Development of new codon optimization tools and development of a synthetic gene expression system in the green alga Chlamydomonas reinhardtii

Summary: Most organisms share the same genetic code, based on three nucleotide codons that encode for one amino acid. However, synonymous codons (which specify a single amino acid) are not used at equal frequency by different species. We are interested in assessing the impact of codon usage in protein production in the green alga Chlamydomonas reinhardtii. We have performed sequence analysis, and developed a platform for measuring the production of a reporter protein, which can be used for testing gene variants. The analysis and materials generated will be useful for transgene design and expression in the alga.

Report and Outcomes:

Project Report Introduction Chlamydomonas reinhardtii is a useful model for plant biotechnology, especially algae. However, the unusual high GC content of Chlamydomonas coding sequences (68%, while Arabidopsis is 44% and human is 52%), strongly biases codon selection towards GC-rich codons, making codon optimization a necessary step for the expression in the alga of genes of other species. Codon optimization consists in replacing infrequent codons of a transgene sequence by synonymous codons used at higher frequency in the host species. This is a necessary step because tRNA molecules abundance is different among species and has co-evolved with codon frequency to optimize translation efficiency. A correlation between codon frequency and the abundance of the isoacceptor tRNA molecules has been reported for a number of organisms, both prokaryotes and eukaryotes (Angov, 2011; Ikemura, 1982). In the case of Chlamydomonas, although tRNA abundance has not been systematically analyzed, there is a good correlation between codon frequency and copy number of the corresponding isoacceptor tRNA gene (Cognat et al., 2008). The majority of codon optimization algorithms use a set of coding sequences of the host to compute codon usage reference values, irrespective of the expression level or translation efficiency of these sequences. This strategy has proven to increase translation efficiency, most likely by the removal of infrequent codons for which the amount of the acceptor tRNA molecule is limiting. However, less is known about the impact in expression level when different sets of genes are used to calculate the reference codon usage, or when the level of codon adaptation varies along the coding sequence. In order to explore codon usage in C. reinhardtii, and to provide the community with a series of tools for codon optimization and protein production, we proposed the following objectives for this OpenPlant fund: 1) to create a web site, with links to tools of codon optimization, describing their utility, and also with links to DNA synthesis companies, including user reviews, 2) to analyze codon usage of different sets of genes, based on expression level, GO term, and pattern of expression, and 3) to assess the impact of codon usage in the production of a protein reporter . During these six months, we have been able to make progress on points 2 and 3, and a detailed description of the work and results is presented below. We think that the website project proposed in point 1 could be useful for synthetic biologists and we

would like to start working on it. We ask for the extra 1000 £ funding to cover website creation, diffusion and maintenance. The remaining funding will be used for DNA synthesis to expand the repertoire of Golden Gate-domesticated DNA parts that will be available to the Chlamydomonas community.

Results 1. CODON USAGE IN Chlamydonas reinhardii SEQUENCES

1.1.General features of codon usage in Chlamydomonas The high GC content of the Chlamydomonas reinhardtii nuclear genome strongly biases codon usage towards G/C rich codons (Table 1). A close look at the data available in http://www.kazusa.or.jp/codon, which is used by most codon optimization algorithms as reference set, reveals a strong bias towards G or C at the wobble position (3rd letter of codon). Codons of nuclear coding sequences end in a G or a C in 86.2 % of the cases. This contrasts with chloroplast gene codons, which are relatively GC-poor (Table 1), and with STOP codons, for which the codon TAA is used in most cases (1.05‰), followed by TGA (0.54 ‰) and TAG (0.44 ‰) (Table 2, "kazusa table"). In addition to the strong selection of high GC-content codons, a number of amino acids shows bias towards one particular GC-rich synonymous codon (e.g. in the case of Gly, GGC is used in 61.99‰ of the cases, while GGG is used in 9.72‰). This preference correlates with the copy number of the isoaceptor tRNA genes, 17 to 1 copy of the corresponding Gly-tRNAs genes (Cognat et al., 2008).

1.2. Codon usage in different gene datasets The previous data of codon usage comes from the web site http://www.kazusa.or.jp/codon/, and it was calculated from 846 CDSs. The sequences that are used to compute these values are selected only by the quality of the sequence data, and not by their similarity in expression level or translation efficiency. We then were interested in testing whether codon usage changes when it is computed from genes with high translation efficiency exclusively, or from genes with a common pattern of expression. There is no genome-wide data of translation efficiency in C. reinhardtii so far; however, data from other organisms shows a positive correlation between mRNA levels and translation efficiency (Gingold & Pilpel, 2011). Similarly, genes related to translation and ribosome biogenesis/structure are among the most efficiently translated genes (Ingolia, Ghaemmaghami, Newman, & Weissman, 2009). In order to obtain an insight into the differential use of codons in C. reinhardtii, we computed codon usage from the following gene datasets: -Kazusa codon usage table (Nakamura, Gojobori, & Ikemura, 1999). -All CDSs, obtain from primary transcripts only (17,741 sequences) -Highly expressed genes, (see below for their identification) -Genes associated with the GO term "structural constituent of ribosome" -Genes associated with the GO term "ribosome biogenesis" -Genes associated with the GO term "translation" -Genes associated with the GO term "photosynthesis" -Nitrogen starvation up-regulated transcripts (Schmollinger et al., 2014) -Nitrogen starvation up-regulated proteins (Schmollinger et al., 2014) -Nitrogen starvation-down regulated proteins(Schmollinger et al., 2014) -Clusters 1 to 18 of cell cycle regulated transcripts during Chlamydomonas life cycle (Zones, Blaby, Merchant, & Umen, 2015) -P-type gamete specific genes (Lopez et al., 2015) -M-type gamete specific genes (Lopez et al., 2015) -Zygote-specific genes (Lopez et al., 2015) Codon usage was computed for each gene dataset, and the total number and frequency (‰) of each codon is given in Table 2.

For each amino acid, the synonymous codon with the highest frequency is highlighted. Codon usage was calculated using C. reinhardtii primary transcripts only (Creinhardtii_281_v5.5.cds_primaryTranscriptOnly.fa, 17,741 sequences), and gene sequences were downloaded from the most recent genome database (v5.5). Differentially expressed genes were obtained from published work, and Phytozome website was used to search for genes associated with specific GO terms. Highly expressed genes were identified in the following way: Tentative Consensus (TC) sequences and constitutive ESTs data were retrieved from the Dana Farber Cancer Institute (ftp://occams.dfci.harvard.edu/pub/bio/tgi/data/Chlamydomonas_reinhardtii/). TCs were blasted against C. reinhardtii primary transcripts and the top hits were kept. Similarly, genes were ranked based on their normalized read counts from an RNA-Seq experiment (Valli et al., 2016). For the latter dataset, read counts from three biological repeats were averaged. Both ESTs and RNA-Seq read counts were normalized by the length of the CDS, and approximately the 1% top genes in each list (167 and 196 genes, respectively) were compared. 116 genes were common to both top lists, and were defined as "highly expressed genes". RNA for the RNA-Seq data was extracted from cells growing in TAP medium, 25°C, continuous light. We do not know about the growth conditions used for the EST experiment, but the high overlap with the RNA-Seq data suggests that similar growth conditions were used. Table 2 shows that the majority of amino acids are encoded by a highly favoured codon among all synonymous codons, and the same codon is preferentially used in all gene datasets analyzed. However, we found four amino acids (Ala, Pro, Ser and Thr) for which the preferred codon changes in different gene datasets. Ala and Pro show a similar pattern: codons GCG (Ala) and CCG (Pro) are the most used codons in "All CDSs" dataset, while in the "Kazusa table", "Highly expressed genes", "Ribosome biogenesis", "Photosynthesis", "Translation" and "Structure constituents of the ribosome" datasets, these amino acids are preferentially encoded by GCC and CCC, respectively. The codon GCC is also preferentially used in the Nitrogen Starvation groups, while the codon CCC is also the most used in one Nitrogen Starvation group and two cell cycle clusters. These alternative preferential codons are recognized by different tRNA molecules, indicating that there is an actual change in the tRNA pool selection among gene groups. Amino acids Ser and Thr use two different preferred codons in different datasets: the "Highly expressed genes" and "Structural constituents of the ribosome" use the codon TCG for Ser more frequently than the codon AGC, which is more frequent in the Kazusa reference table and "All CDSs" datasets. In the case of Thr, only the first cluster of genes of the cell cycle uses slightly more frequently the codon ACG, although the alternative codon ACG, which is dominant in the other gene datasets, has a similar frequency. Based on this analysis, we conclude that codon selection in C. reinhartii is very robust, characterized by a strong preference towards GC-rich codons and with only few alternative preferred codons. Differential codon usage only occurs for the amino acids Ala, Pro, Ser and Thr. Regarding codon optimization of transgene sequences, as most algorithms use Kazusa codon usage as reference, codon selection mirrors the codon selection of genes which potentially are translated with high efficiency ("Highly expressed genes" and ribosome/translation associated genes in our list), except in the case of the amino acid Ser.

1.3. Codon selection along coding sequences Codon optimization tools recode a transgene sequence using only favourite synonymous codons for each amino acid. However, endogenous sequences with such an ideal codon selection are unlikely to occur. We analyzed how distant is codon selection to the most frequent codon usage for each coding sequences of C. reinhardtii by calculating the codon adaptation index (CAI). The CAI of each gene is computed as the geometric mean of the relative adaptiveness () of all the codons in the coding sequence. The relative adaptiveness for codon of kind i that codes for amino acid j is defined as i = Xij / Xjmax , where Xij is the number of occurrences of codon i and Xjmax is the maximum Xij for amino acid j. Then CAI is calculated using the formula: where i(l) is the relative adaptiveness for the l-th codon in the gene and L is the number of codons in the gene (Sharp & Li, 1987). Figure 1A shows the frequency distribution of the codon adaptation index (CAI) for all genes in C. reinhardtii, using the codon usage of "All CDSs" as reference value. No gene has a CAI value of 1 and, therefore, there is no gene that is encoded only by favourite codons. Interestingly, when CAI is calculated only for the "Highly expressed genes" dataset, the distribution is shifted towards 1, indicating that codon usage of this group is closer to the preferential codon selection (Figure 1B). We also looked at a group of genes with no reads in the RNA-Seq experiment ("unmapped gene"), which probably comprises pseudogenes and genes with very low expression. The CAI distribution for this gene group look similar to the distribution of "All CDSs". We then examined how codon usage changes along the coding sequences. Analysis of the "Highly expressed genes" dataset revealed that unfavourable codons are interspersed within stretches of favourite codons, as it is shown in Figure 2A with an example of a coding sequence where the relative adaptiveness value (j) for each codon is indicated. Figure 2B shows a heatmap of j values for the first 200 codons of sequences longer than 600 nts from the "Highly expressed genes" dataset. We noticed that the frequency of unfavourable codons was specially higher in the 5' end of the ORF, which can be seen more clearly by averaging the j value (Figure 2C). We observe that the first codon after the ATG tends to be an unfavourable codon, and then codon adaptiveness increases along the first one hundred codon stretch, creating a sort of "ramp" of codon adaptiveness. This ramp is also observed in the set of "unmapped genes", although it is less pronounced than in the "Highly expressed genes" dataset. The latter observation of a stretch of incremental codon adaptiveness at the beginning of coding sequences could be explained by a mis-annotation of C. reinhardtii ORFs start codons, and these regions being actually part of the 5' UTRs and not ORF. However, similar stretches of low codon adaptiveness have been found in coding sequences of other species, and they have been proposed to act as ramps of translation, regulating ribosome trafficking along the mRNA molecule (Tuller et al., 2010). In this model, the first 30 to 50 residues of a protein are encoded by unfavourable codons that slow down translation initiation, allowing more efficient ribosome binding on the mRNA molecule and avoiding ribosome bottlenecking. Additionally, they could also have a role in coordinating translation with the in vivo folding of the protein.

1.4. Codon optimization of transgenes in C. reinhardtii. Our results show that, despite the high robustness in codon selection observed, there is some variability in codon usage for a number of amino acids. However, Ser is the only amino acid that differs between the set of

highly expressed genes and the set of genes used to compute "kazusa" codon usage table, which is the reference set for codon optimization tools. It will be interesting to test the effect of alternative codons for Ser in protein production. A second observation from our analysis is the putative presence of a stretch of low frequently used codons at the beginning of the ORF. It will be also interesting to test its influence in protein production. These factors can be easily tested using the platform that we have developed (see below). Codon usage in C. reinhardtii has attracted recent interest. In a paper from Ralph Bock lab (Barahimipour et al., 2015), the relative contribution of the GC content and codon usage in protein production has been studied. By expressing different YFP gene variants authors concluded that codon usage rather than GC content was a key determinant of translation efficiency. In contrast, GC content was important for gene expression at the level of chromatin structure. It is noteworthy that the YFP gene variants were fully optimized using kazusa codon usage frequency, and none of the factors mentioned before (alternative codons and ramp presence) were tested.

2. PLATFORM FOR TESTING EXPRESSION OF GENE VARIANTS In addition to the analysis of C. reinhardtii gene sequences, we also proposed to build a platform for testing the impact of codon usage in the production of a fluorescent protein reporter. Our experimental design, shown in Figure 3, consisted in the expression of different gene variants of a fluorescent protein (mVenus) that differed in the choice of synonymous codons. These gene variants are expressed using Chlamydomonas transcription regulatory sequences (promoters, terminators and UTRs). The construct also has a gene that confers resistance to an antibiotic, which allows the selection of cells that have incorporated the construct into their genome. The DNA cassette containing these genes can be integrated in any part of the genome, and therefore, its expression is strongly influenced by the transcriptional activity of the neighbouring region. To correct for this positional effect, we also included in the construct a second fluorescent marker (mRuby in the figure) that would allow to normalize the mVenus signal. The sequence of this second reporter is the same in all constructs. Chlamydomonas cell wall-less mutant CC-1883 was our initial choice as recipient strain since it has a high efficiency of transformation. To compensate for the variability between clones, we proposed to analyze a large pool of clones by FACS. The analysis of numerous clones is necessary because we estimate that less than 30% of clones will harbour the complete cassette. During the duration of this grant, we have been able to implement a versatile cloning strategy and test the expression of fluorescent reporters in different strain backgrounds. Although we have not been able to complete the experiment proposed yet, the progress that we have already made has proven to be useful for applications of transgene expression in Chlamydomonas.

2.1. Cloning strategy We decided to adopt a modular cloning strategy based on Weber et al. (Weber, Engler, Gruetzner, Werner, & Marillonnet, 2011). The strategy uses type IIS restriction endonucleases for the assembly of DNA sequences, and provides an inexpensive and fast method for multi-gene construction. Adoption of the syntax suggested by Patron et al (2015) (Patron, 2015) contributes to the standardization of the DNA parts constructed (Figure 4). This syntax consists in 12 fusion sites that enable the assembly and exchange of the different transcriptional units.

2.2. DNA parts A large fraction of the funding given to this project was used to synthesize DNA parts required for the constructs (Figure 5). These DNA parts were transcription regulatory sequences (promoters, UTRs and terminators) and also coding sequences corresponding to fluorescence reporters and drug resistant genes. These sequences were selected from a set of gene expression tools that have been extensively used for gene expression in Chlamydomonas. In all cases, DNA sequences were "domesticated" by removing the restriction enzyme sites that are required for the cloning strategy (BpiI and BsaI). Silent mutations were introduced when these sites fell on coding sequences. Domestication of HSP70-RBCS2 and PSAD promoters, RBCS2 5'UTR, and RBCS2 and PSAD 3'UTR-terminator sequences did not disrupt signals within for strong expression, as it is inferred from the high number of antibiotic resistant clones obtained when these sequences were used to drive the expression of a gene that confer resistance to the antibiotic paromomycin. The selection of fluorescent reporters was based on two recently published articles (Lauersen, Kruse, & Mussgnug, 2015; Rasala et al., 2013)). Rasala et al. 2013 shows that the autofluorescence of C. reinhardtii cells is low in channels 515/550, 554/590 and 575/608 nm (Ex/Em), and therefore, we chose mVenus/YFP, dTomato and mCherry fluorescent proteins. The mVenus gene described in Lauersen et al. (2015) gave a strong signal that could be detected in a fluorescent plate reader. This gene harbours an intron that authors suggest to boost expression. Regarding the second reporters, we have not been able so far to find a suitable fluorescent protein that could be used together with mVenus. We have expressed the red fluorescent protein mRuby2, also described by Lauersen, but we could not detect signal. We are currently testing a number of other red fluorescent proteins (mKate2, mPlum, mStrawberry and mCherry), some of them tested for first time in Chlamydomonas to our knowledge.

2.3. Chassis, transformation and fluorescence measurement The initial strain of choice to express the DNA constructs was the CC-1883. This is a cell wall-less mutant strain which can be easily transformed by the glass bead method (Neupert, Shao, Lu, & Bock, 2012), and it is routinely used in the Baulcombe lab. Transformants that integrated and expressed an antibiotic resistant gene were easily recovered in this background. However, expression levels of a fluorescent reporter were not sufficiently high and failed to be detected in a fluorescent plate reader. This low expression level of transgenes is a general feature of Chlamydomonas strains. Few years ago, Ralph Bock screened for mutant strains that efficiently expressed transgenes, isolating two different mutants, UVM11 and UVM4 (Neupert, Karcher, & Bock, 2009). We requested the UVM11 strain and confirmed that the mVenus fluorescent reporter was produced in this background to detectable levels. Therefore, we have switched to this strain in order to achieve high expression of our reporters. Although our initial plan was to use FACS for the quantification of fluorescence, we eventually decided to use fluorescence microplate readers instead. Plate readers provide a convenient method for quantifying the fluorescence of a large number of clones. We think that the sensitivity of the plate reader will be sufficient to detect differences in translation efficiency of our gene variants, provided a second fluorescent reporter is used for normalization. In the case that this method fails to detect these differences, an alternative method consisting in measuring fluorescence from native polyacrilamide gels can be

used (Barahimipour et al., 2015). This method was recently developed by Ralph Bock group to measure differences of translation efficiency of a YFP.

2.4. Conclusions of our "wet lab" work We have presented here the progress developing a platform that can be used to measure the impact of codon usage in the production of a fluorescence protein reporter. Although the initially proposed experiment has not been carried out yet, the tools that we have created and the technology that we have implemented are already useful to the Chlamydomonas community for testing gene expression. With the funding received in this project we have created the first Chlamydomonas-specific MoClo DNA parts following the consensus syntax for gene assembly in Plant Synthetic Biology. Some of these parts are already being used by other researchers of the Plant Sciences department of the University of Cambridge, who will also contribute with new parts to increase the repertoire of DNA parts available for Chlamydomonas research. We think that this will be a major contribution for the development of synthetic biology in Chlamydomonas. Unfortunately, we have not been able to use exclusively open technology in this project. One main component, the efficient transgene expressing strain UVM11, was obtained under MTA signing, and therefore, its distribution is restricted. We think that a potential follow-up of this project could be the creation of alternative open source technology.

3. LINKS TO AVAILABLE MATERIAL -RNA Seq data used to define highly expressed genes: https://www.ebi.ac.uk/arrayexpress/experiments/E-MTAB-3852/ -ESTs data used to define highly expressed genes: ftp://occams.dfci.harvard.edu/pub/bio/tgi/data/Chlamydomonas_reinhardtii/ -Script to compute number of codons in a set of genes: https://github.com/mariellev/OpenPlant -Chlamydomonas MoClo DNA parts: These parts will be deposited in the Chlamydomonas Resource Center (http://www.chlamycollection.org/) together with many other parts that are being constructed in the department of Plant Sciences of the University of Cambridge. Plasmid sequences will be available in this website too.

ATTACHED FILES Table 1. GC content of codon usage (data from Kazusa.org.jp) Table 2. Codon usage of gene datasets that differ in expression level, associated-GO term and pattern of expression. Figure 1. Distribution of coding sequences by their Codon Adaptation Index (CAI) using codon usage from the "All CDSs" set as reference. A, All CDS. B, Highly expressed gene set. C, Unmapped genes. Figure 2. Distribution of codon adaptiveness (wj) along coding sequences ("All CDSs" codon usage is used as reference). A, Example coding sequence. B, Heatmap of codon adaptiveness (wj) for the first 200 amino acids of sequences longer than 600 nts in the "Highly expressed genes" dataset. C, Average of codon adaptativeness using sequences in B. D, Average codon adaptativeness of sequences longer than 600 nts from the unmapped gene dataset. Figure 3. Experimental design for testing the impact of codon usage in the expression of a fluorescent protein reporter Figure 4. Gene syntax adopted for the assembly and exchange of DNA parts (modified from Patron et al. 2015). Figure 5. C. reinhardtii DNA parts constructed in this project.

ACKNOWLEDGMENTS We would like to thank David Baulcombe for helpful discussions and giving his consent to carry out part of this project in his lab; Jim Haseloff and members of his lab for suggestions, reagents and access to equipment; Ralph Bock for sharing Chlamydomonas reinhardtii strains; Matthew Hartley for his support throughout the project. This project was funded by the OpenPlant Fund (July 2015).

REFERENCES Angov, E. (2011). Codon usage: Nature's roadmap to expression and folding of proteins. Biotechnology Journal, 6(6), 650–659. doi:10.1002/biot.201000332 Barahimipour, R., Strenkert, D., Neupert, J., Schroda, M., Merchant, S. S., & Bock, R. (2015). Dissecting the contributions of GC content and codon usage to gene expression in the model alga, 1–14. doi:10.1111/tpj.13033 Cognat, V., Deragon, J., Vinogradova, E., Salinas, T., Remacle, C., & Marechal-Drouard. (2008). On the Evolution and Expression of Chlamydomonas reinhardtii Nucleus-Encoded Transfer RNA Genes ´. Genetics, 123(May), 113–123. doi:10.1534/genetics.107.085688 Gingold, H., & Pilpel, Y. (2011). Determinants of translation efficiency and accuracy. Molecular Systems Biology, 7(481), 481. doi:10.1038/msb.2011.14 Ikemura, T. (1982). Correlation between the abundance of yeast transfer RNAs and the occurrence of the respective codons in protein genes. Journal of Molecular Biology, 158(4), 573–597. doi:10.1016/0022-2836(82)90250-9 Ingolia, N. T., Ghaemmaghami, S., Newman, J. R. S., & Weissman, J. S. (2009). Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling. Science (New York, N.Y.), 324(April), 218–23. doi:10.1126/science.1168978 Lauersen, K. J., Kruse, O., & Mussgnug, J. H. (2015). Targeted expression of nuclear transgenes in Chlamydomonas reinhardtii with a versatile, modular vector toolkit. Applied Microbiology and Biotechnology. doi:10.1007/s00253-014-6354-7 Lopez, D. A., Hamaji, T., Kropat, J., De Hoff, P., Morselli, M., Rubbi, L., … Pellegrini, M. (2015). Dynamic changes in the transcriptome and methylome of Chlamydomonas reinhardtii throughout its life cycle. Plant Physiology, 169(October), pp.00861.2015. doi:10.1104/pp.15.00861 Nakamura, Y., Gojobori, T., & Ikemura, T. (1999). Codon usage tabulated from the international DNA sequence databases; its status 1999. Nucleic Acids Research, 27(1), 292. doi:10.1093/nar/27.1.292 Neupert, J., Karcher, D., & Bock, R. (2009). Generation of Chlamydomonas strains that efficiently express nuclear transgenes. Plant Journal, 57(6), 1140–1150. doi:10.1111/j.1365-313X.2008.03746.x Neupert, J., Shao, N., Lu, Y., & Bock, R. (2012). Genetic transformation of the model green alga Chlamydomonas reinhardtii. In Transgenic Plants: Methods and Protocols, Methods in Molecular Biology (Vol. 847, pp. 35–47). doi:10.1007/978-1-61779-558-9 Patron, N. J. . . . et al. (2015). Standards for plant synthetic biology : a common syntax for exchange of DNA parts. New Phytologist, 208, 13–19. Rasala, B. a., Barrera, D. J., Ng, J., Plucinak, T. M., Rosenberg, J. N., Weeks, D. P., … Mayfield, S. P. (2013). Expanding the spectral palette of fluorescent proteins for the green microalga Chlamydomonas reinhardtii. Plant Journal, 74(4), 545–556. doi:10.1111/tpj.12165 Schmollinger, S., Mühlhaus, T., Boyle, N. R., Blaby, I. K., Casero, D., Mettler, T., … Merchant, S. S. (2014). Nitrogen-Sparing Mechanisms in Chlamydomonas Affect the Transcriptome, the Proteome, and Photosynthetic Metabolism. The Plant Cell, 26(4), 1410–1435. doi:10.1105/tpc.113.122523 Sharp, P. M., & Li, W.-H. (1987). The codon adaptation index - a measure of directional synonymous codon usage bias, and its potential applications. Nucleic Acids Research, 15(3),

1281–1295. Tuller, T., Carmi, A., Vestsigian, K., Navon, S., Dorfan, Y., Zaborske, J., … Pilpel, Y. (2010). An Evolutionarily Conserved Mechanism for Controlling the Efficiency of Protein Translation. Cell, 141(2), 344–354. doi:10.1016/j.cell.2010.03.031 Valli, A. A., Santos, B. A. C. M., Hnatova, S., Bassett, A. R., Molnar, A., Chung, B. Y., & Baulcombe, D. C. (2016). Most microRNAs in the single-cell alga Chlamydomonas reinhardtii are produced by Dicer-like 3-mediated cleavage of introns and untranslated regions of coding RNAs, 1–11. doi:10.1101/gr.199703.115.Freely Weber, E., Engler, C., Gruetzner, R., Werner, S., & Marillonnet, S. (2011). A modular cloning system for standardized assembly of multigene constructs. PLoS ONE, 6(2). doi:10.1371/journal.pone.0016765 Zones, J. M., Blaby, I. K., Merchant, S. S., & Umen, J. G. (2015). High-Resolution Pro fi ling of a Synchronized Diurnal Transcriptome from Chlamydomonas reinhardtii Reveals Continuous Cell and Metabolic Differentiation, 1–28. doi:10.1105/tpc.15.00498

Expenditure:
EXPENDITURE BREAKDOWN
DNA part synthesis: IDTDNA technology: £2500
Sequencing vouchers (Source Bioscience): £495
Molecular biology and alga-specific reagents: £1005
Total: £4000

Are you claiming the additional £1000 follow-on funding?: Yes

Follow-On Plans:
FOLLOW-ON WORK We would like to apply for £1000 funding extension to accomplish the following objectives: -creation of a website dedicated to gene synthesis, with links to tools of codon optimization, resources and DNA synthesis companies, including user reviews, -extension of the repertoire of MoClo DNA parts, including new promoter sequences and other fluorescence reporters.