CGSENS: Visualization of CG methylation using a fluorescence protein biosensor

PRIMARY CONTACT FOR THE TEAM

Sara Lopez-Gomollon, Department of Plant Sciences, University of Cambridge
sl750@cam.ac.uk

TEAM

Dr Marino Exposito-Rodriguez (mexpos@essex.ac.uk), School of Biological Sciences, University of Essex.

Marino Exposito is a postdoc in Prof Mullineaux group. He has experience in molecular biology and protein biochemistry. He has broad experience using genetically-encoded biosensors based on engineered fluorescent proteins to monitor oxidative stress and pH changes in plants (*1*). He recently published a paper in *Nature Communications* where fluctuations of H_20_2 in plant cells were recorded *in vivo* and in real time using this technology (*2*). Currently, he is developing a new set of biosensors to monitor changes in lipid peroxidation and dehydroascorbic acid as part of a BBSRC grant.
https://www.researchgate.net/profile/Marino_Exposito/publications

Dr Sara Lopez-Gomollon (sl750@cam.ac.uk), Department of Plant Sciences, University of Cambridge.

Sara is a molecular biologist in Prof Baulcombe group working on the effect of small RNAs and the epigenetic changes produced by them in transgressive phenotypes observed in tomato hybrid populations as part of an ERC grant. She is interested in the variation of gene regulation and DNA methylation during development in plants. She is very experienced in molecular cloning, including GoldenGate, and in protein biochemistry.
https://www.researchgate.net/profile/Sara_Lopez-Gomollon2/publications

This project maps onto the aims of the OpenPlant fund because the Fund will "*support early stage investigation of new tools and technologies open source*". It is possible that these pilot studies will generate research that can be used to support subsequent applications for grants and fellowships. However, the primary focus of projects is expected to be "*technology development rather than data generation*". This project is expressly about technology development using "*low-cost instrumentation for quantifying markers in the DNA*" and will clearly deliver an "*improved technological approach*" allowing researchers "*to study methylation dynamics, further enhancing resolution, sensitivity and precision in a cell-free system*".

SUMMARY

Cytosine DNA methylation is an epigenetic mark critical in diverse biological functions, such as gene regulation or genome stability. So far, the cellular epigenetic landscape has been pictured using in-vitro approaches that provide information about DNA methylation at specific loci or genome-wide, but lacking an in vivo spatiotemporal resolution. The **aim** of this proof-of-concept project is to design and validate a fluorescent protein biosensor for DNA methylation in CG context. The validation of this fluorescent biosensor will facilitate its use as a tool to monitor fluctuations in DNA methylation in living organisms.

INTRODUCTION

DNA methylation is a key epigenetic modification in many organisms that plays a role in gene regulation, imprinting or genome stability. It involves the addition of a methyl group to the fifth carbon of cytosine. It is evolutionarily ancient and associated with gene silencing in eukaryotes. Defects in DNA methylation produce pleiotropic morphological effects in plants and embryonic lethality in

mammals. In plants, cytosine methylation is predominantly found in CG context, but also in CHG and CHH (with H = A, T, or C). **The methylation pattern is dynamic and tissue-specific, to accommodate modifications in gene regulation needed during plant developmen**t (*3*). These complex patterns of methylation highlight the significance of profiling DNA methylation to answer biological questions.

One of the most widely used techniques, genome-wide bisulphite sequencing, (*4*) provide single-nucleotide resolution of the DNA methylation state at a specific time point. (*5*). However, these assays are *in vitro* and require a considerable amount of starting material. Efforts have been made to develop alternative methods to monitor methylation changes using *in vivo* approaches (*6*). **Of particular interest is the development of non-invasive techniques that would allow us to monitor the spatiotemporal (single-cell resolution) variation of DNA methylation during development or in response to external stimuli.**

In vivo fluorescent protein biosensors that allow, by live imaging, the non-invasive dynamic quantification and spatial distribution of small molecules as well as protein-protein interactions and conformational change of receptors, have been widely used in a variety of model organisms and have the potential for widespread applications in fundamental biology and biotechnology. **The technology for *in vivo* biosensors, increasingly being developed out of synthetic biology design principles**, **is set to revolutionise our understanding of cellular metabolism** (*7*).

There are some examples of biosensors to visualise DNA methylation, however they have several limitations. An *in vitro* luciferase-based sensor of DNA Methylation has been developed for mammalian cells, which requires the presence of chemicals to induce fluorescence (Promega). In addition, this luciferase-based sensor cannot be used for *in vivo* live imaging at cellular resolution, therefore it cannot report cell-specific information regarding DNA methylation in different cell types. In a recent publication a biosensor was reported to monitor DNA methylation in plants, consisting of a methylated DNA (methyl-CpG) binding domain (MBD) linked to VENUS, an improved version of the Yellow Fluorescent Protein (YFP). The authors reported DNA methylation dynamics at single-cell level, but the fluorescence signal is emitted independently of the interaction of the biosensor to DNA. This results in a very low signal-to-noise ratio (SNR) (*4*), so subtle changes in the methylation dynamics may not be recorded.

AIM

We propose the design of a **new generation of biosensor (*CGSENS)* with higher SNR based on bimolecular fluorescence complementation (BiFC) to visualize protein-DNA interactions to study the global CG DNA methylation dynamics**. CGSENS will consist of a methyl-CpG binding domain (MBD) linked to the amino (CGSENS-N) or carboxyl (CGSENS-C) end of a split VENUS. When two methylated cytosines in a CG context are in proximity, they will be recognized by the MBD domains that will bring in close proximity the amino and carboxyl ends of Venus, resulting in a reassembly of the protein and concomitant fluorescence signal (Figure 1). To assay the interaction of the biosensor with DNA we will use a cell-free system, where the CGSENS will be incubated with DNA with different levels of methylation. The level of restored VENUS activity can be measured using a plate reader, and the fluorescence levels will be indicative of the percentage of DNA methylation (Figure 2).

Features of CGSENS biosensors:

- CGSENS biosensor is **not invasive and independent of any chemical**, which could facilitate its use as a genome-integrated biosensor in living organisms.
- It is based on **structural complementation** between two non-fluorescent N-terminal and C-terminal fragments of a fluorescent protein, minimizing the background signal and providing a specific signal and a high SNR.
- The fluorescence emission of the reconstructed VENUS moiety in CGSENS can be easily quantified using **standard fluorescence plate readers**.

OBJECTIVES AND METHODOLOGY

Objectives 1-3 – To test the principles of *in vitro* protein-methylated DNA interactions base on BiFC to study global CG DNA methylation

1.- **We will design and chemically synthesize CGSENS.** Following the principle of Venus-based BiFC (*8*) we will design two constructs named **CGSENS-N** and **CGSEN-C** (Figure 1). These constructs encode the methylated DNA (methyl-CpG) binding domain (MBD) of AthMBD6 (*10*) fused in-frame with N- or C-terminal BiFC fragments (N/C-Venus), respectively. We chose this MBD domain because it has been published that it interacts with methylated CG in vitro and in vivo (*4*, *11*), showing that its affinity for methylated cytosines in other contexts (CHH, CHG) is very limited. Also, MBD proteins are very specific for methylated CG (at the nanomolar scale) compared to CG (micromolar scale) (*4*, *12*). DNA sequences for CGSENS-N and CGSENS-C will be domesticated so they would not contain any BsaI or BpiI recognition sites to facilitate further GoldenGate cloning. Sequences and elements designed will be GoldenGate compatible when possible (*9*) and the plasmids created in this project will be deposited in the non-profitable plasmid database AddGene (https://www.addgene.org). ***Months 0-1.***

2.- **We will express and purify CGSENS recombinant protein from *E. coli***. For the biosensors to be made in this project, we will take a well-established strategy for producing recombinant protein. We will express the recombinant proteins in *E. coli* using 6His-tagging to N termini of CGSENS-N/C via commercially synthesised pET plasmids. The protocol for purification of His-tagged proteins by Ni^{2+}NTA affinity has been extensively used by Dr. Exposito. Protein purity and integrity will be checked by PAGE and Western blotting against the His tag. In order to avoid any interference from the His tag, we will introduce a suitable amino acid sequence for the removal of the polyhistidine-tag using Factor Xa. Factor Xa protease cleaves off the His tag peptide behind the arginine residue of the protease recognition site (IEGR) and results in a recombinant protein free of any vector-derived amino acids at the N-terminus. After digestion, Factor Xa Protease can be removed using Xa Removal Resin. Protein purity and integrity will be checked by PAGE. ***Months 2-4.***

3.- **The interaction of CGSENS to DNA will be characterised by *in vitro* binding experiments**. DNA reference samples with different degrees of methylation will be used to establish the dynamic range of the biosensor. Having purified recombinant GCSENS-N and CGSENS-C, we will initially test the biosensor for its ability to bind methylated CG in the presence of unmethylated or fully methylated genomic DNA (Sigma-Aldrich). Titration experiments will be done to evaluate the functional dynamic range of CGSENS using a mix of different proportions of fully methylated and unmethylated DNA to obtain a range from 0-100% methylated DNA (Figure 2). The experiment will be done in a 96-well plate and the affinity recorded as fluorescence signal quantified on a fluorescence plate reader***. Month 5-6.***

FACILITIES

The laboratories in the School of Biological Sciences, University of Essex (of which MER is member), are well equipped for molecular biology, recombinant protein expression and protein purification. They are well-equipped for *in vitro* assays of protein function using individual samples and plate-reader formats.

The laboratories in the Plant Sciences Department, University of Cambridge (of which SLG is member), are well equipped for molecular biology and protein biochemistry. In the department there are fluorescence plant readers that can be used for titration experiments.

BENEFITS AND OUTCOMES

At the end of this project we expect to have *in vitro* validation of the functionality of CGSENS, a new biosensor able to detect methylated cytosines in close proximity. The biosensor is novel because it is specific and standalone. It is specific because the signal is obtained only when the biosensor is recognising CG methylation, and it is standalone because the fluorescence is not due to the presence

of additional chemicals. We can foresee future *in vivo* applications for GCSENS. First by transient experiments in *N. benthamiana* and transfection in mammal cells we expect to validate the results obtained *in vitro* by *in vivo* approaches. The next step will be to obtain transgenic lines where GCSENS is expressed to monitor changes in DNA methylation during development or in response to external stimuli.

This project fits in the OpenPlant interests as it is an original project to create tools for plant biology. The outcome is tangible and open: at the end of the project we will have created a set of plasmids, GoldenGate compatible and be available for the community in open repositories. Information about the plasmids will be available in the OpenPlant website.

This project also benefits from being interdisciplinary and innovative, and for creating new collaborations between the University of Essex and the University of Cambridge. Dr. Exposito will bring the expertise in biosensor design and Dr. Lopez-Gomollon will bring the experience in epigenetics and molecular biology.

SPONSOR FOR THE RESEARCH AND COST CENTRE

Prof Sir David Baulcombe (dcb40@cam.ac.uk), Department of Plant Sciences, University of Cambridge.

I confirm that I have the full support of the sponsor listed above and that they can be added to the OpenPlant Fund mailing list to receive project updates (to which they can unsubscribe at any time).

BUDGET

PRODUCT	COMPANY	CODE	PRICE
Synthesis of DNA elements for CGSENS-N and CGSENS-C (G-Block gene fragment)	IDT		GBP 220.00
pQE-30 Xa vector (25 ug)	Qiagen	33203	GBP 269.00
Ni-NTA Spin Kit (50 Spin columns)	Qiagen	31314	GBP 367.00
Ni-NTA Superflow (25 ml nickel-charged resin)	Qiagen	30450	GBP 332.00
Factor- Xa Protease 250 ug (1mg/ml)	NEB	P8010L	GBP 228.00
Xa Removal Resin (2 x 2.5 ml)	Qiagen	33213	GBP 350.00
NuPAGE™ 4-12% Bis-Tris Protein Gels, 1.0 mm, 10-well (2 boxes)	Invitrogen	NP0321BOX	GBP 244.00
PageRuler™ Plus Prestained Protein Ladder, 10 to 250 kDa	Thermo	26619	GBP 110.00
Anti-6X His tag® antibody [HIS.H8]	Abcam	Ab18184	GBP 285.00
Gilson PIPETMAN Neo 8 channel P8x20N 2-20uL	Gilson	Ab18184	GBP 285.00
CpGenome™ Rat Methylated Genomic DNA Standard (5ug) (2 vials)	Sigma-Aldrich	F14401	GBP 552.00
CpGenome™ Rat Unmethylated Genomic DNA Standard (5ug) (2 vials)	Sigma-Aldrich	S7860	GBP 322.00
TOTAL AMOUNT			**GBP 3621.00**

Dr Exposito has additional funding from a BBSRC project.

BIBLIOGRAPHY

1. M. Exposito-Rodriguez, P. P. Laissue, G. R. Littlejohn, N. Smirnoff, P. M. Mullineaux, The use of HyPer to examine spatial and temporal changes in H2O2 in high light-exposed plants. *Methods Enzymol.* **527**, 185–201 (2013).
2. M. Exposito-Rodriguez, P. P. Laissue, G. Yvon-Durocher, N. Smirnoff, P. M. Mullineaux, Photosynthesis-dependent H 2 O 2 transfer from chloroplasts to nuclei provides a high-light

signalling mechanism. *Nat Commun.* **8**, 4 (2017).
3. J. A. Law, S. E. Jacobsen, Establising, maintaining and modifying DNA methylation patterns in plants and animals. *Nat Rev Genet.* **11**, 204–220 (2011).
4. M. Ingouff *et al.*, Live-cell analysis of DNA methylation during sexual reproduction in Arabidopsis reveals context and sex-specific dynamics controlled by noncanonical RdDM. *Genes Dev.* **31**, 72–83 (2017).
5. W.-S. Yong, F.-M. Hsu, P.-Y. Chen, Profiling genome-wide DNA methylation. *Epigenetics Chromatin*. **9**, 26 (2016).
6. A. H. Badran *et al.*, Evaluating the global CpG methylation status of native DNA utilizing a bipartite split-luciferase sensor. *Anal. Chem.* **83**, 7151–7157 (2011).
7. A. J. Meyer, T. P. Dick, Fluorescent Protein-Based Redox Probes. *Antioxid. Redox Signal.* **13**, 621–650 (2010).
8. K. Ohashi, T. Kiuchi, K. Shoji, K. Sampei, K. Mizuno, Visualization of cofilin-actin and Ras-Raf interactions by bimolecular fluorescence complementation assays using a new pair of split Venus fragments. *Biotechniques*. **52**, 45–50 (2012).
9. C. Engler *et al.*, A Golden Gate Modular Cloning Toolbox for Plants. *ACS Synth. Biol.* **3**, 140220125704008 (2014).
10. A. Zemach *et al.*, The arabidopsis nucleosome remodeler DDM1 allows DNA methyltransferases to access H1-containing heterochromatin. *Cell*. **153**, 193–205 (2013).
11. A. Zemach, G. Grafi, Characterization of Arabidopsis thaliana methyl-CpG-binding domain (MBD) proteins. *Plant J.* **34**, 565–572 (2003).
12. M. F. Fraga *et al.*, The affinity of different MBD proteins for a specific methylated locus depends on their intrinsic binding properties. *Nucleic Acids Res.* **31**, 1765–74 (2003).