 96-well microplates by inoculating cells (1:100) in LB broth (100 μL) supplemented with 1%w/v D(+)-glucose (Fisher Chemical #G0500). Plates were incubated for 18 h at 37℃ without shaking and then developed by addition of the developing solution (0.1%w/v bromothymol blue in 400 mM Tris, pH7.5) in a ratio 1:20. Images were obtained using a Kodak ESPC315 Flatbed scanner. Design of the calculator-like display, full adder, and subtractor are shown in Figures 2, 3 and in Supplementary material (Figure S3). Raw figures were deposited at https://doi.org/10.7488/ds/2497.

7.4   Results

In the distributed logic system of ParAlleL each input bit has two forms, ZERO and ONE, each of which is essential to certain output agents and inhibitory to others. Thus each agent reacts only to a certain combination of input bits, allowing generation of any arbitrary pattern of outputs for any pattern of inputs. In the implementation shown here, each input bit comes in two forms, each being an antibiotic lethal to sensitive strains. In this case, bit A is represented by ampicillin for zero, chloramphenicol for one, bit B by kanamycin for zero, tetracycline for one, and bit C by gentamicin for zero, spectinomycin for one. Thus, four strains are needed to implement any operation with two input bits, and eight strains for three input bits. In contrast to other cell-based logic schemes, only very minimal genetic engineering is required, essentially transformation with 3 different antibiotic resistance markers.

Cells show a global response concordant with the behavior expected for a 1 bit, 2 bit, or 3 bit system (Figure 1). For instance, when the input 101 (chloramphenicol, tetracycline and spectinomycin) is added to the system growth is only observed in the corresponding CTS cells, which carry the proper resistance markers. The response time of the system is around 12 h (Figure S1) but plates were developed at 18 h to avoid false negatives or positives.

In order to further test the ParAlleL system, a digital calculator-like display was designed (Figure 2A). In this case, multiple subcircuit cells are mixed in one well and the global response displays a number from 0 to 7 when the proper binary input is applied. Numbers represent the total eight possible values encoded within 3-bit binary inputs. Input configuration versatility was proven by representing bit
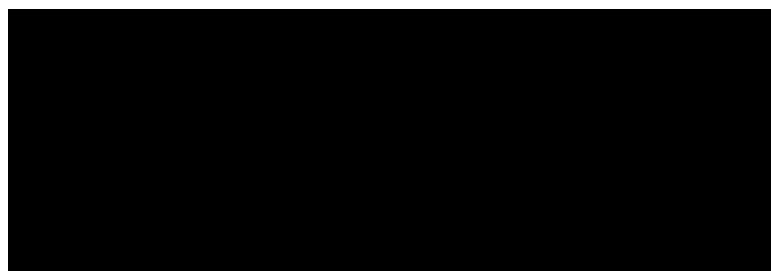
Figure 7.1: ParAlleL responding to 1 bit, 2 bit, and 3 bit inputs. ParAlleL subcircuit cells were spatially distributed in different wells (vertically) and exposed to specified 1 bit, 2 bit, or 3 bit inputs (top of each column). Cells were inoculated (1:100) in LB supplemented with 1% w/v glucose. After 18 h of incubation at 37°C, plates were developed by addition of 0.05 volumes of the developing solution.
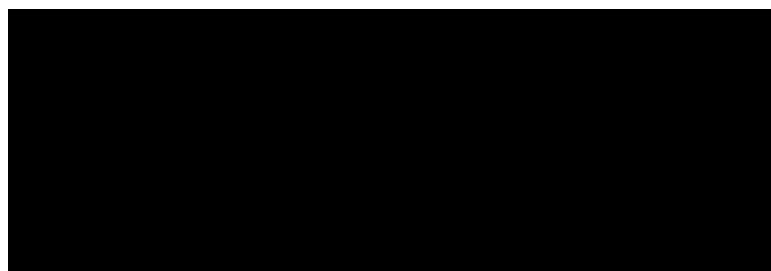
Figure 7.2: Digital calculator-like display using 3 bit ParAlleL. Figure shows all numerals from zero to seven based on the 8 binary inputs provided. (A) Subcircuit cells were mixed and distributed in a 3 × 5 matrix and inoculated (1:100) in LB supplemented with 1%w/v glucose. Plate was developed after 18 h of incubation, by addition of 0.05 volumes of developing solution. (B) Output number results obtained by addition of each 3 bit antibiotic combination.

A in this case by gentamicin for zero, spectinomycin for one, bit B by kanamycin for zero, tetracycline for one, and bit C by ampicillin for zero, chloramphenicol for one. For instance, once input 110 represented by Ch/Km/Gm is added in the system, the number 6 is displayed (Figure 2).

Finally, a full adder and a full subtractor were designed. A full-adder adds three binary inputs, often denoted as A, B, and Cin, generating a Sum result (S) and a Carry-out (Cout). A full subtractor, on the other hand, has a minuend (X), a subtrahend (Y) and an additional Borrow-in (Bin) as inputs. The subtraction operation produces a difference (D) and a Borrow (Bout) (Figure 3). In order to
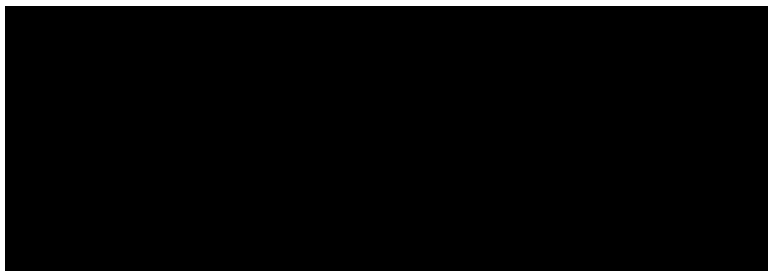
Figure 7.3: Full adder and subtractor using the 3 bit ParAlleL system. The figure shows results of addition and subtraction using the ParAlleL for 3 bit system. Cells were mixed as shown in Figure S3 and inoculated (1:100) in LB supplemented with 1%w/v glucose. After 18 hours of incubation, the plate was developed by addition of the 0.05 volumes of developing solution.

generate the full adder and subtractor, multiple subcircuit cells were mixed and distributed in two different wells (Figure S3). One well represents the solution (S) or difference (D) and a second one the carry (Cout) or borrow (Bout), for the adder and subtractor, respectively (Figure 3). A yellow color represents growth and a positive output 1, a blue color represents no growth and a binary 0 output instead.

7.5   Conclusions/Discussion

Subcircuits that solve complex calculations in parallel have been extensively used for computation in order to reduce the total computation time. Translating this approach to biological systems would allow us to analyze complex processes, currently difficult in synthetic biology, as multiple simple sub-circuits.

In our proof of concept, we present a biological information processing system, ParAlleL, capable of exploiting the parallelism in mixed bacterial cultures. ParAlleL decomposes the analysis of 2 and 3 bit complex inputs, into 4 and 8

sub circuits, respectively (Figure 1). Each sub-circuit corresponds to a different *E. coli* strain carrying a different combination of antibiotic resistance markers (Table S1). As an example, in the 3 bit system the input 000 is represented by the antibiotics ampicillin, tetracycline and gentamicin (Figure 1). When this input is entered into the system, all cells that are not encoded for responding to 000 will die, but cells carrying the proper plasmid combination, pSB4A5, pSB1T3, and pSEVA621 will not (Figures 1.0, 1.1), therefore, a live/dead response (output) is achieved in all sub circuits, the output of each well being one (growth) or zero (failure to grow) (Figure 1.1 and Figure 1.4).
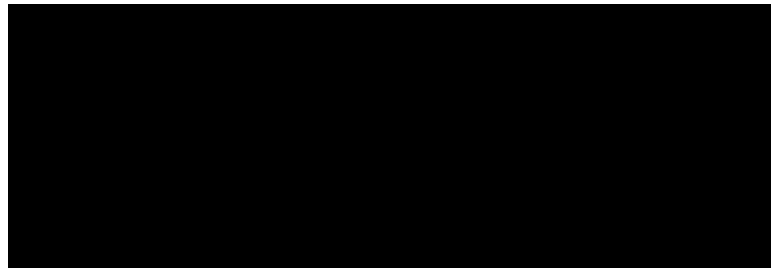
Figure 7.4: Plasmids carried by subcircuit cells. Plasmids carrying antibiotic resistances for Ampicillin (pSB4A5); Chloramphenicol (pSB4C5); Tetracycline (pSB1T3); Kanamycin (pSB1K3); Gentamicin (pSEVA631) and Spectinomycin (pMO9075) represented in SBOL format. Plasmid sequences available at https://doi.org/10.7488/ds/2497.

ParAlleL uses cellular consortia instead of a single type of cell. A similar approach has been developed by **??** using eukaryotic cells, and even showing the possibility of generating transient memory. However, that approach requires a sophisticated design as it relies on a secreted intermediate molecule (hormone-like) that must be kept at the right production level, and that should be previously activated by X (Repressor) and Y (SsrA-tagged protein) degradation. Furthermore, since the output of the circuit is distributed among different consortia, the concentration of the secreted molecule can differ according to the number of cells simultaneously producing it. This kind of multicellular approach and others based on single cells require sophisticated wiring design **????**. By contrast, ParAlleL requires very minimal genetic modification and little tuning to obtain reliable outputs (Figures 2, 3).

The implementation of ParAlleL presented here is simple, but its further development to useful applications presents a number of challenges. Firstly, expansion to 4 bits and beyond would require further well-behaved and non-cross-reacting antibiotic resistance markers, and would probably lead to even greater disparities
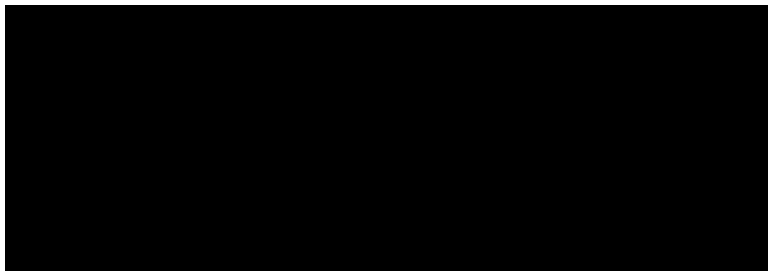
Figure 7.5: ParAlleL subcircuit cell growth curves. Growth curves of the 3-bit subcircuit cells in LB + 0.1% glucose with their respective antibiotic combinations. A: Carbenicilin (100 μg/mL), C: Chloramphenicol (20 μg/mL) T: Tetracycline (10 μg/mL), K: Kanamycin (50 μg/mL), G: Gentamicin (10 μg/mL). S: Spectinomycin (50 μg/mL). Overnight culture (0.01 volume) was used as inoculum.

in growth rate than those observed in the three-bit system (Figure S1). This could be addressed, and the flexibility and usefulness of the system increased, by moving away from direct use of antibiotics to a system using tightly controlled inducible promoters, each controlling a lethal "death gene," such as *ccdB*, with either the presence or absence of the inducer leading to lethality. In this way the system could be made to respond to different combinations of useful inputs, for example in the construction of multiplexed biosensors.

However, to achieve the full potential of ParAlleL, it will be necessary to generate layered systems in which the output from one layer serves as input to another layer. This might be accomplished via quorum sensing, but the implementation would be rather complex and limited. A more attractive option is to transfer the same concept, using a set of agents which each responds to a single combination of inputs, to an alternative system. For example, the concept could be implemented in a cell-free system, in which inputs may be present as small molecules interacting with transcription factors, or as DNA or RNA oligonucleotides. Outputs from the first layer, in the form of DNA or RNA molecules generated by DNA replication (e.g., PCR) or transcription, could then serve as inputs to a second layer. For example, in a PCR-based approach, one of a set of templates would be amplified based on which primer oligonucleotides were added; the output PCR product could then be processed to generate another set of oligonucleotides, or used directly to initiate priming on a second set of templates via formation of three-way junctions. Alternatively, in a transcription-based system, output RNAs from a first layer computation could act as guide RNAs directing binding of CRISPR-based transcription factors to further templates to generate a new set of guide

RNAs, eventually leading to production of an output RNA such as a mRNA leading to translation of a visually detectable signal.

Another interesting aspect of ParAlleL is that, since the final output is the result of a population-based calculation, these systems may show a level of "Byzantine fault tolerance," allowing reliable outcomes even in the face of levels of noise which are unavoidable in biological systems. This would represent a new level of robustness in biological computation systems.

7.6   Data Availability

All datasets generated for this study are included in the manuscript and/or the supplementary files. More Data is available at https://doi.org/10.7488/ds/2497

7.7   Acknowledgments

7.8   Supplementary Material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fbioe.2019.00046/full#supplementary-material

Chapter 8

TXO: Transcription-Only genetic circuits as a novel cell-free approach for Synthetic Biology

8.1 Abstract

While synthetic biology represents a promising approach to solve real-world problems, the use of genetically modified organisms is a cause of legal and environmental concerns. Cell-free systems have emerged as a possible solution but much work is needed to optimize their functionality and simplify their usage for Synthetic Biology. Here we present TXO, transcription-only genetic circuits, independent of translation or post-translation maturation. RNA aptamers are used as reaction output allowing the generation of fast, reliable and simple-to-design transcriptional units. TXO cell-free reactions and their possible applications are a promising new tool for fast and simple bench-to-market genetic circuit and biosensor applications.

8.2 Introduction

Cells, including microbial, plant and mammalian cells, have been long engineered for responding to a plethora of environmental factors, benefiting from their intrinsic ability to process the entire cycle from the recognition of a specific target, to analysis of the information, and generation of an output signal **?**. However, cell-based systems present also unavoidable disadvantages during their usage, including the risk of releasing genetically modified organisms (GMOs) into the environment, accumulation of mutations and genetic instability, uncontrollable

side-reactions within the cell, and issues with viability maintenance during long-term storage **????**.

In order to solve these issues, cell-free systems (CFS), also known as transcription-translation (TX-TL) systems, have emerged as a promising alternative. CFS not only inherit most of the benefits from cell-based systems, but also avoid major challenges that they face. Typically CFS are based on non-living cell-extracts or reconstituted systems that by not presenting a living prospect solve problems related to biosafety, cross-reactivity and long-term functionality maintenance. CFS are rather robust to factors that pose cellular stress, including sensitivity to toxic pollutants and to the typical "overload problem" observed when foreign "genetic circuits" are introduced into the cell **?**.

TX-TL systems share with cells a problem hindering their usage in real-time detection. Response time in TX-TL systems may need a long time for transcription, translation and post-translational maturation to occur, in order to finally produce a detectable signal. For instance, Pardee and colleagues **?** developed a cell-free paper toehold-switch sensor for detection of the Zika virus that takes hours for the final visual results to come out **?**. Paper-based cell-free sensors that detected heavy metals and gamma-hydroxybutyrate by expressing sfGFP needed at least one hour for producing a measurable response **?**. Considering the need for in-situ field analysis, an improvement of cell-free systems response time is necessary **??**.

Since CFS are open and easily modifiable systems **?**, time-consuming translation and post-translational modification can be removed, allowing novel approaches with faster signal generation. Alternative output signals independent of translation and post-translational modification can provide the system with a simpler composition, lower resource requirements and, most importantly, a shorter response time. Here, we introduce a transcription-only (TXO) cell-free system using RNA aptamers as output signals rather than translated proteins.

Fluorescent RNA aptamers are short ribonucleotide sequences that become fluorescent when bound to specific fluorophores. RNA aptamers, such as Spinach **?**, iSpinach **?** and Broccoli **?** have shown excellent properties for generating bright and stable fluorescence. In the past few years, RNA aptamers have been widely applied in live-cell imaging **?** and biosensing **?**. However, these previously reported aptamer-based sensors acted as both sensing elements (binding to

target molecules) and signal generator at the same time **??**, requiring aptamers specifically designed for each target.

In TXO systems the functions for input detection (sensing), processing (analysis) and output generation (reporting) are decoupled, creating modular systems where aptamers can be used as universal signal generators. To exemplify the versatility of such system, here we used a simple cell-free extract from the microorganism *Cupriavidus metallidurans* CH34, which carries most of the necessary heavy-metal (e.g., Hg^+, Pb^{2+}, Cu^{2+}, Zn^{2+}, Ni^{2+}, among others) responsive transcription factors and the RNA polymerase which they control **?**. Since all necessary transcription factors are provided directly by the cell-free extract only the synthetic transcription unit must be changed to detect these different metal ions. Combining of the easy to design TXO system, with the high sensitivity of the sensing elements and the fast output signal generation provided by RNA aptamers, TXO is a promising tool for accelerating bench-to-market Synthetic Biology, most specifically where in-situ, simple to design and rapid applications are needed.

8.3 Methodology

Reagents and culture conditions

Fluorophores DFHBI (Tocris #5609), DFHO (Tocris #6434), DMABI and 2-HBI (provided by Dr. Jaffrey's laboratory at Cornell University) were dissolved in DMSO (Fisher Scientific #BP231) to prepare 2 mM stock solutions (for further information see Table S1). Our optimised transcription/detection buffer (OTDB) was prepared as a 10X stock as follows: Tris base 400 mM pH 7.5 adjusted with HCl (Sigma-Aldrich, #T1503); $MgCl_2\cdot6(H_2O)$ 60 mM (Sigma-Aldrich, #M2670); DTT 100 mM (Melford, #MB1015) and Spermidine 20 mM (Alfa Aesar, #A19096.03). Metal-ion solutions were prepared from soluble salts of analytical grade: $CuSO_4\cdot5H_2O$ (Sigma-Aldrich #C8027), Arsenite $NaAsO_2$ (Sigma-Aldrich #202673), $Pb(NO_3)_2$ (Sigma-Aldrich #228621), $HgCl_2$ (Sigma-Aldrich #215465), $ZnCl_2$ (Sigma Aldrich #229997) and $NiCl_2\cdot6H_2O$ (Fisher Scientific #AC27051) in double-deionised water and filter-sterilised before use.

Strains and cell-free extract preparation

C. metallidurans CH34 cells were cultivated on a Tris-buffered mineral medium (MM284) **?** supplemented with succinate 0.4 %w/v, at 30°C with constant shaking at 180 RPM. *Escherichia coli* JM109 cells were grown on Luria Broth (LB) at 37°C. For cell-extract preparation, cells were collected at middle exponential phase (OD 0.6-0.8) by centrifugation at 3,900 × g and 4°C for 10 min. Cells were resuspended in one tenth of the centrifuged volume in a non-ionic lysis buffer, which contains Triton X-100 0.1%v/v (Sigma-Aldrich #T9284), Tris base (40 mM) adjusted to pH 7.5 with HCl and lysozyme 50 mg/mL (Sigma-Aldrich #L6876) for 30 min at 4°C. The lysed extract was centrifuged at 12,000 × g, 4°C for 20 min and the supernatant was immediately used or saved at -80°C for later use.

Construction of RNA aptamers

Aptamer sequences (see Table **??**) protected within the F30 scaffold (upper and lower sections) were synthesised as 120 bp oligonucleotide sequences (Sigma-Aldrich), and then extended using the 5 bp shared on the 3' of each oligonucleotide for the annealing (see Figure **??** and Table S2) A PCR reaction adding a T7 promoter/terminator pair on each aptamer (see Table **??**) was carried out using Q5 High-Fidelity DNA Polymerase (New England Biolabs Inc.(NEB) #M0491) and the following protocol: 95°C for 5 min (1 cycle), 95°C for 20 s, 54°C for 10s and 72°C for 20 s (35 cycles), with a final elongation at 72°C for 5 min. Amplification was analysed in a 1.5 % (w/v) agarose gel and purified from the PCR reaction using the QIAquick PCR Purification Kit (Qiagen, #28104), and subsequently cloned in to the pMini-T plasmid using the PCR Cloning Kit (NEB #E1202). Plasmids were extracted from *E.coli* JM109 transformed cells using the QIAprep Spin Miniprep Kit (Qiagen) and sequence verified by Sanger sequencing using the BigDye Terminator v3.1 Cycle Sequencing Kit (Thermo Fisher Scientific) and an Applied Biosystems 3730XL DNA Analyzer (Edinburgh Genomics).

Biosensor transcriptional unit design

Diverse inducible promoters and a T7 terminator were added at the ends of the aptamer F30::iSpinach via PCR **?**. The PCR was carried out using Phusion
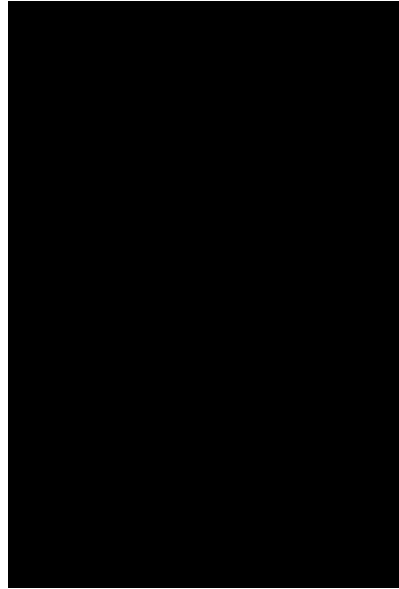
Figure 8.1: Schematic diagrams of TXO cell free system with minimal requirements. **(a)** Sequence and production process of diverse RNA aptamers via PCR, showing addition of BsmBI sites for further Golden Gate manipulation. **(b)** In a TXO system there are minimal requirements such as an inducer, moderator/repressor, RNA polymerase and a final RNA output herewith represented as a fluorescent aptamer.

High-Fidelity DNA Polymerase (NEB #M0530) and oligonucleotides containing the promoters from clusters *copAB*, *arsBCD* from *E. coli* MG1655 and *merTPAD*, *pbrAB*, *czcN*, *cnrH* from *C. metallidurans* CH34 (see Table S2). Amplified products were analysed and cloned for replication as described above.

RNA aptamer expression and detection.

Some previous protocols for *in vitro* transcription of fluorescent RNA aptamers **??** include purification and concentration steps that complicate the use of aptamers for real-time detection. A combination of both main steps, expression and detection, was achieved at room temperature (25°C) by using the OTDB buffer. *In vitro* transcription and detection assays were carried out using OTDB 1X, NTPs (2 mM each), Fluorophores 10 μM (see Table **??**), NEB T7 RNA Polymerase (1.5 U/μL), linear target DNA (650 nM), using 15 μL total reaction volume in 384 Well Small Volume^{TM} LoBase Black Microplates (Greiner Bio-One). Fluorescence was measured using a FluoStar Omega (BMG Labtech) plate reader, 10 mm bandpass, using 20 flashes per well and an orbital averaging of 4 mm, with 5 seconds of orbital shaking prior to each measurement, and filters as close as possible to the wavelength specified in literature (see Table **??**). Fluorescence was thereafter imaged using a Safe Imager^{TM} Blue-Light

Transilluminator (Invitrogen) with an amber filter unit and a digital camera (Sony, DSC-RX100).

Heavy metal detection by TXO assays

Transcriptional regulators MerR (Q5NUU7), PbrR (Q58AJ5), CzcS (Q44007)/CzcR (Q44006), CnrH (P37978), ArsR (Q1LCN6) and CueR (A0A2L0XBJ2) were provided into the working reaction directly in the form of cell extract from *C. metallidurans* CH34 cells (30 %v/v). Reactions were supplemented with increasing concentrations (0.3 μM, 2 μM and 7 μM) of Hg^{2+}, Cu^{2+}, Co^{2+}, Ni^{2+}, Pb^{2+} and As^{3+}. Cell-free metal contact induction was performed in 384 Well Small Volume^{TM} LoBase Black Microplates (Greiner Bio-One) at 25°C, using DFHBI (10 μM) as fluorophore, and analysed in a FluoStar Omega (BMG Labtech) plate reader as described above.

8.4   Results

Results

Fluorogenic RNA aptamers and TXO systems

Commonly, cell-free systems are associated with the use of TX-TL systems instead of living cells for the production of valuable compounds or the generation of translational biosensors. Their dependence on translation, however, increases the complexity and cost of the minimal reaction mixture needed for them to work. In order to generate an alternative approach, we evaluated the use of multiple fluorescent RNA aptamers as possible output signals for the generation of transcription-only cell-free systems.

Table 8.1: RNA Aptamers and their respective characteristics

Aptamer	Emitted Colour	Size (bp)	Ex/Em (nm)	Fluorophore	Reference
Broccoli	Green	47	470/530	DFHBI	?
Broccoli2X	Green	157	470/530	DFHBI	?
Spinach2	Green	95	480/520	DFHBI	?
iSpinach*	Green	78	480/520	DFHBI	?
Corn	Orange	45	480/520	DFHO	?
2-4**	Cyan	100	400/460	DMABI	?
17-3***	Yellow	101	400/550	DMABI	?
6-8****	Red	101	400/600	2-HBI	?

Ex/Em extracted from respective Ref. Hereafter named *Green, **Cyan, ***Yellow, ****Red due to their protection within the F30 scaffold. Further fluorophores information available in Table S1

Eight previously characterised fluorescent aptamers were initially screened (see Table **??**). These were synthesised as 120 bp oligonucleotides and amplified in an overlapping PCR using the 5 bp that they share at their 3' end (see Material and Methods, Figure **??** and Table S2). Additionally, unprotected aptamers (Green, Cyan, Yellow, Red, Corn) were inserted within the F30 RNA scaffold, an engineered version of the naturally occurring phi-29 viral RNA three-way junction motif, not cleaved by cellular nucleases **?**. Additionally, BsmBI sites were added at both ends of each aptamer to allow insertion or replacement of the RNA aptamers via Golden Gate assembly if required in future applications (see Figure **??**). RNA expression and fluorescence detection were performed at room temperature (25°C) in order to simulate field conditions as closely as possible. Furthermore, in order to simplify their use we combined the expression and detection steps by using the novel OTDB buffer (see Materials and Methods), avoiding purification steps needed in some previous cell-free publications. Aptamers were selected primarily by difference in the colour emitted (green, red, orange, yellow, cyan), stability once protected within the F30 scaffold **?**, length, fluorophore needed, and expression level under our previously mentioned experimental conditions (see Table **??**).

The linear PCR-generated fragments can be immediately used for *in vitro* transcription and fluorescence detection. We tested the expression of nine different aptamers with five different fluorophores using a plate reader and an inexpensive blue light/red filter system (see Figure **??**). While previous studies regarded aptamers to be specific for a fluorophore (see Table 1), we have found that cross-reactivity is much higher. As shown in Figure **??**a, Spinach, Broccoli and their derivatives, engineered to bind DFHBI, were found to fluoresce with DMABI and DFHO with their respective colours. Broccoli2X was the most unspecific even increasing the fluorescence of ROX, which we were using as positive fluorescence control. On the other hand, DFHO was the most promiscuous fluorophore, which also displayed background fluorescence with the negative control. The aptamer fluorescence appeared very quickly after induction of transcription; after just a few minutes of incubation at room temperature, there was a significant difference between some aptamers and the negative control (see Figure **??**b). However, the fluorescent signal kept increasing for 100 minutes (Figure **??**b), indicating that the RNA polymerase remained active and the NTP pool was not exhausted for this period. Additionally, all F30::aptamers presented high stability over time when left on the bench at room temperature and without further protection.

Aptamer-fluorophore complexes were stable for weeks without bleaching or loss of fluorescence (See Figure S1).
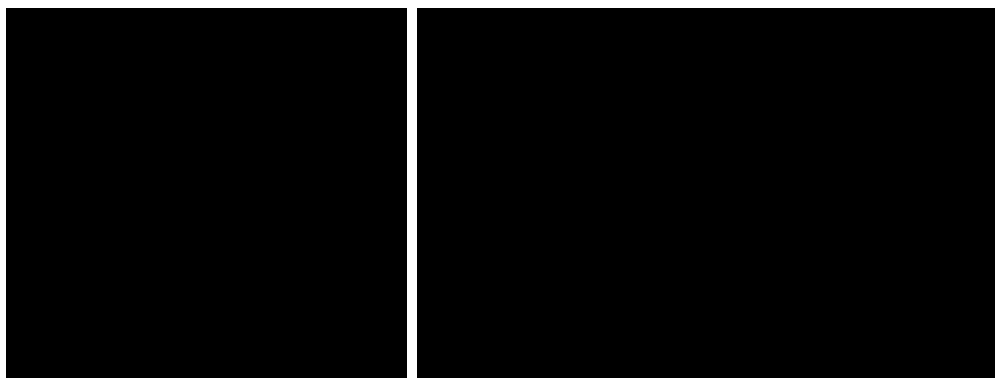
Figure 8.2: Fluorescent RNA aptamers expression. **(a)** Aptamers showing different colours under a blue light/red filter system after plate reader fluorescent measurement (16.5 h) and being observed by naked eye without the need for further equipment. A longer exposed image can be found in Supplementary Figure S1. **(b)** Expression of diverse RNA aptamers over time detected at 480/520 nm. Coloured shades represent the variability (MAD) of three independent tests.

TXO cell-free biosensors

Short aptamers such as the Green aptamer allow incorporation into transcriptional units with ease and simplicity (see Materials and Methods). Considering the stability and optimization of the Green aptamer (F30::iSpinach) for expression under *in vitro* conditions, further experiments were performed with this aptamer. The Green aptamer (F30::iSpinach) was amplified via PCR under the control of diverse heavy metal inducible promoters and the T7 polymersase promoter as control (P*t7*, P*cue*, P*ars*, P*mer*, P*pbr*, P*czc*, P*cnr*, see Material and Methods). Linear PCR products were used in reactions with different heavy metal concentrations (5 μM, 25 μM and 100 μM) with transcription factors and native RNA polymerase supplied from cell lysates (see Figure **??**). Using our TXO cell free system concentrations as low as 2 μM were detected for lead, mercury, zinc, and arsenite. Nickel, on the other hand, was detected at 0.3 μM. All of them reached a 2 fold-change in fluorescence after just 20 min of initiation of the transcription reaction (see Figure **??**) which kept increasing over time reaching after three hours values up to 10-times higher (see Figure S2).

Figure 8.3: Heavy metal detection by TXO Cell-free biosensors. Detection of Pb^{2+}, Hg^{2+}, Zn^{2+}, Ni^{2+}, Cu^{2+} and As^{3+} using the TXO reaction Cell-free system and the Green RNA aptamer cloned with each specified promoter (P*pbr*, P*mer*, P*cnr*, P*czc*, P*cop*, and P*ars*). Constructs were induced using increasing concentrations of the specified heavy-metal. Excitation/Emission of 480/520 nm. Data points at 20 minutes; the time course can be found in Supplementary Figure S2. Error bars represent the variability (MAD) of three independent tests.

8.5   Discussion / Conclusions

TXO systems maintain the advantages of cell-free TX-TL systems in terms of uniformity and avoidance of regulatory issues related to live GMOs, and further increase the simplicity and speed of response at the expense of reduced flexibility, since new proteins can not be generated within the system. Hence they are well suited to applications where simplicity and rapid response are paramount; for example, biosensors and diagnostics for use in field and point-of-care applications. Lack of translational capability means that reporters for system output must be in the form of nucleic acids. We tested eight previously described fluorescent aptamer-fluorophore systems and found that all could generate visible fluorescent responses in an optimized buffer system as a result of transcription with no further processing. Several of the aptamers showed the ability to bind a wider range of fluorophores than what has previously been reported; for example, Broccoli2X binds 4 out of the 5 fluorophores analysed (Figure 2a and S1), surprisingly including ROX (5-Carboxy-X-rhodamine N-succinimidyl ester) commonly used for basal line generation during routine fluorescent assays, such as qPCR. Using TXO systems fluorescent responses can be generated in a matter of minutes, as compared to an hour or more for reported TX-TL systems based on generation of fluorescent proteins or enzymes with chromogenic substrates **??????**. Thus we conclude that aptamer-fluorophore complexes are a suitable output for TXO synthetic biology systems.

A further advantage of TXO systems is simplicity, meaning that they can

be very rapidly generated and tested, even more so than TX-TL systems **?**. The aptamer-based reporter genes are small enough that they can be easily generated by annealing and extension of oligonucleotides (see Figure **??**). PCR was used to add metal-responsive promoters to DNA encoding the F30::iSpinach reporter ('Green'), and the PCR products were used directly for assays, with the necessary transcription factors being supplied by an unpurified lysate from *C. metallidurans* CH34. Fluorescent response to metals was observed within minutes, with detection limits in these unoptimized systems below 0.3 μM (18 ppb) for nickel, and below 2 μM for copper (127 ppb), zinc (131 ppb), arsenic(III) (150 ppb), mercury (403 ppb) and lead (414 ppb). No response was observed for arsenic(V), presumably because the system lacked the reducing power necessary to reduce it to arsenic(III), the form actually detected by the ArsR repressor (Data not shown). Further optimization should reduce these limits of detection. Previous works with RNA aptamers for metal sensing, DasGupta et al. mentioned the use of a truncated form of the Spinach aptamer as a Pb^{2+} sensor, but this could have lead to false results at concentrations <2 ppm due to a lack of sensitivity and a weak output response **?**.

Since TXO systems are considerably simpler than TX-TL systems, with no requirement for ribosomes, tRNA, or associated enzymes, we expect that it will be possible to freeze-dry such systems for storage and distribution as previously reported for TX-TL systems **??**. Fluorescent signals may be further strengthened using alternative fluorophores such as DHFBI-1T **?**. The visibly different colours, as well as different excitation and emission wavelengths for different aptamer-fluorophore complexes, raise the possibility of multiplexed assays as well as more complex approaches such as FRET and quenching. While this publication focuses on the use of aptamers as signal output in a TXO ststem, future research could exploit the advantages of translation independent systems introducing other known functions of RNA molecules. Examples include RNA-based molecule detection (e.g. using riboswitch aptamers) and signal processing (e.g. using RNA-RNA hybridisation networks or crRNA/gRNA) **??????**. Possible improvements to the already existing technology could also imply the use of other RNA aptamers showing different properties, for instance, by improving the structure and conditions of other aptamers published in the original work by Paige *et al*, which emit fluorescence in different wavelength, i.e. red, yellow, cyan (see Table

?? and Figure **??**) or by using ribozymes, such as PS2.M, allowing additional possible transcriptional outputs and broadening the possibilities of TXO systems.

Additionally, TXO systems can be made with no need for GMOs, using in vitro DNA manipulation and non-ionic detergent lysis of wild type ACDPI or Biosafety level 1 cells, plus non toxic chemicals and commercially available enzymes/proteins, so procedures could easily be done in high schools or even at home, leading to a democratization of synthetic biology without the worrying risks of people making bioweapons or causing ecological disasters. Considering that algorithms are simply a sequence of instructions, typically to solve a class of problems or perform a computation, TXO systems offer a step forward in terms of simplicity, ease of implementation, and rapid response, and may be widely useful in synthetic biology for speeding up and simplifying genetic circuit algorithm design and response.

8.6   Data Availability

All datasets generated for this study are included in the manuscript and/or the supplementary files. More Data is available at

8.7   Acknowledgments

8.8   Supplementary Material

Table S1: Fluorophores information

Abbreviation	Formula	Chemical name
DFHBI	$C_{12}H_{10}F_2N_2O_2$	3,5-difluoro-4-hydroxybenzylidene imidazolinone
DMABI*	$C_{14}H_{17}N_3O_1$	4-dimethylaminobenzylidene imidazolinone
2-HBI*	$C_{12}H_{12}N_2O_2$	2-hydroxybenzlidene imidazolinone
DFHO	$C_{12}H_9F_2N_3O_3$	3,5-difluoro-4-hydroxybenzylidene imidazolinone-2-oxime
ROX	$C_{37}H_{33}N_3O_7$	5-Carboxy-X-rhodamine N-succinimidyl ester

provided by Dr. Sammie R. Jaffrey's laboratory

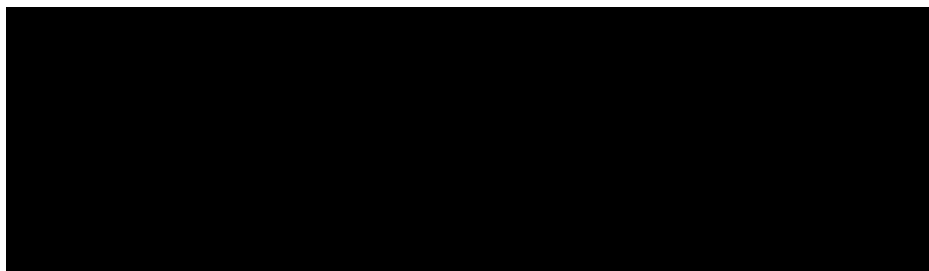**Figure S1:** Fluorescent RNA aptamers observed under a blue light/red filter system. Aptamers showing different colours under a blue light/red filter system after measurement 16.5 h and being observed by naked eye without the need for further equipment. Picture taken with a Sony, DSC-RX100 camera and aperture
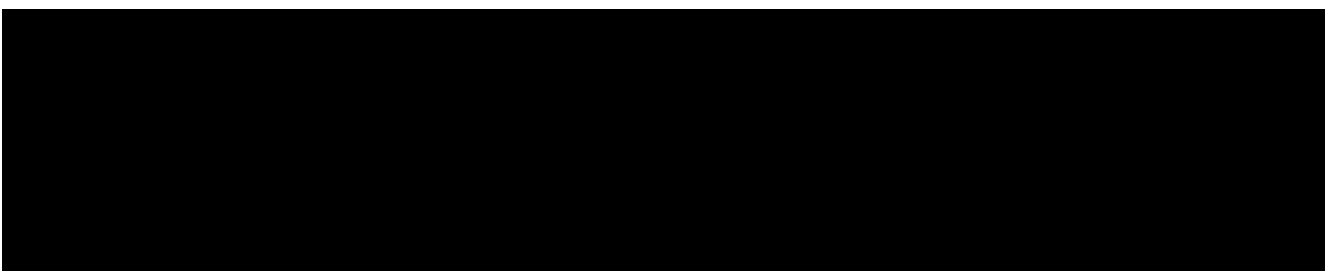

Figure S2: TXO sensors over time with different metals. Coloured shades represent variability of three independent tests.

Table S2: Oligonucleotides used

Oligo	Name	Sequence
F240	$\text{Cyan}_{F30::2-4}Fw$	tgtgggagacgcaactgaatgaacctagagttatgccaggctctgagcctgcttcggcaggtgctatgatcgccagcggtat gcagtccgtaactagtcgcgtctcc
F241	$\text{Cyan}_{F30::2-4}Rv$	gtggggagacgcgactagttacggactgcataccgctggcgatcatagcacctgccgaagcaggctcagagcctggcataa ctctaggttcattcagttgcgtctcc
F242	$\text{Red}_{F30::6-8}Fw$	tgtgggagacgcaactgaatgaaaatggcaaaatattcgagaagctggtctgcttcggcaggattctccaaggggtagatc gtgtattccgtaactagtcgcgtctcc
F243	$\text{Red}_{F30::6-8}Rv$	gtggggagacgcgactagttacggaatacacgatctaccccttggagaatcctgccgaagcagaccagcttctcgaatattt tgccattttcattcagttgcgtctcc
F244	$\text{Yellow}_{F30::17-3}Fw$	tgtgggagacgcaactgaatgaagagcagtagcgagtagttcacaagagctgcttcggcaggatcttgtaggaagtaaatg tgcaaatccgtaactagtcgcgtctcc
F245	$\text{Yellow}_{F30::17-3}Rv$	gtggggagacgcgactagttacggatttgcacatttacttcctacaagatcctgccgaagcagctcttgtgaactactcgcta ctgctcttcattcagttgcgtctcc
F296	$\text{Green}_{F30::iSpinach}Fw$	gaattcgcggccgcttctagagtaatacgactcactatagggttgccatgtgtatgtgggagacgcgactacggtgagggtc gggtccagtagcttcggctactgttgagtagagtgtgg
F297	$\text{Green}_{F30::iSpinach}Rv$	ctgcagcggccgctactagtacccctcaagacccgtttagaggccccaaggggttatttgccatgaatgatcccgaaggatc atcagagtatgtggggagacgtacggagcccaca
F310	P_{pbr}::Green Fw	gaattcgcggccgcttctagagggcgtcggatgggagatgtcttgactctatagtaactagagggtgttaaatcggcaactt gccatgtgtatgtgggagacgcgactacggtgagggtc
F311	P_{mer}::Green Fw	gaattcgcggccgcttctagagatcgcttgactccgtacatgagtacggaagtaaggttacgctatccaatttcaattcgaat tgccatgtg
F312	P_{cop}::Green Fw	gaattcgcggccgcttctagagaatttcttgaccttccccttgctggaaggtttaacctttatcacattgccatgtgtatgtgg gagacgcg
F313	P_{ars}::Green Fw	gaattcgcggccgcttctagaggtatatacacattcgttaagtcatatatgttttttgacttatccgcttcgaagagagacact acctgcaacttgccatgtgtatgtgggagacgcgact
F330	$\text{Corn}_{F30::Corn}Fw$	gatccttcgggatcattcatggcaaataaccccttggggcctctaaacgggtcttgaggggtactagtagcggccgctgcag tactata
F331	$\text{Corn}_{F30::Corn}Rv$	ctcctcagacctccttcctcgcgccgtctcccacatacacatggcaaccctatagtgagtcgtattactctagaagcggccgc gaattc
F404	P_{czcN}::GreenFw	gaattcgcggccgcttctagagggagggcgtctctgggtgtgtgctgaaaatggccaagacagtctatgtcccagaagatg actgtcagattgccgagctttgccatgtgtatgtgggag
F405	P_{cnr}::GreenFw	gaattcgcggccgcttctagagggaggcctgaagccggaacatcgacctgcttacgatcgcgttcttatcgatgcacttgcc atgtgtatgtgggagacg
F408	PS2M_Fw	gaattcgcggccgcttctagagtaatacgactcactatagggggtgggtagggcgggttggataaccccttggggcctctaa acgggtcttgaggggtactagtagcggccgctgcagtat
F409	PS2M_Rv	atactgcagcggccgctactagtacccctcaagacccgtttagaggcccaaggggttatccaacccgccctacccaccccc tatagtgagtcgtattactctagaagcggccgcgaattc
F296-F297	BBK::P_{T7}::Green ::T_{T7}::BBK	gaattcgcggccgcttctagagtaatacgactcactatagggttgccatgtgtatgtgggagacgcgactacggtgagggtc gggtccagtagcttcggctactgttgagtagagtgtgggctccgtagtcgcgtctccccacatactctgatgatccttcg ggatcattcatggcaaataaccccttggggcctctaaacgggtcttgaggggtactagtagcggccgctgcag
F313-F297	BBK::P_{ars}::Greeen ::T_{T7}::BBK	gaattcgcggccgcttctagaggtatatacacattcgttaagtcatatatgttttttgacttatccgcttcgaagagagacac tacctgcaacttgccatgtgtatgtgggagacgcgactacggtgagggtcgggtccagtagcttcggctactgttgagtag agtgtgggctccgtagtcgcgtctccccacatactctgatgatccttcgggatcattcatggcaaataaccccttggggcct ctaaacgggtcttgaggggtactagtagcggccgctgcag
F405-F297	BBK::P_{cnr}::Green ::T_{T7}::BBK	gaattcgcggccgcttctagagggaggcctgaagccggaacatcgacctgcttacgatcgcgttcttatcgatgcacttgc catgtgtatgtgggagacgcgactacggtgagggtcgggtccagtagcttcggctactgttgagtagagtgtgggctccg tagtcgcgtctccccacatactctgatgatccttcgggatcattcatggcaaataaccccttggggcctctaaacgggtctt gaggggtactagtagcggccgctgcag
F312-F297	BBK::P_{cop}::Green ::T_{T7}::BBK	gaattcgcggccgcttctagagaatttcttgaccttccccttgctggaaggtttaacctttatcacattgccatgtgtatgtg ggagacgcgactacggtgagggtcgggtccagtagcttcggctactgttgagtagagtgtgggctccgtagtcgcgtctcc ccacatactctgatgatccttcgggatcattcatggcaaataaccccttggggcctctaaacgggtcttgaggggtactagt agcggccgctgcag

Oligo	Name	Sequence
F404-F297	BBK::P_{czc}::Green ::T_{T7}::BBK	gaattcgcggccgcttctagagggagggcgtctctgggtgtgtgctgaaaatggccaagacagtctatgtcccagaagatg actgtcagattgccgagctttgccatgtgtatgtgggagacgcgactacggtgagggtcgggtccagtagcttcggctactg ttgagtagagtgtgggctccgtagtcgcgtctccccacatactctgatgatccttcgggatcattcatggcaaataacccctt ggggcctctaaacgggtcttgaggggtactagtagcggccgctgcag
F311-F297	BBK::P_{mer}::Green ::T_{T7}::BBK	gaattcgcggccgcttctagagatcgcttgactccgtacatgagtacggaagtaaggttacgctatccaatttcaattcgaa ttgccatgtgtatgtgggagacgcgactacggtgagggtcgggtccagtagcttcggctactgttgagtagagtgtgggctc cgtagtcgcgtctccccacatactctgatgatccttcgggatcattcatggcaaataacccctgggggcctctaaacgggtct tgaggggtactagtagcggccgctgcag
F310-F297	BBK::P_{pbr}::Green ::T_{T7}::BBK	gaattcgcggccgcttctagagggcgtcggatgggagatgtcttgactctatagtaactagagggtgttaaatcggcaactt gccatgtgtatgtgggagacgcgactacggtgagggtcgggtccagtagcttcggctactgttgagtagagtgtgggctccg tagtcgcgtctccccacatactctgatgatccttcgggatcattcatggcaaataacccctgggggcctctaaacgggtcttga ggggtactagtagcggccgctgcag

Chapter 9

Cupria, a novel cell-free system for industrial biotechnology

9.1   Abstract

Cell-free systems (CFS) are an emerging technology where cell extracts are used instead of living genetically-modified cells, thus, not presenting a "living prospect" applicable to current legal regulations. Here we show a versatile cell-free system based on the master survivalist bacteria *Cupriavidus metallidurans* CH34, able of not only sensing environmental variables, such as heavy metals, but also synthesize proteins, produce bioplastics and degrade toxic aromatic compounds. This novel cell-free chassis could become a standard for mass production considering the versatility and robustness showed once compared with standard *E. coli* cell-free systems.

9.2   Introduction

While cell based systems have to be a self-replicating chassis as its greater advantage, cell-free expression systems are a powerful tool in applications where cell based expression systems are too variable, slow, difficult to store or prohibitive due to legislation on the release of genetically modified organisms.

Since interaction with the surrounding environment is direct

As their very nature is to be a defined, precisely controlled system, their lack of a self-replicating functionality, allows them to function in a limited supply of intracellular supplies such as ATP sources, amino acids and nucleotides. In consequence in vitro protein synthesis is limited by two dimensions: the overall maximum protein amount that can be synthesized with the supplied components, and the time during which the cell free expression system stays active

before background processes and chemical deterioration have used up supplies or inactivated the system (**???**.

9.3   Methodology

Strains and Cell free extract preparation

E. coli MG1655 and *C. metallidurans* CH34 strains were grown on Luria Broth (LB) at 37°C and 30°C respectively, both under constant shaking at 180 RPM. Cultures were induced with the respective heavy metal at middle exponential phase and cells collected by centrifugation at 10.000 RPM, 4 C for 5 min. Resuspension was performed using Tris-HCl (40 mM) pH 7.5, followed by a nonionic lysis step using Triton X-100 (0.8%), Tris-HCl (40 mM) pH 7.5 and lysosime 50 mg/mL. Lysed was centrifuged at 10.000 RPM, 4°C for 10 min and supernatant was saved at -80°C or immediately used.

DNA constructs generation

Promoters from clusters *copAB*, *arsBCD* from E. coli MG1655 and *merTPAD*, *pbrAB* from C. metallidurans CH34 were added by PCR at the 5' of the pTUB[F30::iSpinach] aptamers using the Q5 High-Fidelity DNA Polymerase (NEB). Transcriptional regulators MerR, PbrR, ArsR and CueR were provided in to the working reaction directly from the cell extract. pBP, pTUB (Moore et al., 2016), pSB1, pSB4 (iGEM collection), pGEM (Promega), pET28, pET31 (Paige et al., 2011), ptZ57R/T (Autour et al., 2016), pMO (Kuehl et al., 2014), pSEVA (Silva-Rocha et al., 2013) plasmids were used as backbones. All constructs generated were confirmed by Sanger sequencing(Edinburgh Genomics). Primer sequences are listed in Appendix 2. Plasmids were digested using appropriate restriction enzymes (NEB) and excised from agarose 1.3 - 2% gel, and recovered using a Quiaquick gel extraction kit (Quiagen). Ligation steps were performed using T4 DNA ligase and 30 fmol of plasmid per 90 fmol of insert (1:3). Ligations were used to transform chemically competent E. coli JM109 cells selected on the appropiate antibiotic, plasmid DNA was prepared using a QIAprep Spin Miniprep Kit (Qiagen) and constructs further confirmed by PCR and Sanger sequencing. Constructs for expressing crRNA (Figure 10) and the F30 scaffold were synthesized (Sigma-Aldrich) and cloned in to pBP (Figure 7). All RNA aptamers generated by Paige et al., (2011) and Autour (2017), except Spinach,

were cloned by Golden gate assembly in pTUB flanked by an F30 scaffold (Figure 8). Nullomers 2 (5' GTCCGAGCGTA 3') and 5 (5' ATGTTGATGTTGT 3') (see Table 6) were cloned flanking the F30::iSPinach construct in the pGEMTeasy by TA cloning (Figure 11 and 12), using Platinum Taq polymerase High Fidelity (ThermoFisher). Promoters from the copAB and arsBCD operons from E. coli MG1655 and merTPAD, pbrAB operons from C. metallidurans CH34 were added by PCR on the 5' end of the F30::Broccoli (pET28C) and F30::iSpinach (pTUB) aptamers using Q5 High-Fidelity DNA Polymerase (NEB). Plasmids were transformed into chemically competent E. coli as prepared by the standard TSS method (Chung et al., 2014).

Expression and detection of Fluorescent RNA aptamers.

Although there are standard protocols for in vitro transcription of fluorescent RNA aptamers, these include purification and concentration steps that complicate the entire procedure. A combination of both main steps was achieved at room temperature (25°C) using our Optimized Transcription/Detection buffer (OTDB): Tris-HCl (40 mM) pH 7.5; MgCl2*6(H2O) (6 mM); DTT (1 mM); Spermidine (2 mM); NTPs (2 mM each); DFHBI (20 uM); NEB T7 RNA polymerase (1.5 U/uL); Template DNA (650 nM). Fluorescence was measured using a FluoStar Omega (BMG Labtech) plate reader, excitation 480 nm/ emission 520 nm filters (10 mm bandpass), using 20 flashes per well and an orbital averaging of 4 mm, with 5 seconds of orbital shaking prior each measurement.

Heavy metal quantification by fluorescence detection

Stock solutions (100 mM) of HgCl 2 (Sigma-Aldrich), CuSO 4 • 5H 2 O (Sigma-Aldrich), ZnSO 4 • 7H 2 O (Sigma-Aldrich), CdSO 4 (Sigma-Aldrich), Na 2 HAsO 4 • 7H 2 O (Sigma-Aldrich), Pb(NO 3) 2 (VWR), NaAsO 2 (HoneyWell) were prepared and used according to their free metal ion concentration (Table 1). Table 1: Metal stock solution with their free ion concentrations. Metal solution Concentration ppm Free Ion Concentration ppm HgCl 2 100 mM 27152 Hg +2 100 mM 20059 CuSO 4 •5H 2 O 100 mM 24969 Cu +2 100 mM 6355 Pb(NO 3) 2 100 mM 33120 Pb +2 100 mM 20720 ZnSO 4 •7H 2 O 100 mM 28756 Zn +2 100 mM 6538 CdSO 4 100 mM 20487 Cd +2 100 mM 11241 Na 2 HAsO 4 •7H 2 O 100 mM 31201 As +5 100 mM 7492 NaAsO2 100 mM 12991 As +3 100 mM 7492

9.4   Results

9.5   Discussion / Conclusions

Contamination of the environment with heavy metals is very widespread in many regions. Standard analytical techniques such as ICP-MS and ICP-OES are highly accurate but require expensive equipment and highly trained staff. Biosensors offer potential advantages for sensitive, inexpensive field testing, but those based on live genetically modified cells may be difficult to commercialize for regulatory reasons (French et al., 2011). Furthermore, as biosensors depend on living organisms, their usability is affected in real world applications where contamination is widespread and more than one toxic compound is present, i.e. mining soils that are contaminated with extremely high heavy metal concentrations and multiple aromatic compounds derived from petrol usage. Cupriavidus metallidurans is an interesting organism for use in this kind of complex environment (Millacura et al., 2017; Rojas et al 2011), but its use is limited by the same legal regulations as other biosensors. In this thesis, a cell free extract from this organism will be tested for its possible use in genetic circuit design, capable of function on complex environments, granting robustness to the system. In order to improve the analysis of complex environments, genetic circuits will be generated to simultaneously analyze the multiple variables present, being the parallel approach a promising design for decompose the environmental samples into their constituent parts. The modular approach plus the possible robustness of this system, will allow us to obtain an adaptable system that would not rely on the normal cellprocessing/synthesis machinery (interaction between transcription factors, polymerases, ribosomes, and other diverse macromolecules), complying with existing and near-future regulations and which can thus be easily moved into field trials and potentially commercialized.

9.6   Data Availability

All datasets generated for this study are included in the manuscript and/or the supplementary files. More Data is available at https://doi.org/10.7488/ds/2497

9.7 Acknowledgments

9.8 Supplementary Material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fbioe.2019.00046/full#supplementary-material

Chapter 10

Discussion / Conclusions

Here you finish your thesis as a king/queen and drop the microphone LOL

Chapter 11

Appendix

ParAlleL: A Novel Population-Based Approach to Biological Logic Gates

Felipe A. Millacura*, Brendan Largey and Christopher E. French*

School of Biological Sciences, Institute of Quantitative Biology, Biochemistry and Biotechnology, University of Edinburgh, Edinburgh, United Kingdom

In vivo logic gates have proven difficult to combine into larger devices. Our cell-based logic system, ParAlleL, decomposes a large circuit into a collection of small subcircuits working in parallel, each subcircuit responding to a different combination of inputs. A final global output is then generated by a combination of the responses. Using ParAlleL, for the first time a completely functional 3-bit full adder and full subtractor were generated using *Escherichia coli* cells, as well as a calculator-style display that shows a numeric result, from 0 to 7, when the proper 3 bit binary inputs are introduced into the system. ParAlleL demonstrates the use of a parallel approach for the design of cell-based logic gates that facilitates the generation and analysis of complex processes, without the need for complex genetic engineering.

Keywords: *Escherichia coli*, 3-bits, full adder, full subtractor, calculator-like display, parallel approach

INTRODUCTION

A major challenge in the field of synthetic biology is the construction of complex logic circuits that analyze variables as in electronics; where a single circuit accepts one or more binary inputs to generate one or more binary outputs. A cell-based logic network consists of engineered cells producing an output macromolecule only if the corresponding pattern of inputs is present. The mechanism of analysis is commonly based on the use of transcriptional regulators, transcription factors, polymerases, receptors, or recombinases (Brenner et al., 2018). Some examples of genetic circuits mimicking computational behavior are toggle switches, oscillators, boolean logic gates, feedback controllers, and multiplexers. Although there are genetic circuits that simulate computational behavior, the complex engineering of their biological chassis is affected by gene expression noise, mutation, cell death, undefined, and changing extracellular environments and improper interactions with the cellular context (Andrianantoandro et al., 2006). Furthermore, complex genetic engineering is necessary when multiple input variables are analyzed, limiting the processing capacity of the system.

Biological multiplexers analyze one or more signals over a common transmission line using interconnected transcription factors, recombinases, antisense RNA, or CRISPR-like technology (Nielsen and Voigt, 2014; Roquet et al., 2016; Brenner et al., 2018). However, complex genetic engineering is needed for wiring the basic computational units, becoming inefficient for moving beyond simple NOT or AND logic gates or for scaling to 3 bit logic circuits. The complexity of the genetic engineering required can be reduced by using distributed logic circuits, where the computation is distributed among several physically separated cellular consortia that each sense only one signal and respond by secreting a communication molecule (Regot et al., 2011). As a circuit responds to one signal, but not another, due to spatial distribution, a change in the state of the system can be triggered as response, making synthetic learning possible

(Macia et al., 2017; Shipman et al., 2017). Even though the consortium approach makes Boolean circuit design simpler, it still shows a slow response and considerable complexity since each cell needs to recognize, synthesize and secrete a wiring molecule (Macia et al., 2016).

Here we propose an alternative logic architecture, which decomposes a large circuit into a collection of small subcircuits acting in parallel (hereafter ParAlleL). Rather than having a single type of agent (such as a genetically engineered cell) doing the computation, ParAlleL has separate types of agent that each react to a different combination of inputs. A final output is then generated by combination of the responses, making all kinds of binary operation possible. As an example, here we show the implementation of this concept using cells resistant to different combinations of antibiotics, with the response indicated by growth. This is used to demonstrate a completely functional 3 bit full adder and full subtractor, as well as a calculator-style display that shows digits from 0 to 7 based on three binary input bits.

METHODOLOGY

Reagents and Stock Solution Preparations

Antibiotic stock solutions were prepared as follows: 100 mg/ml carbenicilin disodium salt (Sigma-Aldrich #C1389), 50 mg/ml kanamycin sulfate (PanReac Applichem #A1493), 20 mg/ml chloramphenicol (Acros Organics #22792), 10 mg/ml tetracycline hydrochloryde (Duchefa Biochemie #T0150), 10 mg/ml gentamicin sulfate (Melford #G0124), and 50 mg/ml spectinomycin.HCl (LKT Labs #S6018). Developing solution contained 0.1 %w/v bromothymol blue (Sigma-Aldrich #114421) and 400 mM Trizma base pH7.5 (Sigma-Aldrich #T1503).

Generation of Subcircuit Cells

E. coli JM109 was transformed with 200–300 pg of plasmid pSB4A5 (AmpR) or pSB4C5 (ChlR) (Registry of Standard Biological Parts) and selected on 100 μg/ml carbenicilin (Am) or 20 μg/ml chloramphenicol (Ch), respectively. Cells carrying the first bit plasmid were made chemically competent (Chung et al., 1989) and transformed with 200–300 pg of the 2nd bit plasmid, pSB1T3 (TetR) or pSB1K3 (KanR) (Registry of Standard Biological Parts). Selection was performed with the first antibiotic (Am or Ch) and the addition of 10 μg/ml Tetracycline (Tc) or 50 μg/ml kanamycin (Km), obtaining the two-bit combinations Km/Am (KA), Tc/Am (TA), Km/Ch (KC), and Tc/Ch (TC). This set of strains is sufficient to implement all two-bit binary operations.

The third bit layer was generated by transforming these four strains with pSEVA631 (GenR) (Silva-Rocha et al., 2012; GenBank JX560348) or pMO9075 (SpmR) (Keller et al., 2011). Resulting strains were selected on the 2-bit antibiotic combinations plus 10 μg/ml gentamicin (Gm) or 50 μg/ml spectinomycin (Sm). This gave 8 strains, designated ATG (Am/Tc/Gm), AKG (Am/Km/Gm), ATS (Am/Tc/Sm), AKS (Am/Km/Sm), CTG (Ch/Tc/Gm), CTS (Ch/Tc/Sm), CKG (Ch/Km/Gm), CKS (Ch/Km/Sm) based on their resistance markers. This set of strains is sufficient to implement all three-bit binary operations. Plasmid specifications are listed in **Figure S2**, **Tables S1, S3**, with further information about these antibiotics in **Table S2**. Plasmid sequences are available in different formats at https://doi.org/10.7488/ds/2497.

Three-Bit Logic Operations

Tests were performed in 96-well microplates by inoculating cells (1:100) in LB broth (100 μL) supplemented with 1%w/v D(+)-glucose (Fisher Chemical #G0500). Plates were incubated for 18 h at 37°C without shaking and then developed by addition of the developing solution (0.1%w/v bromothymol blue in 400 mM

FIGURE 1 | ParAlleL responding to 1 bit, 2 bit, and 3 bit inputs. ParAlleL subcircuit cells were spatially distributed in different wells (vertically) and exposed to specified 1 bit, 2 bit, or 3 bit inputs (top of each column). Cells were inoculated (1:100) in LB supplemented with 1% w/v glucose. After 18 h of incubation at 37°C, plates were developed by addition of 0.05 volumes of the developing solution.
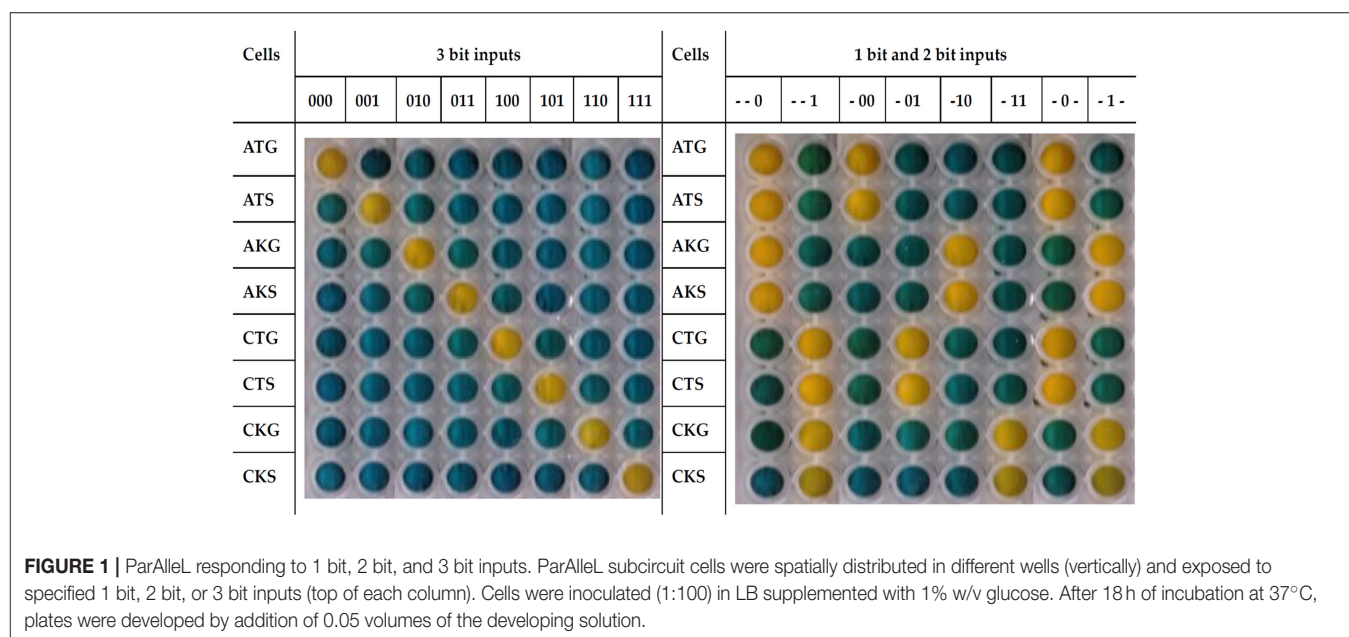
A Calculator-like display design

A	B	C
D	---	E
F	G	H
I	---	J
K	L	M

No. to display	A	B	C	D	E	F	G	H	I	J	K	L	M
000 (0)	ATG	ATG	ATG	ATG	ATG	ATG	----	ATG	ATG	ATG	ATG	ATG	ATG
001 (1)	----	----	ATS	----	ATS	----	----	ATS	----	ATS	----	----	ATS
010 (2)	AKG	AKG	AKG	----	AKG	AKG	AKG	AKG	AKG	----	AKG	AKG	AKG
011 (3)	AKS	AKS	AKS	----	AKS	AKS	AKS	AKS	----	AKS	AKS	AKS	AKS
100 (4)	CTG	----	CTG	CTG	CTG	CTG	CTG	CTG	----	CTG	----	----	CTG
101 (5)	CTS	CTS	CTS	CTS	----	CTS	CTS	CTS	----	CTS	CTS	CTS	CTS
110 (6)	CKG	CKG	CKG	CKG	----	CKG	CKG	CKG	CKG	CKG	CKG	CKG	CKG
111 (7)	CKS	CKS	CKS	----	CKS	----	----	CKS	----	CKS	----	----	CKS

B ParAlleL 3 bit display outputs

ATG (0) 000   ATS (1) 001   AKG(2) 010   AKS(3) 011

CTG (4) 100   CTS (5) 101   CKG(6) 110   CKS(7) 111
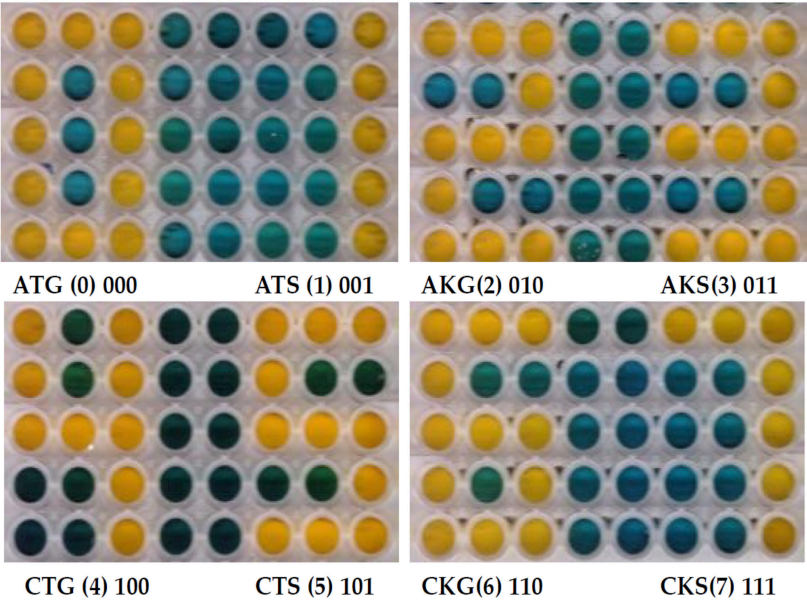
FIGURE 2 | Digital calculator-like display using 3 bit ParAlleL. Figure shows all numerals from zero to seven based on the 8 binary inputs provided. **(A)** Subcircuit cells were mixed and distributed in a 3 × 5 matrix and inoculated (1:100) in LB supplemented with 1%w/v glucose. Plate was developed after 18 h of incubation, by addition of 0.05 volumes of developing solution. **(B)** Output number results obtained by addition of each 3 bit antibiotic combination.

Tris, pH7.5) in a ratio 1:20. Images were obtained using a Kodak ESPC315 Flatbed scanner. Design of the calculator-like display, full adder, and subtractor are shown in **Figures 2**, **3** and in Supplementary material (**Figure S3**). Raw figures were deposited at https://doi.org/10.7488/ds/2497.

RESULTS

In the distributed logic system of ParAlleL each input bit has two forms, ZERO and ONE, each of which is essential to certain output agents and inhibitory to others. Thus each agent reacts only to a certain combination of input bits, allowing generation of any arbitrary pattern of outputs for any pattern of inputs. In the implementation shown here, each input bit comes in two forms, each being an antibiotic lethal to sensitive strains. In this case, bit A is represented by ampicillin for zero, chloramphenicol for one, bit B by kanamycin for zero, tetracycline for one, and bit C by gentamicin for zero, spectinomycin for one. Thus, four strains

are needed to implement any operation with two input bits, and eight strains for three input bits. In contrast to other cell-based logic schemes, only very minimal genetic engineering is required, essentially transformation with 3 different antibiotic resistance markers.

Cells show a global response concordant with the behavior expected for a 1 bit, 2 bit, or 3 bit system (**Figure 1**). For instance, when the input 101 (chloramphenicol, tetracycline and spectinomycin) is added to the system growth is only observed in the corresponding CTS cells, which carry the proper resistance markers. The response time of the system is around 12 h (**Figure S1**) but plates were developed at 18 h to avoid false negatives or positives.

In order to further test the ParAlleL system, a digital calculator-like display was designed (**Figure 2A**). In this case, multiple subcircuit cells are mixed in one well and the global response displays a number from 0 to 7 when the proper binary input is applied. Numbers represent the total eight possible values encoded within 3-bit binary inputs.
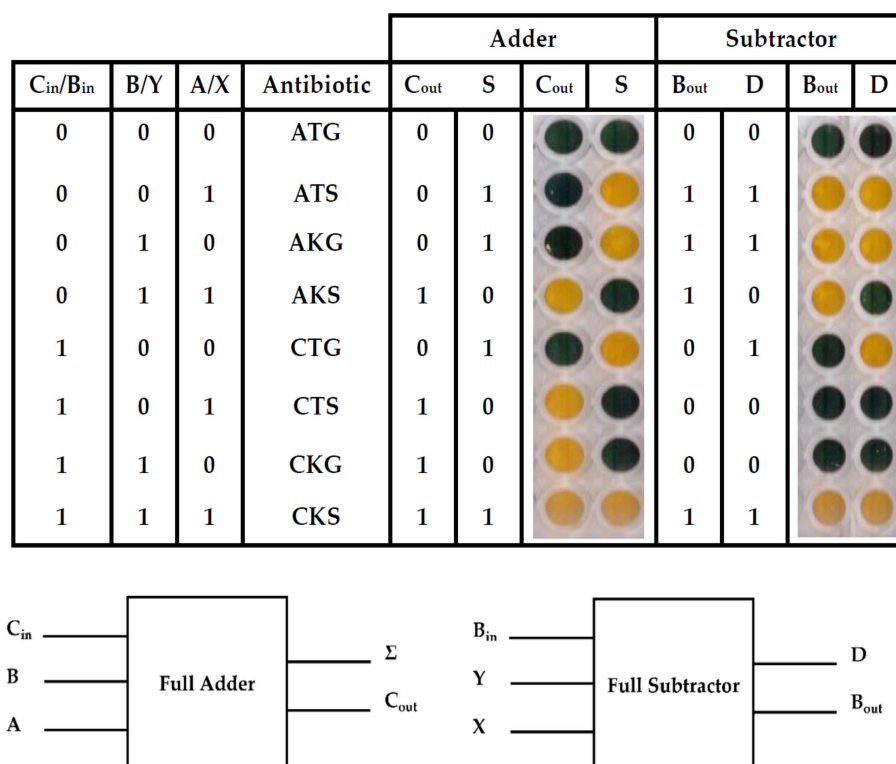
C_{in}/B_{in}	B/Y	A/X	Antibiotic	Adder				Subtractor			
				C_{out}	S	C_{out}	S	B_{out}	D	B_{out}	D
0	0	0	ATG	0	0			0	0		
0	0	1	ATS	0	1			1	1		
0	1	0	AKG	0	1			1	1		
0	1	1	AKS	1	0			1	0		
1	0	0	CTG	0	1			0	1		
1	0	1	CTS	1	0			0	0		
1	1	0	CKG	1	0			0	0		
1	1	1	CKS	1	1			1	1		

FIGURE 3 | Full adder and subtractor using the 3 bit ParAlleL system. The figure shows results of addition and subtraction using the ParAlleL for 3 bit system. Cells were mixed as shown in **Figure S3** and inoculated (1:100) in LB supplemented with 1%w/v glucose. After 18 hours of incubation, the plate was developed by addition of the 0.05 volumes of developing solution.

Input configuration versatility was proven by representing bit A in this case by gentamicin for zero, spectinomycin for one, bit B by kanamycin for zero, tetracycline for one, and bit C by ampicillin for zero, chloramphenicol for one. For instance, once input 110 represented by Ch/Km/Gm is added in the system, the number 6 is displayed (**Figure 2**).

Finally, a full adder and a full subtractor were designed. A full-adder adds three binary inputs, often denoted as A, B, and C_{in}, generating a Sum result (S) and a Carry-out (C_{out}). A full subtractor, on the other hand, has a minuend (X), a subtrahend (Y) and an additional Borrow-in (B_{in}) as inputs. The subtraction operation produces a difference (D) and a Borrow (B_{out}) (**Figure 3**).

In order to generate the full adder and subtractor, multiple subcircuit cells were mixed and distributed in two different wells (**Figure S3**). One well represents the solution (S) or difference (D) and a second one the carry (C_{out}) or borrow (B_{out}), for the adder and subtractor, respectively (**Figure 3**). A yellow color represents growth and a positive output 1, a blue color represents no growth and a binary 0 output instead.

CONCLUSIONS/DISCUSSION

Subcircuits that solve complex calculations in parallel have been extensively used for computation in order to reduce the total computation time. Translating this approach to biological

systems would allow us to analyze complex processes, currently difficult in synthetic biology, as multiple simple sub-circuits.

In our proof of concept, we present a biological information processing system, ParAlleL, capable of exploiting the parallelism in mixed bacterial cultures. ParAlleL decomposes the analysis of 2 and 3 bit complex inputs, into 4 and 8 sub circuits, respectively (**Figure 1**). Each sub-circuit corresponds to a different *E. coli* strain carrying a different combination of antibiotic resistance markers (**Table S1**). As an example, in the 3 bit system the input 000 is represented by the antibiotics ampicillin, tetracycline and gentamicin (**Figure 1**). When this input is entered into the system, all cells that are not encoded for responding to 000 will die, but cells carrying the proper plasmid combination, pSB4A5, pSB1T3, and pSEVA621 will not (**Figures 1**, **2**), therefore, a live/dead response (output) is achieved in all sub circuits, the output of each well being one (growth) or zero (failure to grow) (**Figure 2** and **Figure S2**).

ParAlleL uses cellular consortia instead of a single type of cell. A similar approach has been developed by Macia et al. (2016, 2017) using eukaryotic cells, and even showing the possibility of generating transient memory. However, that approach requires a sophisticated design as it relies on a secreted intermediate molecule (hormone-like) that must be kept at the right production level, and that should be previously activated by X (Repressor) and Y (SsrA-tagged protein) degradation. Furthermore, since the output of the circuit is distributed among different consortia, the concentration of the secreted molecule

can differ according to the number of cells simultaneously producing it. This kind of multicellular approach and others based on single cells require sophisticated wiring design (Silva-Rocha and de Lorenzo, 2008; Siuti et al., 2013; Macia et al., 2016, 2017). By contrast, ParAlleL requires very minimal genetic modification and little tuning to obtain reliable outputs (**Figures 2**, **3**).

The implementation of ParAlleL presented here is simple, but its further development to useful applications presents a number of challenges. Firstly, expansion to 4 bits and beyond would require further well-behaved and non-cross-reacting antibiotic resistance markers, and would probably lead to even greater disparities in growth rate than those observed in the three-bit system (**Figure S1**). This could be addressed, and the flexibility and usefulness of the system increased, by moving away from direct use of antibiotics to a system using tightly controlled inducible promoters, each controlling a lethal "death gene," such as *ccdB*, with either the presence or absence of the inducer leading to lethality. In this way the system could be made to respond to different combinations of useful inputs, for example in the construction of multiplexed biosensors.

However, to achieve the full potential of ParAlleL, it will be necessary to generate layered systems in which the output from one layer serves as input to another layer. This might be accomplished via quorum sensing, but the implementation would be rather complex and limited. A more attractive option is to transfer the same concept, using a set of agents which each responds to a single combination of inputs, to an alternative system. For example, the concept could be implemented in a cell-free system, in which inputs may be present as small molecules interacting with transcription factors, or as DNA or RNA oligonucleotides. Outputs from the first layer, in the form of DNA or RNA molecules generated by DNA replication (e.g., PCR) or transcription, could then serve as inputs to a second layer. For example, in a PCR-based approach, one of a set of templates would be amplified based on which primer oligonucleotides were added; the output PCR product could then be processed to generate another set of oligonucleotides, or used directly to initiate priming on a second set of templates via formation of three-way junctions. Alternatively, in a transcription-based system, output RNAs from a first layer computation could act as guide RNAs directing binding of CRISPR-based transcription factors to further templates to generate a new set of guide RNAs, eventually leading to production of an output RNA such as a mRNA leading to translation of a visually detectable signal.

Another interesting aspect of ParAlleL is that, since the final output is the result of a population-based calculation, these systems may show a level of "Byzantine fault tolerance," allowing reliable outcomes even in the face of levels of noise which are unavoidable in biological systems. This would represent a new level of robustness in biological computation systems.

DATA AVAILABILITY

All datasets generated for this study are included in the manuscript and/or the supplementary files. More Data is available at https://doi.org/10.7488/ds/2497.

AUTHOR CONTRIBUTIONS

FM and CF conceived and designed the experiments. FM and BL performed the experiments. FM and CF analyzed and validated the data. FM and CF contributed reagents, materials, and analysis tools. FM and CF wrote—reviewed & edited the manuscript.

FUNDING

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fbioe.2019.00046/full#supplementary-material

REFERENCES

Andrianantoandro, E., Basu, S., Karig, D. K., and Weiss, R. (2006). Synthetic biology: new engineering rules for an emerging discipline. *Mol. Syst. Biol.* 2:2006.0028. doi: 10.1038/msb4100073

Brenner, M. J., Cho, J. H., Wong, N. M. L., and Wong, W. W. (2018). Synthetic biology: immunotherapy by design. *Annu. Rev. Biomed. Eng.* 20, 95–118. doi: 10.1146/annurev-bioeng-062117-121147

Chung, C. T., Niemela, S. L., and Miller, R. H. (1989). One-step preparation of competent Escherichia coli: transformation and storage of bacterial cells in the same solution. *Proc. Natl. Acad. Sci. U.S.A.* 86, 2172–2175. doi: 10.1073/pnas.86.7.2172

Keller, K. L., Wall, J. D., and Chhabra, S. (2011). "Methods for engineering sulfate reducing bacteria of the genus *Desulfovibrio*," in *Methods in Enzymology. Vol. 497.* ed C. Voigt (California, CA: Academic Press), 503–517.

Macia, J., Manzoni, R., Conde, N., Urrios, A., de Nadal, E., Solé, R., et al. (2016). Implementation of complex biological logic circuits using spatially distributed multicellular consortia. *PLoS Comput. Biol.* 12:e1004685. doi: 10.1371/journal.pcbi.1004685

Macia, J., Vidiella, B., and Solé, R. V. (2017). Synthetic associative learning in engineered multicellular consortia. *J. R. Soc. Interface* 14:20170158. doi: 10.1098/rsif.2017.0158

Millacura, F. A., Largey, B., and French, C. E. (2018). BioLogic, a parallel approach to cell-based logic gates. *bioRxiv* [Preprint]. doi: 10.1101/303032

Nielsen, A. A., and Voigt, C. A. (2014). Multi-input CRISPR/Cas genetic circuits that interface host regulatory networks. *Mol. Syst. Biol.* 10:763. doi: 10.15252/msb.20145735

Regot, S., Macia, J., Conde, N., Furukawa, K., Kjellén, J., Peeters, T., et al. (2011). Distributed biological computation with multicellular engineered networks. *Nature* 469, 207–211. doi: 10.1038/nature09679

Roquet, N., Soleimany, A. P., Ferris, A. C., Aaronson, S., and Lu, T. K. (2016). Synthetic recombinase-based state machines in living cells. *Science* 353:aad8559. doi: 10.1126/science.aad8559

Shipman, S. L., Nivala, J., Macklis, J. D., and Church, G. M. (2017). CRISPR–Cas encoding of a digital movie into the genomes of a population of living bacteria. *Nature* 547, 345–349. doi: 10.1038/nature23017

Silva-Rocha, R., and de Lorenzo, V. (2008). Mining logic gates in prokaryotic transcriptional regulation networks. *FEBS Lett.* 582, 1237–1244. doi: 10.1016/j.febslet.2008.01.060

Silva-Rocha, R., Martínez-García, E., Calles, B., Chavarría, M., Arce-Rodríguez, A., de Las Heras, A., et al. (2012). The Standard European Vector Architecture (SEVA): a coherent platform for the analysis and deployment of complex prokaryotic phenotypes. *Nucleic Acids Res.* 41, D666–D675. doi: 10.1093/nar/gks1119

Siuti, P., Yazbek, J., and Lu, T. K. (2013). Synthetic circuits integrating logic and memory in living cells. *Nat. Biotechnol.* 31, 448–452. doi: 10.1038/nbt.2510

Unintentional Genomic Changes Endow *Cupriavidus metallidurans* with an Augmented Heavy-Metal Resistance

Felipe A. Millacura [1], Paul J. Janssen [2], Pieter Monsieurs [2], Ann Janssen [2], Ann Provoost [2], Rob Van Houdt [2] and Luis A. Rojas [3,*]

[1] School of Biological Sciences, University of Edinburgh, Edinburgh EH9 3JQ, UK; s1647595@sms.ed.ac.uk
[2] Interdisciplinary Biosciences, Belgian Nuclear Research Centre, SCK●CEN, 2400 Mol, Belgium; pjanssen@sckcen.be (P.J.J.); pieter.monsieurs@sckcen.be (P.M.); ann.janssen@sckcen.be (A.J); ann.provoost@sckcen.be (A.P.); rvhoudto@sckcen.be (R.V.H.)
[3] Chemistry Department, Faculty of Sciences, Universidad Católica del Norte, UCN, Antofagasta 1240000, Chile
*   Correspondence: luis.rojas02@ucn.cl; Tel.: +56-55-235-5629

Abstract: For the past three decades, *Cupriavidus metallidurans* has been one of the major model organisms for bacterial tolerance to heavy metals. Its type strain CH34 contains at least 24 gene clusters distributed over four replicons, allowing for intricate and multilayered metal responses. To gain organic mercury resistance in CH34, broad-spectrum *mer* genes were introduced in a previous work via conjugation of the IncP-1β plasmid pTP6. However, we recently noted that this CH34-derived strain, MSR33, unexpectedly showed an increased resistance to other metals (i.e., Co^{2+}, Ni^{2+}, and Cd^{2+}). To thoroughly investigate this phenomenon, we resequenced the entire genome of MSR33 and compared its DNA sequence and basal gene expression profile to those of its parental strain CH34. Genome comparison identified 11 insertions or deletions (INDELs) and nine single nucleotide polymorphisms (SNPs), whereas transcriptomic analysis displayed 107 differentially expressed genes. Sequence data implicated the transposition of IS*1088* in higher Co^{2+} and Ni^{2+} resistances and altered gene expression, although the precise mechanisms of the augmented Cd^{2+} resistance in MSR33 remains elusive. Our work indicates that conjugation procedures involving large complex genomes and extensive mobilomes may pose a considerable risk toward the introduction of unwanted, undocumented genetic changes. Special efforts are needed for the applied use and further development of small nonconjugative broad-host plasmid vectors, ideally involving CRISPR-related and advanced biosynthetic technologies.

Keywords: *Cupriavidus*; heavy metals; genomic islands; genomic rearrangements; metal resistance genes

1. Introduction

Since life appeared on Earth some 3.7 billion years ago, microorganisms have undergone molecular changes to adapt (i.e., respond to selection) to the harsh conditions of their natural habitats, including extreme temperatures, pH, salinity, UV and ionizing radiation, and heavy metals [1]. In addition, global fluctuations in the composition of the atmosphere, the oceans, and the earth crust have elicited genomic changes in microbes over eons of time and hence contributed to microbial diversity [2]. Certain microorganisms had been adapted to heavy metals and radionuclides prior to human appearance. However, during the past few hundred years, anthropogenic influences have forced microbes to also adapt to pollutants previously nonexistent (xenobiotics) [3–6], as well as to increased concentrations of heavy metals and radionuclides [7–9].

Members of the beta-proteobacterial genus *Cupriavidus* are prime examples of microbial endurance, possessing a variety of genomic islands involved in the resistance to heavy metals or the degradation of aromatics or xenobiotics [10–20]. They all typically display a bipartite chromosomal structure with one chromosomal replicon bearing the marks of a plasmid-type maintenance and replication system henceforth called "chromid". In addition, most *Cupriavidus* strains carry one or two dispensable megaplasmids with a size of 100 kb or more. The model organism for heavy metal resistance, *Cupriavidus metallidurans* strain CH34, carries two megaplasmids, pMOL28 and pMOL30. Together, the four CH34 replicons encode resistance markers for a plethora of heavy metals including copper, nickel, zinc, cobalt, cadmium, chrome, lead, silver, gold, mercury, caesium, selenium, strontium, and uranium [12,21–25]. These resistances are mainly related to a variety of metal reduction and efflux systems [26–28]. Aromatics degradation, on the other hand, is carried out by various bacterial multicomponent mono- and di-oxygenases solely encoded by genes on the chromosome and chromid [29].

In an effort to improve the inorganic and organic mercury resistance of *C. metallidurans* CH34 and thus improve its utility in cleaning up mercury-contaminated environments, the IncP-1β plasmid pTP6, providing additional *mer* genes [30], was introduced in CH34 by biparental mating, leading to strain MSR33 [31]. These extra *mer* genes are part of a transposon, Tn*50580*, which is necessary for broad-spectrum (organomercury) resistance, including two genes, *merG* and *merB*, not native to CH34 (i.e., CH34 only contains two narrow-spectrum mercury resistance *merRTPADE* operons, one on each megaplasmid, conferring resistance to inorganic mercury). MerB plays a key role in methylmercury degradation through its unique ability to cleave the carbon-mercury bond in methylmercury and the subsequent shuttling of ionic mercury to MerA to reduce it to the less harmful elemental mercury.

In comparison to its parental strain CH34, strain MSR33 became 240% more resistant to inorganic mercury and gained resistance to methylmercury by incorporating the previously nonpresent *merBG* genes. Other metal resistances (as tested for chrome and copper) were seemingly unaffected, and the pTP6 plasmid was stably maintained for over 70 generations under nonselective pressure [31]. However, when we recently tested the resistances for both strains to additional metals (i.e., cadmium, cobalt, and nickel), we noted a significant increase of metal resistance for strain MSR33 compared to CH34 (this study). As this was fully unexpected, since the only difference between the two strains should be an additional *mer* gene dosage, implicating only an improved mercury resistance, we set out to investigate the reasons for this phenomenon. Considering the genome plasticity [32–34] and the intricate relationships between metal resistance loci [24] in *C. metallidurans* CH34, we decided to sequence the full genome of strain MSR33 and compare the sequence of its replicons with the corresponding replicons of the parental strain CH34 and with plasmid pTP6. We also performed microarray-based expression analysis on both CH34 and MSR33 gene sets to determine whether genomic differences could be correlated with differences in the expression of individual genes.

2. Materials and Methods

2.1. Strains and Culture Conditions

Cupriavidus metallidurans strains CH34 and MSR33, obtained from respectively the SCK•CEN (Mol, Belgium) and the Univesidad Católica del Norte (UCN) (Antofagasta, Chile) culture collections, were cultivated at 30 °C and 200 revolutions per minute (rpm) on a shaker in dark, aerobic conditions in a Tris-buffered mineral medium (MM284) [35] with 0.4% (*w/v*) succinate as the sole carbon source. *Escherichia coli* JM109 pTP6 (also obtained from the UCN culture collection) was cultured at 37 °C on an M9 minimal medium [36] supplemented with 0.4% (*w/v*) glucose.

2.2. Synthetic Construct Generation

The *merTPAGB*$_1$ gene cluster of pTP6 was amplified by PCR using Phusion High-Fidelity DNA polymerase (ThermoFisher, Aalst, Belgium) with primer pair pTP6mer_Fw-Rv (Table S1). In tandem,

the broad-host-range cloning vector pBBR1MCS-2 [37] was linearized by PCR with the same DNA polymerase and primer pair pBBR1MCS-2_GA_Fw-Rv (Table S1), providing homologous ends with the amplified *merTPAGB$_1$* locus. These compatible PCR products were end-ligated using the Invitrogen GeneArt® Seamless Cloning and Assembly Enzyme Mix (ThermoFisher). After transforming *E. coli* DG1 with the ligation mix and selection on Lysogeny Broth (LB) with 50 µg/mL kanamycin (Km), four randomly chosen transformants were tested by DNA digestion and fragment sizing for correct plasmid construction holding the *merTBAGB$_1$* genes. The plasmid gene construct of one transformant was further verified by sequencing its insert using forward and reverse cloning primers (Table S1) prior to electroporation into *C. metallidurans* CH34 on an Eppendorf 2510 Electroporator (Eppendorf, Aarschot, Belgium) using conditions as described previously [38].

2.3. Estimation of Bacterial Tolerance to Metals

Strains CH34 and MSR33 were grown overnight in MM284 liquid media with 0.4% (*w/v*) succinate as the sole carbon source and thereafter used as a pre-inoculum (1% *v/v*) for a freshly prepared 200 µL culture supplemented with increased concentrations of Hg^{2+} (from 0.0625 mM to 8 mM), Cd^{2+} (from 0.25 mM to 8 mM), and increasing steps (1 mM) of Co^{2+} or Ni^{2+} (from 5 mM to 20 mM). Cultures for metal contact were grown in microtiter plates at 30 °C on a rotary shaker at 120 rpm. Metal ion solutions were prepared from soluble salts of analytical grade (CdCl$_2$, HgCl$_2$, CoCl$_2$·6H$_2$O, and NiCl$_2$·6H$_2$O) in double-deionized water and were filter-sterilized before use. The lowest metal concentration that prevented growth after 48 h (i.e., showing no growth as measured at OD$_{600}$ by a CLARIOstar microplate reader (BMG LabTech, Offenburg, Germany)), was considered the minimum inhibitory concentration (MIC) (Table 1 and Table S2). All MIC analyses were performed using biological triplicates.

Table 1. Minimal inhibitory concentration (MIC) of heavy metals for *Cupriavidus metallidurans* strains MSR33 and CH34.

Strain/Metal in mM	Hg^{2+}	Cd^{2+}	Ni^{2+}	Co^{2+}
C. metallidurans CH34	0.01	2.0	10.0	11.0
C. metallidurans MSR33	0.10	4.0	12.0	20.0
C. metallidurans CH34 (pBBR::*merTPAGB$_1$*)	0.10	2.0	ND [1]	ND [1]

[1] ND: not determined.

2.4. Plasmid Copy Number Determination

Single-copy (i.e., "replicon-unique") genes were taken as representatives of the chromosome (*cadA*), chromid (*zniA*), and plasmids pMOL30 (*nccA*), pMOL28 (*cnrA*), and pTP6 (*merG*). Primer pairs were designed to amplify 150 bp of each gene (Table S1). Real-time PCR was performed on a 7500 Applied Biosystems Fast Real-Time PCR System (ThermoFisher) using QiaGen RT^2 Sybr Green Rox qPCR Mastermix (ThermoFisher), 20 ng of MSR33 or CH34 genomic DNA as a template, and 0.2 µM of each primer. To reduce nonspecific amplification, we incubated this mixture at 95 °C for 10 min as part of a hot start PCR setup. Next, a 40-cycle amplification and quantification protocol (15 s at 95 °C, 15 s at 58 °C, and 30 s at 60 °C) was performed with a single fluorescence measurement for each cycle. Finally, a melting curve program (15 s at 95 °C, 60 s at 60 °C, 30 s at 95 °C, and 15 s at 60 °C) was carried out. Plasmid copy numbers for both strains were determined using the absolute method (allowing estimates of both the absolute and relative number of plasmids per cell), following earlier described protocols [39]. Standard curves were created with 7500 Fast Software v2.3 of Applied Biosystems (Foster City, CA, USA), using serial (10-fold) dilutions of genomic DNA in a linear range from 20 ng to 0.2 pg. The qPCR efficiencies were calculated from slopes of the log-linear portion of the calibration curves from the equation $E = 10^{(1/slope)}$. Using the linear equation obtained from each calibration curve, log DNA copy numbers were derived by intersecting the obtained C_t values. All analyses were done in triplicate.

2.5. Illumina Sequencing and Assembly

MSR33 cells were grown overnight in MM284 liquid medium, and genomic DNA was extracted by using the Qiagen QIAamp DNA Mini Kit (ThermoFisher), following the instructions of the manufacturer. DNA quantity and quality were measured using a DropSense (Trinean, Piscataway, NJ, USA). Genome sequencing was performed on a HiSeq 2500 apparatus (Illumina, San Diego, CA, USA) using 2 × 250 bp paired-end reads. The reads were trimmed using the Trimmomatic tool [40] and their quality assessed using in-house Perl and shell scripts in combination with SAMtools [41], BEDTools [42], and a Burrows–Wheeler aligner with maximum exact matches (bwa-mem) [43].

The entire genome of *C. metallidurans* CH34 was sequenced and largely annotated [12,22]. Sequences for a chromosome (NC_007973.1), a chromid (NC_007974.2), and the large plasmids pMOL28 (NC_007972.2) and pMOL30 (NC_007971.2) were all obtained from GenBank and used as a reference for the assembly and annotation of trimmed sequences of the MSR33 genome. The DNA sequence of pTP6 plasmid, also obtained from Genbank (AM048832), was used for sequence comparisons between the CH34 and MSR33 sequence data sets. The full *C. metallidurans* MSR33 genome sequence (this study) is available from the NCBI Sequence Read Archive (SRA) under accession number PRJNA493617.

2.6. Total RNA Isolation and Microarray

Cupriavidus metallidurans MSR33 and CH34 cells were both cultivated in triplicate on MM284 medium supplemented with 0.4% (*w/v*) succinate. Cell samples were taken at the middle exponential phase (OD$_{600}$ 0.6–0.7) and centrifuged at 16,000 × *g* and 4 °C for 5 min. Pellets were quick-frozen with liquid nitrogen and kept at −80 °C for further analysis. Total RNA extraction was performed as described previously [24] by using an SV Total RNA Isolation System kit (Promega Benelux, Leiden, the Netherlands) according to the manufacturer's recommendations. Samples were cleaned and concentrated using a Nucleospin RNA cleanup XS kit (Macherey-Nagel, Düren, Germany). Concentrated RNA samples (10–20 μg) were retrotranscribed using the Invitrogen Superscript^{TM} Direct cDNA Labeling System (ThermoFisher) and labeled by incorporation of Cy3-dCTP (ref PA53021, control condition) and Cy-5dCTP (ref PA55021, experimental condition) by Pronto!^{TM} Long Oligo/cDNA Hybridization Solution (supplied with the Corning® Pronto! Universal Microarray Hybridization kit from Merck/Sigma-Aldrich, Overijse, Belgium), following the manufacturer's instructions. The microarrays we used were designed with 60-*mer* probes for 6205 Open Reading Frames that were spotted in triplicate onto glass slides (UltraGPS, Corning, NY, USA) using a MicroGrid system (BioRobotics, Cambridge, UK) at the microarray platform at SCK•CEN (Mol, Belgium). The spotted slides were cross-linked and placed in the presoaking solutions from the Pronto Kit (Promega, Madison, WI, USA). Analyses were performed on RNAs retrieved from CH34 and MSR33 cells, using respectively Cy3-dCTP and Cy5-dCTP incorporation and determination of Cy3/Cy5 signal intensity ratios. Labeled cDNA was resuspended in the universal hybridization buffer (Pronto kit), mixed, and added to the spotted slide for overnight hybridization at 42 °C in a Tecan HS4800 Pro hybridization station (Tecan Group Ltd., Männedorf, Switzerland). Afterwards, the slide was washed according to Pronto kit's protocol. Slides were scanned (at 532 and 635 nm) using the GenePix Personal 4100A microarray scanner (Molecular Devices, San Jose, CA, USA). All post-hybridization analyses were performed as described before [24]. In brief, spot signals were qualified using GenePix Pro v.6.0.1 software, and raw median density data were imported into R version 3.3.2 (https://cran.rstudio.com/) for statistical analysis using the LIMMA package version 2.15.15 (http://bioinf.wehi.edu.au/limma/), as available from Bioconductor (https://bioconductor.org). Background correction, normalizations, *t*-statistics, and *p*-value corrections were done as before [24]. Only log-transformed expression results with a *p*-value > 0.05 were considered for data interpretation (Table S2).

3. Results and Discussion

3.1. The Influence of Plasmid pTP6 on Increased Heavy Metal Resistance in MSR33

The *C. metallidurans* strains CH34 and MSR33 have the same genetic background, but MSR33 has, compared to its parental strain CH34, an extra plasmid 54 kb in size [31]. This plasmid, pTP6, is a broad-host-range IncP-1β plasmid originally isolated from mercury-polluted sediments [30]. It carries a transposon with *mer* genes that are not native to CH34 (i.e., *merG*) that encode an organomercurial transporter, and a pair of duplicate *merB* genes that encode periplasmic organomercurial lyases (Table S3). These additional *mer* genes in strain MSR33, via plasmid pTP6, grant this strain a 2.4-fold increased resistance for Hg^{2+} and a 16-fold increased resistance for CH_3Hg^+ [31]. In our hands, we noted a much-improved Hg^{2+} resistance for strain MSR33, with a 10-fold increase in comparison to its parental strain CH34 (Table 1). As an added note, in contrast to Rojas et al. [31], who performed MIC analyses on solid media, we performed our MIC analyses in a liquid medium, increasing metal bioavailability. Hence, sensitivity (i.e., the MIC (Hg) for CH34) was 0.05 mM in Rojas et al.'s study [31] but 0.01 mM in our study]. Surprisingly, when we tested MICs for the metals cadmium, nickel, and cobalt, we found a 2-fold increased resistance in MSR33 to Cd^{2+} and Co^{2+} and a 1.2-fold increased resistance to Ni^{2+} (Table 1).

The only genes on pTP6 with relevance to metal resistance are mercury resistance genes situated on transposon Tn*50580* as the two clusters *merR₁TPAGB₁* and *merR₂B₂D₂E* [30]. All other genes are typical for the backbones of self-transmissible and promiscuous IncP-1 plasmids from subgroups α and β and are involved in replication (*trfA*, *ssb*), plasmid maintenance, partitioning, control (*kfrABC*, *incC*, *korABC*, *kluAB*, *klcAB*, *kleABEF*, and a remnant of the resolvase gene *parA*), conjugal transfer (*traCDEFGHIJKL*), mating pair formation (*trbABCDEFGHIJKLMNO*), and transposition (*tniABQR* as part of Tn*50580*). Two more genes, *upf30.5* and *upf31.0*, are located downstream of *trbP* and encode, respectively, a putative outer membrane protein and a site-specific methylase (Figure S1). Except for the *upf* genes and the plasmid maintenance, partitioning, and control genes, pTP6 genes have at least one counterpart on one of the replicons of CH34. Taken together, we did not expect the pTP6 genes, other than the above-mentioned *mer* genes, to play any significant role in the augmented metal resistance of strain MSR33. Nonetheless, we decided to generate a new synthetic construct by cloning the *merTPAGB₁* gene cluster in the low copy number broad-host-range cloning vector pBBR1MCS-2 [37]. This small plasmid only contained two genes, *rep* and *mob*, involved in, respectively, plasmid replication and mobilization, as well as a kanamycin resistance marker (Km^R). Strain CH34 transformed with this new construct reached the same level of inorganic mercury resistance as the MSR33 strain but did not show an increase in cadmium resistance (Table 1). From this we deduced that the *mer* genes of pTP6 exerted a positive effect on host resistance to mercury. However, neither the concomitant increase of cadmium resistance in MSR33 nor its higher resistance to nickel and cobalt could be readily explained by the presence of the auxiliary, pTP6-associated mercury resistance genes in this strain.

We also determined the effect of various mixtures containing both mercury and cadmium on the growth of strains CH34 and MSR33. In general, the combination of the two metals was expected to be more toxic to cells than the corresponding metals alone. Indeed, strain CH34 was capable of growing in up to 1 mM of Cd^{2+} when combined with 6.25 μM of Hg^{2+}, and cellular growth was diminished beyond these threshold metal concentrations (Table S4). Instead, strain MSR33 was capable of growing in up to 4 mM of Cd^{2+} when combined with 6.25 μM of Hg^{2+} (Table S4). However, even in combination with mercury, cadmium resistance in MSR33 was still twice as high as the cadmium resistance in CH34. Moreover, strain MSR33 showed much higher tolerance to mercury in combination with cadmium, and growth was only affected at the threshold metal concentrations of 125 μM Cd^{2+} and 100 μM Hg^{2+} (Table S4).

To exclude the possibility that the increased resistance in strain MSR33 to Hg^{2+} and Cd^{2+}, either alone or in combination, was a mere effect of gene dosage, we also determined the plasmid copy number for all replicons in strain MSR33 and strain CH34 using quantitative PCR (Table S5). The presence of the pTP6 plasmid in strain MSR33 (plasmid copy number (*PCN*) = 1.8) did not alter

the relative copy numbers for the chromid and plasmids pMOL28 and pMOL30 with respect to the calculated chromosomal copy number taken as a reference. In addition, the Tn*50580* transposon carrying the broad spectrum *mer* gene cluster on plasmid pTP6 did not transpose (verified by genome sequence analysis present in Section 3.2).

The CH34 strain carries on its genome a total of four *mer* gene clusters: *merRTPA* on the chromosome, one complete *merRTPADE* on both plasmids pMOL28 and pMOL30, and one truncated *merRTΔP*, also on pMOL30 [12]. One could argue that the presence of a single copy of pTP6 in strain MSR33 raises the number of *mer* genes in strain MSR33, with one unit for genes *T*, *P*, *A*, *D*, and *E*, and two units for gene *R* (a second *merR* gene is located on the right-hand part of Tn*50580* on pTP6 [30] (Table S6), but the *merG* or *merB* genes of pTP6 were not considered here as they only play a role in organomercurial resistance). Considering the calculated PCN values of all replicons, the theoretical abundance of these genes increased by roughly 50–78% (Table S6). *MerR* and *merD* are transcriptional regulators that compete for the same operator sequence in the *merR*–*merT* intergenic region. The *merT* product is an inner membrane protein involved in the transport of Hg^{2+} ions into the cell cytoplasm. The *merP* product is a small periplasmic Hg^{2+}-sequestering protein that shuttles Hg^{2+} to the mercurial reductase MerA, which converts it into the significantly less toxic Hg(0) that then is allowed to leave the cell by passive diffusion. The *merE* product, finally, is another inner membrane protein and may play a role in the uptake of both CH_3-Hg^+ and Hg^{2+}. All these genes appeared to be intact, and there was no reason for us to assume that any of the multiple-copy genes would be dysfunctional or, with respect to each other (from gene to gene or copy to copy), would be differently transcribed or expressed (owing to limitations of gene-specific primer or probe design for multiple copies of these genes in qPCR or hybridization-based microarray procedures, no *mer* gene-specific expression data are available). A plausible explanation for the positive effect of the pTP6 *mer* genes on host mercury resistance could thus lay in the stoichiometry of *mer* gene products, particularly those involved in Hg^{2+} sequestration (*merP*) and transport (*merT/merE*). Nevertheless, the increased resistance to Cd^{2+} in MSR33 cannot be readily explained by a stoichiometric change in *mer* gene products. Also, when a single *merTPAGB* was introduced on a small plasmid into strain CH34, the increased Hg^{2+} resistance was still there, but the increased Cd^{2+} resistance was no longer seen (Table 1). From this we had to conclude that the MSR33 genetic background, besides the extra plasmid pTP6, may actually have differed from the genetic background of its parental strain CH34. In other words, the MSR33 genome had undergone genetic changes leading to an improved resistance to cadmium and possibly also to other metals. Such a genomic adaptation appears to be common to IncP-1 plasmid backbones [44]. In order to get to the bottom of this we decided to determine the DNA sequence of the entire genome of strain MSR33 and register in detail which genetic changes occurred with respect to the reference genomes of strain CH34 and plasmid pTP6.

3.2. The C. metallidurans MSR33 Genome Showed Multiple Insertions or Deletions and Single Nucleotide Polymorphisms

The whole-genome resequencing of MSR33 (since the known genome sequences of strain CH34 and plasmid pTP6 served as references, we considered this effort as a resequencing project) revealed a total of eight insertions and three deletions (Table 2), and nine single nucleotide polymorphisms (Table 3), all changes being located predominantly across the four replicons of the CH34 backbone, with only one genomic change occurring in plasmid pTP6 (Table 2, Figure 1). Most of the insertions (six out of eight) were found to be related to IS*1088*, an insertion element belonging to the IS*30* family with a typical size range of 1000–1250 bp [45]. It should be noted at this point that the CH34 genome indigenously harboured nine copies of IS*1088*, distributed on its chromosome and chromid but not on its megaplasmids [12], bringing the total of IS*1088* copies in the MSR33 genome to 15.

The majority of the new IS*1088* copies in the MSR33 genome were located on the chromosome and chromid, where all IS*1088* copies indigenous to CH34 resided, but one IS*1088* copy transposed into the *cnrY* gene (Rmet_6205) of pMOL28 (Table 2). This gene was part of the *cnrYXHCBAT* locus

involved in the inducible cobalt and nickel resistance in strain CH34, and encoded the anti-sigma factor CnrY that tethered, in conjunction with the sensor protein CnrX, the sigma factor CnrH, but released it in the presence of Ni^{2+} or Co^{2+} [46,47]. The sigma factor CnrH promoted transcription of its own locus *cnrYXH*, but also of the structural locus *cnrCBA*, encoding a resistance nodulation division (RND)-driven efflux system [12,48]. The inactivation of the *cnrY* gene in MSR33 by IS*1088* inevitably led to the constitutive derepression of *cnrCBAT* transcription and explained the increased cobalt and nickel resistance we observed for MSR33 (Table 1) (see also gene expression results in Section 3.3). A similar phenomenon was previously seen for spontaneous mutants of a pMOL30-less CH34 derivative, strain AE126 [35], which showed a significantly increased resistance to cobalt and nickel [49] and which was later acknowledged as being an IS- and frameshift-mediated inactivation of *cnrY* and *cnrX* [50].

The other five genes affected by the insertion of an IS*1088* element were Rmet_0312 (*nptA*) and Rmet_2860 (*tauB*), lying on the chromosome; and Rmet_4160 (*pelF*), Rmet_4867 (*acrA*), and Rmet_5682 (*nimB*), lying on the chromid (Table 2). The first three genes encode proteins with general cellular functions, and their inactivation is very unlikely to affect heavy metal resistance in strain MSR33. The fourth gene, *acrA*, encodes a membrane fusion protein and is part of an intact *acrABC* operon whose gene products form a tripartite multidrug efflux system. The last gene, *nimB*, is involved in efflux-mediated heavy metal resistance, encoding also a membrane fusion protein resembling other membrane metal-binding fusion proteins in structure and function (e.g., CzcB, CnrB, CusB, and ZneB) by forming a periplasmic bridge between the cytoplasmic porter and the outer membrane channel [48]. Nonetheless, taking also into account that the *nimA* gene is already inactivated in strain CH34 (and MSR33) by the presence of the insertion sequence element IS*Rme3* [12], it is hard to see how the inactivation of *nimB* would result in the increased metal resistance we observed in strain MSR33. Two insertions were not attributed to IS*1088*. One appeared to be the result of a Tn3-related transposition event affecting gene Rmet_5388, encoding a tentative ApbE-like lipoprotein, while the other concerned an unknown mutational event in gene Rmet_5508 resulting in the insertion of a nucleotide triplet (+CTT) (Table 2). This gene encodes a long-chain fatty-acid CoA-ligase and also underwent a triplet deletion (-CGG) just a few nucleotides downstream of the triplet insert. As a combined result, the actual change at the protein level remained perfectly in-frame and gave a protein of the same length, but led to an altered peptide sequence at positions 149–153 (i.e., Xxx-Leu-**Arg**-**Phe**-**Ala**-Gln-Xxx in CH34 to Xxx-Leu-**Phe**-**Ala**-**Lys**-Gln-Xxx in MSR33 (amino acidic sequence change from **Arg**-**Phe**-**Ala** to **Phe**-**Ala**-**Lys**). The third deletion in MSR33 occurred in plasmid pTP6, effectively destroying the genes *upf30.5*, *upf31.0*, and *parA*, immediately preceding Tn*50580*. Except for *cnrY* and *nimB*, none of the aforementioned genes are in any way associated with metal resistance.

Table 2. Insertions (INs) and deletions (DELs) present in the genome of *C. metallidurans* MSR33.

Type	Replicon (#)	Start	End	Size (bp)	Description	Targeted Gene(s)	Function (MaGe Annotation) (*)
IN	CHR1	328395	328395	1102	IS*1088*	Rmet_0312	putative transporter
IN	CHR1	3106728	3106728	1102	IS*1088*	Rmet_2857 (*tauB*)	taurine ABC transporter ATP-binding protein
DEL	CHR2	602035	602490	455		Rmet_4033	LysR family transcriptional regulator
IN	CHR2	741818	741818	1102	IS*1088*	Rmet_4160 (*pelF*)	EPS biosynthesis, biofilm formation
IN	CHR2	1529231	1529231	1102	IS*1088*	Rmet_4867 (*acrA*)	membrane fusion protein, multidrug efflux
IN	CHR2	2113815	2113815	256	*tniA*/Tn3	Rmet_5388 (*apbE*)	ApbE-like lipoprotein
IN	CHR2	2253560	2253560	3	+CTT	Rmet_5508	long-chain-fatty-acid-CoA ligase
DEL	CHR2	2253566	2253568	3	−CGG	Rmet_5508	long-chain-fatty-acid-CoA ligase
IN	CHR2	2440975	2440975	1104	IS*1088*	Rmet_5682 (*nimB*)	membrane fusion protein, heavy metal transport
IN	pMOL28	53484	53484	1104	IS*1088*	Rmet_6205 (*cnrY*)	antisigma factor
DEL	pTP6	26155	27385	1230		*upf30.5*, *upf31.0*, *parA*	respectively an outer membrane protein, a DNA methylase, and a plasmid partition protein

Chromosome: CHR1; chromid: CHR2. * Genomic and Metabolics analysis software (www.genoscope.cns.fr).

Table 3. Single nucleotide polymorphisms (SNPs) detected in the genome of *C. metallidurans* MSR33.

Replicon (#)	Position	SNP	SNP Type	Affected Gene	Gene Description *
CHR1	333850	A→G	intergenic (+201/−75)	Rmet_0314 →/→ *ssb1*	putative transporter, major facilitator family/single-stranded DNA-binding protein (helix-destabilizing protein)
CHR1	645608	A→G	A23A (GCA→GCG)	Rmet_0598 →	Ser/Thr protein phosphatase family protein
CHR1	2400725	G→A	intergenic (−77/+116)	*greA* ←/← *carB*	transcription elongation factor/carbamoyl-phosphate synthase large subunit
CHR1	3418147	A→G	V295A (GTC→GCC)	*dppF* ←	dipeptide transporter; ATP-binding component of ABC superfamily
CHR1	3444412	A→G	V17A (GTC→GCC)	NirJ ←	heme d_1 biosynthesis protein
CHR1	3456113	A→G	V76A (GTG→GCG)	acyP ←	acylphosphatase
CHR2	2253543	A→T	V158E (GTG→GAG)	Rmet_5508 ←	long-chain-fatty-acid-CoA ligase
CHR2	2253553	G→T	P155T (CCG→ACG)	Rmet_5508 ←	long-chain-fatty-acid-CoA ligase
CHR2	2529357	T→C	G264G (GGA→GGG)	Rmet_5769 ←	esterase

\# Chromosome: *CHR1*; chromid: *CHR2*; arrows show positive → or negative ← gene orientation; * www.genoscope.cns.fr.
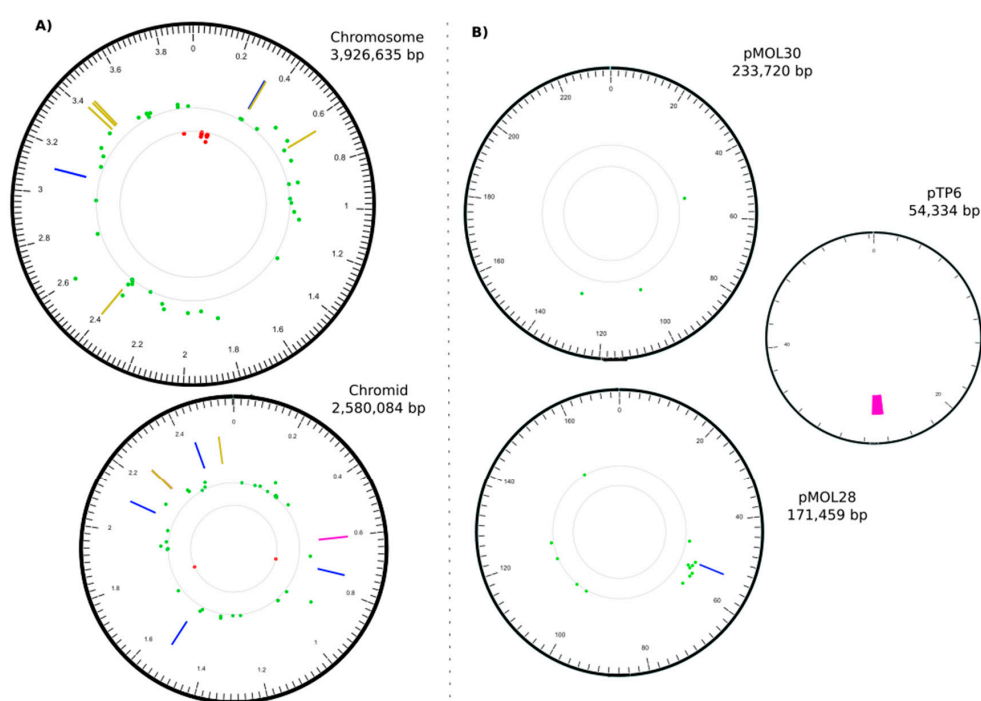

Figure 1. Correlation map between Insertions or Deletions (INDELs), Single nucleotide polymorphisms (SNPs), and transcriptional changes found for *C. metallidurans* MSR33. Concentric circles (ring) displayed from the outside inwards: (**A**) In ring 1, chromosome and chromid size scale in megabases (Mb) using a 20 kilobase (kb) window; (**B**) plasmids size scale in kb using a 2 kb window. In ring 2, position of insertion (blue), deletion (pink), and SNPs (mustard). INDELs and SNPs listed in Tables 2 and 3. In ring 3, each dot represents a single gene basal expression, overexpression (log$_2$ *ratio* > 1, green), or repression (log$_2$ *ratio* < −1, red), with a *p*-value < 0.05. Plotted genes listed by function in Table S7. Circos plot created with Circa (http://omgenomics.com/circa).

In addition to the eight insertions and three deletions in the MSR33 genome, sequence analysis revealed the presence of nine single nucleotide polymorphisms (Table 3). Two of those occurred in intergenic regions on the chromosome, without apparent disruption of gene regulatory elements, while the other seven occurred in protein-encoding genes (four on the chromosome and three on the chromid). Except for the "silent" mutation (i.e., no aa change) at position 645608, these Single nucleotide polymorphisms (SNPs) caused aa substitutions in the corresponding gene products (Table 3). No SNPs were detected in any of the plasmids. Remarkably, gene Rmet_5508 lying on the chromid (CHR2) once again was a target for mutation, displaying two SNPs in the immediate vicinity of the aforementioned triplet insertion and deletion in this gene (Table 3), bringing the full change of this

region from **RFA**QK**P**AY**V** to **FAK**QK**T**AY**E** (changes are in bold and underlined). It is uncertain whether these protein changes would have any effect on the cellular and metabolic functions in MSR33.

　　Taken together, of the 11 Insertions or Deletions (INDELs) and nine SNPs identified by the whole genome resequencing of the *C. metallidurans* strain MSR33, only the *cnrY* inactivation by IS*1088* and the concomitant derepression of *cnrB* (see above) may be directly linked to the observed augmented heavy metal resistance in this strain, at least for Co^{2+} and Ni^{2+} (see above). None of the other genomic changes seemed to play a role in this augmentation. The augmented resistance for Cd^{2+} in strain MSR33 (Table 1), however, remains a puzzle. The fact that such augmentation for Cd^{2+} was only noted for MSR33 with an altered genome (i.e., with 17 INDELs and six SNPs), but not in CH34 transformed with plasmid pBBR::*merTPAGB*$_1$ (Table 1) strongly indicates that the MSR33 genetic background was at play. In basic terms, bacterial resistance to toxic metals depends on two cellular processes, metal binding and metal transport, with the former generally being an intrinsic part of the latter. It is well established that many proteins or peptides that mediate the transport, buffering, or detoxification of metal ions in living cells have metal-binding domains (MBDs) in which certain amino acid residues (e.g., cysteine), as well as their structural layout and relative position to each other, play a key role in metal selectivity and specificity [51,52]. While some of these proteins might be highly metal-specific, other proteins follow a more relaxed, nonspecific mode of metal binding. For instance, divalent metal uptake in *C. metallidurans* is governed by a battery of redundant transporters that display a minimal degree of metal cation selectivity [53]. Depending on the environment, this may lead to a cytoplasmic pool of unsolicited metal ions that at some point, particularly when reaching a toxic threshold, need to be removed by the cell. In *C. metallidurans* this was done by one of three efflux systems: Cation diffusion facilitators (CDF), P-type ATPases, and the earlier mentioned RND-driven transenvelope transporters (HME-RND). Their main task in *C. metallidurans*, because of this bacterium's adaptation to metal-rich environments, was to balance the cytoplasmic and periplasmic concentrations of unwanted transition metals by entering the cellular arena and going into competition for metal cations with the "frivolous" metal uptake systems. Interestingly, all three types of metal efflux systems seemed to possess some degree of frivolity toward metal ions as well, albeit perhaps not as outspoken as for the metal uptake systems. The *C. metallidurans* CzcD exporter (Rmet_5979), for instance, allowed as a CDF protein Zn^{2+}, Cd^{2+}, or Co^{2+} as a substrate [54], whereas the DmeF and FieF exporters (Rmet_0198 and Rmet_3406) displayed as CDF family members broad metal specificity for Zn^{2+}, Cd^{2+}, Co^{2+}, and Ni^{2+} [55]. It is worth mentioning that disruption of the *dmeF* gene in strain CH34 dramatically lowered the resistance for Co^{2+} (but not for Zn^{2+}, Cd^{2+}, and Ni^{2+}), indicating a complex interplay between the DmeF exporter and the CzcCBA and CnrCBA efflux pumps (possibly partially obscured by the action of other metal resistance systems) [55]. Moreover, CDF proteins can play diverse roles and may possess different metal ion selectivity depending on the environmental conditions (i.e., by adjusted K_d values for certain metals) [56]. In addition, the eight metal resistance-related P_{1B}-type ATPases currently identified in strain CH34 can be subdivided into two groups according to their substrate profile [28]: Those that extrude Cu^+ and Ag^+ (CupA and CupF) and those that extrude Zn^{2+}, Cd^{2+}, Co^{2+}, or Pb^{2+} (ZntA, CadA, PbrA, and CzcP). These exporters mainly differ in the presence of unique amino acid sequences in their transmembrane MBDs, hence defining their metal specificity. But even within a subgroup, differences may exist in terms of metal affinity. For example, CzcP encoded by plasmid pMOL30 is unable to mediate Zn resistance on its own but rather augments the metal exportability of the ZntA, CadA, and PbrA exporters [57]. In a similar fashion, the five active HME-RND efflux systems in strain CH34 displayed a limited substrate spectrum, pumping out either the monovalent metal cations Cu^+ and Ag^+ (CusA, SilA) or the divalent metal cations Zn^{2+}, Ni^{2+}, and Co^{2+}, with occasionally also Cd^{2+} (ZniA, CnrA, CzcA) [28] (Figure 2). In such HME-RND systems, two steps of heavy-metal extrusion were discerned, the periplasmic and the transenvelope efflux (Figure 2). Each step involved the interaction of metals with MBDs within the Membrane Fusion Protein (MFP) and RND proteins. Sometimes, the delivery of periplasmic metal ions to the typical $C_3B_6A_3$-complex is facilitated by a small periplasmic metallochaperone, as is the

case for the *E. coli* CusCBFA system [58] (and likely, based on CusF aa sequence similarities, also the CusCBAF complex of strain CH34). Little is known about the substrate specificity of the metal-binding proteins of HME-RND efflux complexes. Apparently, metal-induced conformational changes in the $C_3B_6A_3$-complex are required in order to create a proper metal-guiding $C_3B_6A_3$ channel for metal export to take place [48,59–61].

Figure 2. Transcriptional changes in *C. metallidurans* MSR33 with respect to CH34, with both strains grown under equal and nonselective conditions (see methods). Bar graphs show the significantly (*p*-value < 0.05) higher expression (log$_2$ ratio > +1) and lower expression (log$_2$ ratio < −1) of MSR33 genes (with CH34 gene expression levels as reference). Transcriptional changes from genes pertaining to all replicons are depicted. Genes indicated with an * are part of pMOL30.

As mentioned, it is not inconceivable that the genetic changes in MSR33 instigated cellular conditions or pleiotropic effects that were generally favourable for Cd^{2+} detoxification and hence led to the observed improvement in Cd^{2+} resistance. Possibly, this involved the temporal recruitment of one or more metal binding export proteins, from known metal resistant systems or from hitherto unknown export systems, able to bind Cd^{2+}. The transition metals cadmium and mercury belong to Group 12 of chemical elements in the periodic table, together with zinc and copernicium. Although these four metals differ in significant respects, they also have common properties. Particularly, Cd and Hg are similar in their outer shell electron configuration ($d^{10}s^2$) and atomic radius (ca. 150 pm), and their cations both have a high affinity for sulfhydryl groups in cellular compounds and proteins (i.e., in methionine and cysteine residues). From this perspective, competition between Cd^{2+} and Hg^{2+} for certain MBDs cannot be excluded. Bacterial evolution and adaptation to new or rapidly changing environments implies a delicate balance between the safeguarding of genome integrity and the tolerance for genome instability. A too-rigid genome inevitably will lead to the demise of innovative power and hence adaptability of the host, whereas a too plastic or "fluid" genome may lead to disadvantageous mutations and cell growth arrest, or even cell death. This balance between beneficiary and perilous change in a bacterial genome also relates to the general fitness and the energy household of its host. Members of the genus *Cupriavidus*, and in particular *C. metallidurans*, appear to be masters in adaptation as they are home to a wide variety of habitats, often in extreme conditions [62–65]. The introduction of the 54 kb plasmid pTP6 into a strain already carrying two large replicons of 3.9 and 2.6 Mb (chromosome and chromid, respectively) and two megasized plasmids of 171 and 234 kb (pMOL28 and pMOL30, respectively) could be seen as a serious additional burden to the host regardless of whether or not phenotypical or physiological changes occur.

Because the charting of genomic changes in MSR33 with respect to its parental strain CH34 did not provide us with any clues or direct evidence on the involvement of certain genetic loci or of any of the known metal resistance determinants on the observed augmented Cd^{2+} resistance, we decided to compare the basal gene expression data for strains CH34 and MSR33 using RNA microarray technology in an attempt to associate their gene expression profiles with strain-specific physiological behaviour, with a focus on differentially expressed (DE) genes that might be involved in the cellular detoxification of heavy metals such as Cd^{2+}.

3.3. Transcriptional Analysis of Strain MSR33

Strains CH34 and MSR33, which were equally grown in nonselective conditions without any metal-related stress, were investigated for basal gene expression levels, and their expression profiles were compared. A total of 107 DE genes showed statistically significant changes in their expression (Figure 2). Affected genes pertained to the main chromosome (55 genes), the chromid (36 genes), and the plasmids pMOL30 (3 genes) and pMOL28 (13 genes) (Figure 2). In general terms, the products of these 107 genes could be grouped according to their predicted annotated function [12]: Catalytic function (35 genes), transport (24 genes), transcriptional regulation (11 genes), recombination (6 genes), movement and chemotaxis (9 genes), and miscellaneous (22 genes) (Table S7). Most of these genes had a higher expression in strain MSR33, with only 16 genes in strain MSR33 showing a lower expression level. What is immediately striking is that all genes of the *cnrYXHCBAT* operon on plasmid pMOL28 had a significantly higher expression (Figure 2), with $\log_2$ fold changes ranging from four to six. As we know from MSR33 whole genome sequence analysis, this high expression of the *cnr* locus in strain MSR33 was the direct result of IS*1088*-mediated *cnrY* inactivation and hence derepression of the *cnr* locus, explaining the increased resistance we observed for strain MSR33 to Co^{2+} and Ni^{2+} (see previous sections). Equally noticeable is the complete absence of an altered expression of the *czc* locus on pMOL30, strongly suggesting that the augmented Cd^{+2} resistance we see for strain MSR33 was independent of this locus. This, in fact, corroborates earlier findings about the pMOL30-less CH34 derivative AE126 (which, like MSR33, also has an IS*1088*-mediated inactivated *cnrY* gene on the remaining plasmid pMOL28): Vandecraen et al. [50] showed, next to a heightened resistance

to Zn^{2+}, Co^{2+}, and Ni^{2+}, a 2-fold increased resistance to Cd^{2+}. Adding another level of complexity, when strains AE126 and AE104 (a CH34 derivative lacking both pMOL28 and pMOL30 plasmids) [35] were transformed with pTP6, these strains, like MSR33, gained an improvement in Cd^{2+} resistance, albeit to a lesser extent (Rojas LA, personal communication).

This would indicate that the augmentation of Cd^{2+} resistance in MSR33 by pTP6 conjugation should be seen as a layered process brought about by multiple factors and possibly diverse mechanisms supporting each other. We cannot say at this point what these mechanisms precisely are and how and when they are triggered, as we have no information about the genomic changes in pTP6 conjugants of AE104 and AE126 (as pTP6 conjugation in CH34 causes genomic changes, this would most likely also be the case for pTP6 conjugation in strains AE104 and AE126, but not necessarily involving the same genomic changes). Clearly, further studies are needed to understand the augmented metal resistance in pTP6 conjugants of CH34 and its derivatives, including (1) the resequencing of pTP6-conjugated AE104 and AE126 strains and (2) the extensive RNAseq-based genetic response analyses for a wider range of heavy metals in all three pTP6-conjugated strains. The additional possibility that some of the observed genetic changes were already introduced to the recipient CH34 strain prior to conjugation with pTP6 cannot be entirely excluded. Lastly, the plasmid-curing procedures used to obtain strains AE104 and AE126 (i.e., applying mitomycin C, nalidixic acid, or hydroxyurea to growing CH34 cells [35]) may also have had mutagenic effects or may have induced transposition activity. In this respect, it would be best, in the frame of future studies, to resequence these strains as well.

A very high $\log_2$ fold difference in the expression of >4 was also noted in strain MSR33 for the Rmet_4229 gene, a *dctA* paralogue whose product was functionally annotated as a C4-dicarboxylate transporter and which is unlikely to have any connection to metal detoxification or resistance, and gene Rmet_2382, originally identified in CH34 as a transposase-encoding *tnpA* gene (IS*1088*) (Figure 2). Intermediate high $\log_2$ fold changes of >2 were seen in strain MSR33 for another 17 genes (Table S7), whereas the remaining 72 genes showed a $\log_2$ fold change between one and two. None of these genes is thought to be involved in metal binding, metal detoxification, or metal resistance. Among the genes with lowered expression in strain MSR33, we noted the *pelG* gene (Rmet_4161), which is part of the *pelABCDEFG* operon required to produce an extracellular polysaccharide that has been implicated in biofilm development [66]. Our sequence data confirmed that the IS*1088* element transposed into the *pelF* gene (Table 2), thereby disrupting expression of the *pelG* gene. This could explain the complete lack of biofilm formation in strain MSR33 reported to us by P. Alviz in a personal communication.

In conclusion, the genome of MSR33 underwent eight insertions, three deletions, and nine SNPs. At least seven of the insertions were due to the action of mobile genetic elements, with their presence fully confirmed by sequence data (implicating IS*1088* in six cases), whereas one small insertion and all three deletions in strain MSR33 may have been the result of DNA recombination or transposition events. The *C. metallidurans* genome is known to be ridden with a very high number of mobile genetic elements, with 57 IS elements, 19 other transposable elements, and 16 genomic islands for its type strain CH34 [12,32,34]. In concordance with this genomic fluidity, *C. metallidurans* displays a highly versatile metabolism and an inherent ability to inhabit a variety of harsh environments [9,62–65]. This adaptability has not come about overnight but is the wonderful result of microbial evolution over long periods of time. In a time in which large chunks of DNA were retrieved from the environment (e.g., by plasmid transfer or gene exchanges), adaptation was brought about by DNA mutations and natural selection and molecular inventions took place, steadily moulding the genome into its present large (6.9 Mb) and highly malleable form, providing the bacterium with a vast array of possibilities for rapid genetic responses (hence its well-chosen epithet as "Master Survivalist") [12]. However, tinkering with this hugely evolved and dynamic genome holds intrinsic dangers. Although the plasmid pTP6 was maintained stably in strain CH34 (i.e., MSR33) for over 70 generations under nonselective conditions [31], it has now become clear from our study at hand that the receiving host's genome underwent multiple changes in the form of 11 INDELs and 9 SNPs, affecting the physiology and heavy metal resistance of the host.

It would be wrong to point the finger at the extra plasmid as the "usual suspect" for these genetic changes, but rather we hold the actual process of conjugation responsible. Conjugative interaction appears to be a strong stimulus for transposition [67–69], and hence it is easy to envisage that, as a result of conjugation procedures, some elements of the extensive mobilome of *C. metallidurans* (with nearly 100 mobile elements) were triggered into action and "moved around", causing genetic changes that led to clearly perceptible but also less visible (and less understood) effects alike. The take-home message here is that the genetic engineering of bacteria with large complex and dynamic genomes should be carried out with much caution and that a strong preference should be given to the new generation of small broad-host-range cloning vectors and CRISPR-based technologies nowadays available [70–74].

References

1. Bakermans, C. *Microbial Evolution under Extreme Conditions*; De Gruyter: Altoona, PA, USA, 2015; Volume 2.

2. Knoll, A.H. Paleobiological perspectives on early microbial evolution. *CSH Perspect. Biol.* **2015**, *7*, a018093. [CrossRef] [PubMed]

3. Springael, D.; Top, E.M. Horizontal gene transfer and microbial adaptation to xenobiotics: New types of mobile genetic elements and lessons from ecological studies. *Trends Microbiol.* **2004**, *12*, 53–58. [CrossRef] [PubMed]

4. Cao, L.; Liu, H.; Zhang, H.; Huang, K.; Gu, T.; Ni, H.; Hong, Q.; Li, S. Characterization of a newly isolated highly effective 3,5,6-trichloro-2-pyridinol degrading strain *Cupriavidus pauculus* P2. *Curr. Microbiol.* **2012**, *65*, 231–236. [CrossRef] [PubMed]

5. Ilori, M.O.; Picardal, F.W.; Aramayo, R.; Adebusoye, S.A.; Obayori, O.S.; Benedik, M.J. Catabolic plasmid specifying polychlorinated biphenyl degradation in *Cupriavidus* sp. strain SK-4: Mobilization and expression in a pseudomonad. *J. Basic Microb.* **2015**, *55*, 338–345. [CrossRef] [PubMed]

6. Mergeay, M.; Van Houdt, R. Adaptation to xenobiotics and toxic compounds by *Cupriavidus* and *Ralstonia* with special reference to *Cupriavidus metallidurans* CH34 and mobile genetic elements. In *Biodegradative Bacteria: How Bacteria Degrade, Survive, Adapt, and Evolve*; Nojiri, H., Tsuda, M., Fukuda, M., Kamagata, Y., Eds.; Springer: Tokyo, Japan, 2014; pp. 105–127.

7. Nies, D.H. Microbial heavy-metal resistance. *Appl. Microbiol. Biotechnol.* **1999**, *51*, 730–750. [CrossRef] [PubMed]

8. Mergeay, M. Bacteria adapted to industrial biotopes: The metal resistant *Ralstonia*. In *Bacterial Stress Responses*; Storz, G., Hengge-Aronis, R., Eds.; ASM Press: Washington, DC, USA, 2000; pp. 403–414.

9. Sobecky, P.A.; Coombs, J.M. Horizontal gene transfer in metal and radionuclide contaminated soils. In *Horizontal Gene Transfer: Genomes in Flux*; Gogarten, M.B., Gogarten, J.P., Olendzenski, L.C., Eds.; Humana Press: Totowa, NJ, USA, 2009; pp. 455–472.

10. Pohlmann, A.; Fricke, W.F.; Reinecke, F.; Kusian, B.; Liesegang, H.; Cramm, R.; Eitinger, T.; Ewering, C.; Potter, M.; Schwartz, E.; et al. Genome sequence of the bioplastic-producing "knallgas" bacterium *Ralstonia eutropha* H16. *Nat. Biotechnol.* **2006**, *24*, 1257–1262. [CrossRef] [PubMed]

11. Amadou, C.; Pascal, G.; Mangenot, S.; Glew, M.; Bontemps, C.; Capela, D.; Carrere, S.; Cruveiller, S.; Dossat, C.; Lajus, A.; et al. Genome sequence of the beta-rhizobium *Cupriavidus taiwanensis* and comparative genomics of rhizobia. *Genome Res.* **2008**, *18*, 1472–1483. [CrossRef] [PubMed]

12. Janssen, P.J.; Van Houdt, R.; Moors, H.; Monsieurs, P.; Morin, N.; Michaux, A.; Benotmane, M.A.; Leys, N.; Vallaeys, T.; Lapidus, A.; et al. The complete genome sequence of *Cupriavidus metallidurans* strain CH34, a master survivalist in harsh and anthropogenic environments. *PLoS ONE* **2010**, *5*, e10433. [CrossRef] [PubMed]

13. Lykidis, A.; Perez-Pantoja, D.; Ledger, T.; Mavromatis, K.; Anderson, I.J.; Ivanova, N.N.; Hooper, S.D.; Lapidus, A.; Lucas, S.; Gonzalez, B.; et al. The complete multipartite genome sequence of *Cupriavidus necator* JMP134, a versatile pollutant degrader. *PLoS ONE* **2010**, *5*, e9729. [CrossRef] [PubMed]

14. Poehlein, A.; Kusian, B.; Friedrich, B.; Daniel, R.; Bowien, B. Complete genome sequence of the type strain *Cupriavidus necator* N-1. *J. Bacteriol.* **2011**, *193*, 5017. [CrossRef] [PubMed]

15. Cserhati, M.; Kriszt, B.; Szoboszlay, S.; Toth, A.; Szabo, I.; Tancsics, A.; Nagy, I.; Horvath, B.; Nagy, I.; Kukolya, J. De novo genome project of *Cupriavidus basilensis* OR16. *J. Bacteriol.* **2012**, *194*, 2109–2110. [CrossRef] [PubMed]

16. Hong, K.W.; Thinagaran, D.A.L.; Gan, H.M.; Yin, W.F.; Chan, K.G. Whole-genome squence of *Cupriavidus* sp. strain BIS7, a heavy-metal-resistant bacterium. *J. Bacteriol.* **2012**, *194*, 6324. [CrossRef] [PubMed]

17. Li, L.G.; Cai, L.; Zhang, T. Genome of *Cupriavidus* sp. HMR-1, a heavy metal-resistant bacterium. *Genome Announc.* **2013**, *1*, e00202-12. [CrossRef] [PubMed]

18. Ray, J.; Waters, R.J.; Skerker, J.M.; Kuehl, J.V.; Price, M.N.; Huang, J.; Chakraborty, R.; Arkin, A.P.; Deutschbauer, A. Complete genome sequence of *Cupriavidus basilensis* 4G11, isolated from the oak ridge field research center site. *Genome Announc.* **2015**, *3*, e00322-15. [CrossRef] [PubMed]

19. Wang, X.Y.; Chen, M.L.; Xiao, J.F.; Hao, L.R.; Crowley, D.E.; Zhang, Z.W.; Yu, J.; Huang, N.; Huo, M.X.; Wu, J.Y. Genome sequence analysis of the naphthenic acid degrading and metal resistant bacterium *Cupriavidus gilardii* CR3. *PLoS ONE* **2015**, *10*, e0132881. [CrossRef] [PubMed]

20. Fang, L.C.; Chen, Y.F.; Zhou, Y.L.; Wang, D.S.; Sun, L.N.; Tang, X.Y.; Hua, R.M. Complete genome sequence of a novel chlorpyrifos degrading bacterium, *Cupriavidus nantongensis* X1. *J. Biotechnol.* **2016**, *227*, 1–2. [CrossRef] [PubMed]

21. Mergeay, M.; Monchy, S.; Vallaeys, T.; Auquier, V.; Benotmane, A.; Bertin, P.; Taghavi, S.; Dunn, J.; van der Lelie, D.; Wattiez, R. *Ralstonia metallidurans*, a bacterium specifically adapted to toxic metals: Towards a catalogue of metal-responsive genes. *FEMS Microbiol. Rev.* **2003**, *27*, 385–410. [CrossRef]

22. Monchy, S.; Benotmane, M.A.; Janssen, P.; Vallaeys, T.; Taghavi, S.; van der Lelie, D.; Mergeay, M. Plasmids pMOL28 and pMOL30 of *Cupriavidus metallidurans* are specialized in the maximal viable response to heavy metals. *J. Bacteriol.* **2007**, *189*, 7417–7425. [CrossRef] [PubMed]

23. Avoscan, L.; Untereiner, G.; Degrouard, J.; Carriere, M.; Gouget, B. Uranium and selenium resistance in *Cupriavidus metallidurans* CH34. *Toxicol. Lett.* **2007**, *172*, S157. [CrossRef]

24. Monsieurs, P.; Moors, H.; Van Houdt, R.; Janssen, P.J.; Janssen, A.; Coninx, I.; Mergeay, M.; Leys, N. Heavy metal resistance in *Cupriavidus metallidurans* CH34 is governed by an intricate transcriptional network. *Biometals* **2011**, *24*, 1133–1151. [CrossRef] [PubMed]

25. Ben Salem, I.; Sghaier, H.; Monsieurs, P.; Moors, H.; Van Houdt, R.; Fattouch, S.; Saidi, M.; Landolsi, A.; Leys, N. Strontium-induced genomic responses of *Cupriavidus metallidurans* and strontium bioprecipitation as strontium carbonate. *Ann. Microbiol.* **2013**, *63*, 833–844. [CrossRef]

26. Van Houdt, R.; Mergeay, M. Genomic context of metal response genes in *Cupriavidus metallidurans* with a focus on strain CH34. In *Metal Response in Cupriavidus Metallidurans: Volume I: From Habitats to Genes and Proteins*; Mergeay, M., Van Houdt, R., Eds.; Springer International Publishing: Cham, Switzerland, 2015; pp. 21–44.

27. Vandenbussche, G.; Mergeay, M.; Van Houdt, R. Metal response in *Cupriavidus metallidurans*: Insights into the structure-function relationship of proteins. In *Metal Response in Cupriavidus Metallidurans: Volume II: Insights into the Structure-Function Relationship of Proteins*; Springer International Publishing: Cham, Switzerland, 2015; pp. 1–70.

28. Nies, D.H. The biological chemistry of the transition metal "transportome" of *Cupriavidus metallidurans*. *Metallomics* **2016**, *8*, 481–507. [CrossRef] [PubMed]

29. Millacura, F.A.; Cardenas, F.; Mendez, V.; Seeger, M.; Rojas, L.A. Degradation of benzene by the heavy-metal resistant bacterium *Cupriavidus metallidurans* CH34 reveals its catabolic potential for aromatic compounds. *bioRxiv* **2017**. [CrossRef]

30. Smalla, K.; Haines, A.S.; Jones, K.; Krogerrecklenfort, E.; Heuer, H.; Schloter, M.; Thomas, C.M. Increased abundance of IncP-1 beta plasmids and mercury resistance genes in mercury-polluted river sediments: First discovery of IncP-1 beta plasmids with a complex *mer* transposon as the sole accessory element. *Appl. Environ. Microb.* **2006**, *72*, 7253–7259. [CrossRef] [PubMed]

31. Rojas, L.A.; Yanez, C.; Gonzalez, M.; Lobos, S.; Smalla, K.; Seeger, M. Characterization of the metabolically modified heavy metal-resistant *Cupriavidus metallidurans* strain MSR33 generated for mercury bioremediation. *PLoS ONE* **2011**, *6*, e17555. [CrossRef] [PubMed]

32. Van Houdt, R.; Monchy, S.; Leys, N.; Mergeay, M. New mobile genetic elements in *Cupriavidus metallidurans* CH34, their possible roles and occurrence in other bacteria. *Antonie Leeuwenhoek* **2009**, *96*, 205–226. [CrossRef] [PubMed]

33. Mijnendonckx, K.; Provoost, A.; Monsieurs, P.; Leys, N.; Mergeay, M.; Mahillon, J.; Van Houdt, R. Insertion sequence elements in *Cupriavidus metallidurans* CH34: Distribution and role in adaptation. *Plasmid* **2011**, *65*, 193–203. [CrossRef] [PubMed]

34. Van Houdt, R.; Monsieurs, P.; Mijnendonckx, K.; Provoost, A.; Janssen, A.; Mergeay, M.; Leys, N. Variation in genomic islands contribute to genome plasticity in *Cupriavidus metallidurans*. *BMC Genom.* **2012**, *13*, 111. [CrossRef] [PubMed]

35. Mergeay, M.; Nies, D.; Schlegel, H.G.; Gerits, J.; Charles, P.; Van Gijsegem, F. *Alcaligenes eutrophus* CH34 is a facultative chemolithotroph with plasmid-bound resistance to heavy metals. *J. Bacteriol.* **1985**, *162*, 328–334. [PubMed]

36. Maniatis, T.; Fritsch, E.F.; Sambrook, J. *Molecular Cloning: A Laboratory Manual*; Cold Spring Harbor Laboratory: Cold Spring Harbor, NY, USA, 1982.

37. Kovach, M.E.; Elzer, P.H.; Hill, D.S.; Robertson, G.T.; Farris, M.A.; Roop, R.M.; Peterson, K.M. Four new derivatives of the broad-host-range cloning vector pBBR1MCS, carrying different antibiotic-resistance cassettes. *Gene* **1995**, *166*, 175–176. [CrossRef]

38. Taghavi, S.; Vanderlelie, D.; Mergeay, M. Electroporation of *Alcaligenes eutrophus* with (mega) plasmids and genomic DNA fragments. *Appl. Environ. Microb.* **1994**, *60*, 3585–3591.

39. Lee, C.; Kim, J.; Shin, S.G.; Hwang, S. Absolute and relative qPCR quantification of plasmid copy number in *Escherichia coli*. *J. Biotechnol.* **2006**, *123*, 273–280. [CrossRef] [PubMed]

40. Bolger, A.M.; Lohse, M.; Usadel, B. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* **2014**, *30*, 2114–2120. [CrossRef] [PubMed]

41. Li, H.; Handsaker, B.; Wysoker, A.; Fennell, T.; Ruan, J.; Homer, N.; Marth, G.; Abecasis, G.; Durbin, R.; 1000 Genome Project Data Processing Subgroup. The sequence alignment/map format and SAMtools. *Bioinformatics* **2009**, *25*, 2078–2079. [CrossRef] [PubMed]

42. Quinlan, A.R. BEDTools: The Swiss-army tool for genome feature analysis. *Curr. Protoc. Bioinform.* **2014**, *47*, 11–34. [CrossRef] [PubMed]

43. Li, H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv*, 2013; arXiv:1303.3997.

44. Norberg, P.; Bergstrom, M.; Jethava, V.; Dubhashi, D.; Hermansson, M. The IncP-1 plasmid backbone adapts to different host bacterial species and evolves through homologous recombination. *Nat. Commun.* **2011**, *2*, 268. [CrossRef] [PubMed]

45. Siguier, P.; Gourbeyre, E.; Chandler, M. Bacterial insertion sequences: Their genomic impact and diversity. *FEMS Microbiol. Rev.* **2014**, *38*, 865–891. [CrossRef] [PubMed]

46. Grass, G.; Grosse, C.; Nies, D.H. Regulation of the *cnr* cobalt and nickel resistance determinant from *Ralstonia* sp. strain CH34. *J. Bacteriol.* **2000**, *182*, 1390–1398. [CrossRef] [PubMed]

47. Tibazarwa, C.; Wuertz, S.; Mergeay, M.; Wyns, L.; van Der Lelie, D. Regulation of the *cnr* cobalt and nickel resistance determinant of *Ralstonia eutropha* (*Alcaligenes eutrophus*) CH34. *J. Bacteriol.* **2000**, *182*, 1399–1409. [CrossRef] [PubMed]

48. Kim, E.H.; Nies, D.H.; McEvoy, M.M.; Rensing, C. Switch or funnel: How RND-type transport systems control periplasmic metal homeostasis. *J. Bacteriol.* **2011**, *193*, 2381–2387. [CrossRef] [PubMed]

49. Collard, J.M.; Provoost, A.; Taghavi, S.; Mergeay, M. A new type of *Alcaligenes eutrophus* CH34 zinc resistance generated by mutations affecting regulation of the *cnr* cobalt-nickel resistance system. *J. Bacteriol.* **1993**, *175*, 779–784. [CrossRef] [PubMed]

50. Vandecraen, J.; Monsieurs, P.; Mergeay, M.; Leys, N.; Aertsen, A.; Van Houdt, R. Zinc-induced transposition of insertion sequence elements contributes to increased adaptability of *Cupriavidus metallidurans*. *Front. Microbiol.* **2016**, *7*, 359. [CrossRef] [PubMed]

51. Zheng, H.; Chruszcz, M.; Lasota, P.; Lebioda, L.; Minor, W. Data mining of metal ion environments present in protein structures. *J. Inorg. Biochem.* **2008**, *102*, 1765–1776. [CrossRef] [PubMed]

52. Thilakaraj, R.; Raghunathan, K.; Anishetty, S.; Pennathur, G. In silico identification of putative metal binding motifs. *Bioinformatics* **2007**, *23*, 267–271. [CrossRef] [PubMed]

53. Kirsten, A.; Herzberg, M.; Voigt, A.; Seravalli, J.; Grass, G.; Scherer, J.; Nies, D.H. Contributions of five secondary metal uptake systems to metal homeostasis of *Cupriavidus metallidurans* CH34. *J. Bacteriol.* **2011**, *193*, 4652–4663. [CrossRef] [PubMed]

54. Anton, A.; Grosse, C.; Reissmann, J.; Pribyl, T.; Nies, D.H. Czcd is a heavy metal ion transporter involved in regulation of heavy metal resistance in *Ralstonia* sp. strain CH34. *J. Bacteriol.* **1999**, *181*, 6876–6881. [PubMed]

55. Munkelt, D.; Grass, G.; Nies, D.H. The chromosomally encoded cation diffusion facilitator proteins DmeF and FieF from *Wautersia metallidurans* CH34 are transporters of broad metal specificity. *J. Bacteriol.* **2004**, *186*, 8036–8043. [CrossRef] [PubMed]

56. Barber-Zucker, S.; Shaanan, B.; Zarivach, R. Transition metal binding selectivity in proteins and its correlation with the phylogenomic classification of the cation diffusion facilitator protein family. *Sci. Rep.* **2017**, *7*, 16381. [CrossRef] [PubMed]

57. Scherer, J.; Nies, D.H. CzcP is a novel efflux system contributing to transition metal resistance in *Cupriavidus metallidurans* CH34. *Mol. Microbiol.* **2009**, *73*, 601–621. [CrossRef] [PubMed]

58. Delmar, J.A.; Su, C.-C.; Yu, E.W. Bacterial multi-drug efflux transporters. *Annu. Rev. Biophys.* **2014**, *43*, 93–117. [CrossRef] [PubMed]

59. De Angelis, F.; Lee, J.K.; O' Connell, J.D.; Miercke, L.J.W.; Verschueren, K.H.; Srinivasan, V.; Bauvois, C.; Govaerts, C.; Robbins, R.A.; Ruysschaert, J.M.; et al. Metal-induced conformational changes in ZneB suggest an active role of membrane fusion proteins in efflux resistance systems. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 11038–11043. [CrossRef] [PubMed]

60. Long, F.; Su, C.C.; Lei, H.T.; Bolla, J.R.; Do, S.V.; Yu, E.W. Structure and mechanism of the tripartite CusCBA heavy-metal efflux complex. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **2012**, *367*, 1047–1058. [CrossRef] [PubMed]

61. Pak, J.E.; Ekende, E.N.; Kifle, E.G.; O'Connell, J.D.; De Angelis, F.; Tessema, M.B.; Derfoufi, K.M.; Robles-Colmenares, Y.; Robbins, R.A.; Goormaghtigh, E.; et al. Structures of intermediate transport states of ZneA, a Zn(II)/proton antiporter. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 18484–18489. [CrossRef] [PubMed]

62. Sota, M.; Tsuda, M.; Yano, H.; Suzuki, H.; Forney, L.J.; Top, E.M. Region-specific insertion of transposons in combination with selection for high plasmid transferability and stability accounts for the structural similarity of IncP-1 plasmids. *J. Bacteriol.* **2007**, *189*, 3091–3098. [CrossRef] [PubMed]

63. Leys, N.; Baatout, S.; Rosier, C.; Dams, A.; s' Heeren, C.; Wattiez, R.; Mergeay, M. The response of *Cupriavidus metallidurans* CH34 to spaceflight in the international space station. *Antonie Leeuwenhoek* **2009**, *96*, 227–245. [CrossRef] [PubMed]

64. Mijnendonckx, K.; Provoost, A.; Ott, C.M.; Venkateswaran, K.; Mahillon, J.; Leys, N.; Van Houdt, R. Characterization of the survival ability of *Cupriavidus metallidurans* and *Ralstonia pickettii* from space-related environments. *Microb. Ecol.* **2013**, *65*, 347–360. [CrossRef] [PubMed]

65. Byloos, B.; Coninx, I.; Van Hoey, O.; Cockell, C.; Nicholson, N.; Ilyin, V.; Van Houdt, R.; Boon, N.; Leys, N. The impact of space flight on survival and interaction of *Cupriavidus metallidurans* CH34 with basalt, a volcanic moon analog rock. *Front. Microbiol.* **2017**, *8*, 671. [CrossRef] [PubMed]

66. Vasseur, P.; Vallet-Gely, I.; Soscia, C.; Genin, S.; Filloux, A. The *pel* genes of the *Pseudomonas aeruginosa* PAK strain are involved at early and late stages of biofilm formation. *Microbiology* **2005**, *151*, 985–997. [CrossRef] [PubMed]

67. Godoy, V.G.; Fox, M.S. Transposon stability and a role for conjugational transfer in adaptive mutability. *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 7393–7398. [CrossRef] [PubMed]

68. Christie-Oleza, J.A.; Lanfranconi, M.P.; Nogales, B.; Lalucat, J.; Bosch, R. Conjugative interaction induces transposition of ISPst9 in *Pseudomonas stutzeri* AN10. *J. Bacteriol.* **2009**, *191*, 1239–1247. [CrossRef] [PubMed]

69. Baharoglu, Z.; Bikard, D.; Mazel, D. Conjugative DNA transfer induces the bacterial SOS response and promotes antibiotic resistance development through integron activation. *PLoS Genet.* **2010**, *6*, e1001165. [CrossRef] [PubMed]

70. Jain, A.; Srivastava, P. Broad host range plasmids. *FEMS Microbiol. Lett.* **2013**, *348*, 87–96. [CrossRef] [PubMed]

71. Obranic, S.; Babic, F.; Maravic-Vlahovicek, G. Improvement of pBBR1MCS plasmids, a very useful series of broad-host-range cloning vectors. *Plasmid* **2013**, *70*, 263–267. [CrossRef] [PubMed]

72. Tian, P.; Wang, J.; Shen, X.; Rey, J.F.; Yuan, Q.; Yan, Y. Fundamental CRISPR-cas9 tools and current applications in microbial systems. *Synth. Syst. Biotechnol.* **2017**, *2*, 219–225. [CrossRef] [PubMed]

73. Cook, T.B.; Rand, J.M.; Nurani, W.; Courtney, D.K.; Liu, S.A.; Pfleger, B.F. Genetic tools for reliable gene expression and recombineering in *Pseudomonas putida*. *J. Ind. Microbiol. Biotechnol.* **2018**, *45*, 517–527. [CrossRef] [PubMed]

74. Xiong, B.; Li, Z.; Liu, L.; Zhao, D.; Zhang, X.; Bi, C. Genome editing of *Ralstonia eutropha* using an electroporation-based CRISPR-cas9 technique. *Biotechnol. Biofuels* **2018**, *11*, 172. [CrossRef] [PubMed]

Table 11.1: Primers designed and synthesised for this thesis

Numb	Name	
F208	Fw-primer-T7-Spi2-BglII	
F209	Rv-Primer-Spi2-BmtI	
F210	Fw-primerT7-iSpi-BglII	
F211	Rv-Primer-iSpi-BmtI	
F212	Fw-F30-2xdBroccolicheck	
F213	Fw-F30-Broccolicheck	
F214	Fw-Spinach2check	
F215	Fw-iSpinachcheck	
F216	GGA-XbaI-T7-promCpf1crRNA-Fw	
F217	GGA-XbaI-T7-promCpf1crRNA-Rv	
F218	GGA-T7-prom-Cpf1crRNAFw	
F219	GGA-T7-prom-Cpf1-crRNA-Rv	
F220	GGA-As-crRNA-Fw	
F221	GGA-As-crRNA-Rv	
F222	GGA-Lb-crRNA-Fw	
F223	GGA-Lb-crRNA-Rv	
F224	GGA-spacer-Fw	
F225	GGA-spacer-Rv	
F226	GGA-T7-terminator-Fw	
F227	GGA-T7-terminator-Rv	
F228	GGA-T7-terminator-PstI-Fw	
F229	GGA-T7-terminator-PstI-Rv	
F230	EFX-Lb-Cpf1-Fw	
F231	EFX-Lb-Cpf1-Rv	
F232	EFX-As-Cpf1-Fw	
F233	EFX-As-Cpf1-Rv	
F234	GGA-spacer-Fw	
F235	GGA-spacer-Rv	
F236	GGA-T7-terminator-PstI-fixed-scar-Fw	
F237	GGA-T7-terminator-PstI-fixed-scar-RV	
F238	GGA-XbaI-F30-scaffold-part-PstI-Fw	CTAGaGGTCTCA
F239	GGA-XbaI-F30-scaffold-part-PstI-Rv	GGTCTCTT
F240	Cyan-RNA-2-4-Fw	TGTGGGAGACGCAACTGA
F241	Cyan-RNA-2-4-Rv	GTGGGGAGACGCGACTA
F242	Red-RNA-6-8-Fw	
F243	Red-RNA-6-8-Fw	
F244	Yellow-RNA-17-3-Fw	
F245	Yellow-RNA-17-3-Rv	
F246	F30-scaffold-Dw-PstI-Fw	
F247	F30-scaffold-Dw-PstI-Rv	
F248	XbaI-F30-scaffold-Up-Fw	

Table 11.2: ParAlleL 3 bit Full adder/subtractor design

		Adder		Subtractor	
Input	Antibiotic	Cout	S	Bo	D
000	ATG	—	—	—	—
001	ATS	—	+++	+++	+++
010	AKG	—	+++	+++	+++
011	AKS	+++	—	+++	—
100	CTG	—	+++	—	+++
101	CTS	+++	—	—	—
010	CKG	+++	—	—	—
111	CTG	+++	+++	+++	+++

Cells added in specified wells carry resistance markers for:

ATG: Ampicilin, Tetracycline, Gentamicin.

ATS: Ampicilin, Tetracycline, Spectinomycin.

AKG: Ampicillin, Kanamycin, Gentamicin.

AKS: Ampicillin, Kanamycin, Spectinomycin.

CTG: Chloramphenicol, Tetracycline, Gentamicin.

CTS: Chloramphenicol, Tetracycline, Spectinomycin.

CKG: Chloramphenicol, Kanamycin, Gentamicin.

CKS: Chloramphenicol, Kanamycin, Spectinomycin.

Table 11.3: Plasmids used for generating ParAlleL subcircuit strains.

Plasmid	Antibiotic marker	ORI	Copy number	Reference
pSB4A5	AmpR	pSC101	5	http://parts.igem.org/Part:pSB
pSB4C5	CmlR	pSC101	5	http://parts.igem.org/Part:pSB
pSB1T3	TetR	pMB1(der)	100-300	http://parts.igem.org/Part:pSB
pSB1K3	KanR	pMB1(der)	100-300	http://parts.igem.org/Part:pSB
pSEVA631	GenR	pBBR1	medium	https://www.ncbi.nlm.nih.gov/nucco re
pMO9075	SpmR	pBG1	low	Keller, et al., 2011

Table 11.4: Antibiotics and resistance cassettes used on ParAlleL.

Antibiotic	Class	Mode of action
Ampicillin	β-lactam	Bactericidal; Inhibits cell wall synthesis
Kanamycin	Aminoglycoside	Bactericidal; Binds 30S ribosomal subunit; causes m
Chloramphenicol	Chloramphenicol	Bacteriostatic; Binds 50S ribosomal subunit; inhib
Tetracycline	Tetracycline	Bacteriostatic; Binds 16S ribosomal subunit; inhibits protein sy
Gentamicin	Aminoglycoside	Irreversibly binding the 30S subunit of the bacteri
Spectinomycin	Aminocyclitol	It binds to the 30S and interrupts protein synthesis aff

Table 11.5: Plasmid incompatibility groups.

Incompatibility group	Regulation	Comm
pBR322/ColE1/pMB1	Inhibitor-target RNA I	Control processing of pr
IncFII, pT181	RNA	Affecting synthesis
R6K*, P1, F, pSC101, Rts1, P15A*, RK2	Iteron binding	Sequestering of I