Kwansei Gakuin University
Report of Research Outcome

2020/02/04

To President

Department : Science and Technology
Position : Postdoctoral fellow
Name : Vandana Deopa

I report the outcome of the research as follows.

Name of the Fund/Program	□Sabbatical leave with grant □Sabbatical leave with no grant □KGU Joint Research □Individual Special Research ☒**Postdoctoral fellow** ※Please report by designated form as for "International Research Collaboration".
Research Theme	1.Functional characterization of putative chloroplast envelope located transporters belonging to the SLC-4 family by CRISPR/Cas9 nickase (D10A) mediated genome editing in the marine diatoms, *P. tricornutum* and *T. pseudonana*. 2. Study of the Pyshell proteins homologs in *T. pseudonana* and their functional characterization by CRISPR/Cas9 nickase (D10A) mediated genome editing.
Research Site/Venue	Matsuda lab, Building IV, Department of Bioscience, Kwansei Gakuin University, Kobe-Sanda Campus
Research period	2019/06/01 ～ 2020/01/31 (8 months)

◆ **Summary of the research outcome** (approx. 2,500 words)

Please write down the outcomes in detail regarding the research theme above.

Enhancing crop productivity by introduction of carbon concentrating mechanisms has been a highly sought-after area in the field of photosynthetic research. Among the diverse kinds of CCMs present in nature, the diatomic CCM has invoked greater interest because of its endosymbiotic origin which has enabled it to possess genes from algae, as well as animals and bacteria. Notably, the endosymbiotic event bestowed the diatoms with a four layered chloroplast envelope i.e. a double-layered chloroplast envelope that, in turn, is encircled by a double-layered chloroplastic endoplasmic reticulum. Although the information regarding the function of these envelopes is still scarce, it is well established that these play an important role in DIC trafficking. It has been ascertained that different set of transporters are present on these membranes which along with CAs in specific compartment function in conjunction to facilitate the uptake of HCO_3^- into the chloroplast stroma from the cytoplasm (Figure no.1).

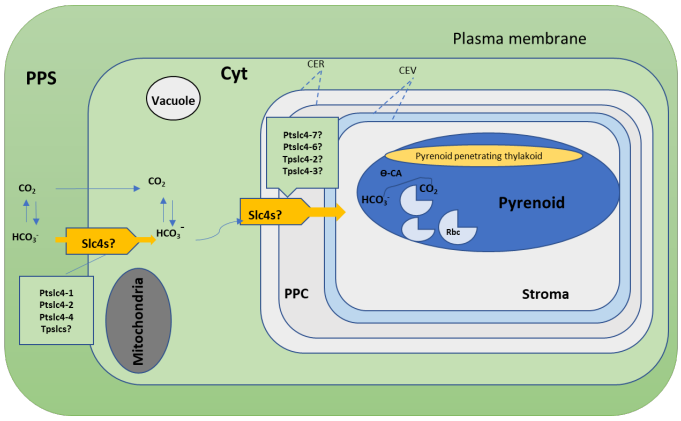

Figure 1. The CO_2 concentrating mechanism in diatoms

Nakajima et al. (2013) identified seven SLC4 genes and three SLC26 genes in the genome of *Phaeodactylum tricornutum*. Out of the seven SLC4 genes, PtSLC4-1, PtSLC4-2, and PtSLC4-4 have been already identified as the SLCs located on the plasma membrane which are involved in regulating the influx of HCO_3^- into the cell (Nakajima et al., 2013). The other two, PtSLC4-6 and PtSLC4-7 encode proteins predicted to localize in the four-layered chloroplast membrane systems as evident by the presence of targeting signals viz., GSA-FTS and SAA-FHT respectively. Similarly, in *Thalassiosira pseudonana*, TpSLC4-2 (SFS-FAP) and TpSLC4-3 (VNA-FPT) have been identified as candidate chloroplast membrane HCO_3^- transporters. The pre-sequences in diatoms significantly differ from those present in green algae or land plants as they possess a bipartite sequence comprising of a signal peptide sequence along with transit peptide.

Despite the advances in the field of diatom research in recent years, the function of the multi-layered chloroplast envelope membranes in diatoms is still largely unexplored at the molecular level and the candidate proteins involved in the entry of HCO_3^- into the stroma requires experimental validation.

Another area of recent interest is the recently discovered pyshell proteins. The confocal analysis of these proteins showed that these proteins were present on the surface of the pyrenoid in chloroplast (proposed by previous studies conducted in lab) of *P.tricornutum* and *T. pseudonana*. Although exact function of these proteins is not known any addition in this knowledge could be a stepping stone in proposition of carbon uptake model in these unicellular photosynthetic organisms.

Since studying the function of proteins through mutagenesis has been widely used these days, an attempt was made to knockout the candidate transporter proteins (in both *P.tricornutum* and *T. pseudonana*) and the pyshell proteins (in *T. pseudonana*) and then study their effects on the mutants. The genes chosen for the study have been described in the table below (Table no.1).

Serial no.	Gene name	Protein id in JGI	Genomic location
1.	Ptslc4-7	45656(Phatr2)	chr_7:834212-836081 (-)
2.	Ptslc4-6	43194(Phatr2)	Chr_1:2279932-2281914 (-)
3.	Tpslc4-2	260936 (Thaps3)	chr_1:1419165-1419927 (+)
4.	Tpslc4-3	267979 (Thaps3)	chr_1:974609-977268 (-)
5.	Tp7881(Tpphyshell1)	7881(Thaps3)	chr_9:155342-156226 (-)
6.	Tp23918(Tpphyshell2)	23918(Thaps3)	chr_9:157777-159128 (+)
7.	Tp7883(Tpphyshell3)	7883(Thaps3)	chr_9:159916-160818 (+)

Table no.1: Genes used for the functional characterization.

For this, CAS9 nickase approach was used (Figure no.2) and guide RNA pairs were shortlisted using the benchling software (https://www.benchling.com/crispr/).

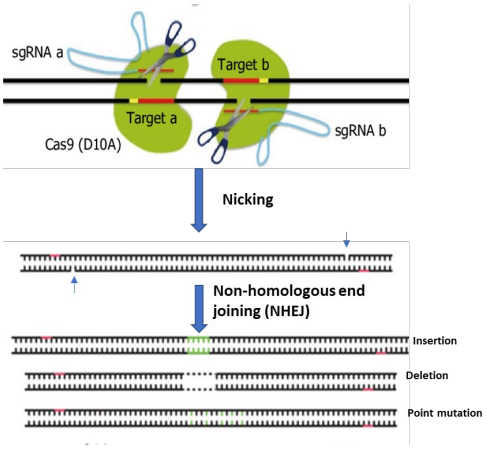

Figure 2: CRISPR/Cas9 nickase system

The guide RNA pair with the out of frame score more than 66 (calculated using the cas-offinder tool (http://www.rgenome.net/cas-offinder/) and those having the least off-target activity were finally used for synthesis of constructs. The primer pair used for the amplification of gRNA for each gene have been depicted in the table below (Table no.2).

S. No.	Primer name	Sequence used (5'to 3') for the initial amplification of gRNA
1.	Fptslc7	AGGTCTCATCGAGGTTGATCCGATGATTGTCAAGTTTAAGAGCTAGAAAT
2.	Rptslc7	GGGTCTCTAAACATTGTGCGAGACAGGCTACGCTCGACTTTGAAGGTGTT
3.	Fptslc6	AGGTCTCATCGAGACTGACCGCCGGAGAAAGACGTTTAAGAGCTAGAAAT
4.	Rptslc6	GGGTCTCTAAACTGCGTTATCTCCGAAGCAATCTCGACTTTGAAGGTGTT
5.	Ftpslc3	AGGTCTCATTGTGAAACCCACCCACGCATATACGTTTAAGAGCTAGAAAT
6.	Rtpslc3	GGGTCTCTAAACAGAATCATCATAAAGGCACTCACAATTTCGGCAAAACG
7.	Ftpslc2	AGGTCTCATTGTGCATGGTACCGATCGCCCCATGTTTAAGAGCTAGAAAT
8.	Rtpslc2	GGGTCTCTAAACCACATTGCTGTAGAACTGACCACAATTTCGGCAAAACG
9.	Ftppyshell1	AGGTCTCATTGTGAGTTGTCCCAGATCTCGGCAGTTTAAGAGCTAGAAAT
10.	Rtppyshell1	GGGTCTCTAAACCCCTGCACAATCACGGGGGACACAATTTCGGCAAAACG
11.	Ftppyshell2	AGGTCTCATTGTGTAACGGCATTGAAGGTACGAGTTTAAGAGCTAGAAAT
12.	Rtppyshell2	GGGTCTCTAAACCGGTGTTGGGGCCGCGGGGACACAATTTCGGCAAAACG
13.	Ftppyshell3	AGGTCTCATTGTGATCAATGGATGGGTTGGCGAGTTTAAGAGCTAGAAAT
14.	Rtppyshell3	GGGTCTCTAAACTCTTGAGCAACACTTGCACTCACAATTTCGGCAAAACG

Table no. 2: The primer sequences used for gRNA amplification from the template vector.

The gRNAs were cloned in the insertion vector using BsaI restriction enzyme sites. The methodology followed for cloning and the vector used have been illustrated in the figure below (Figure3A and 3B).

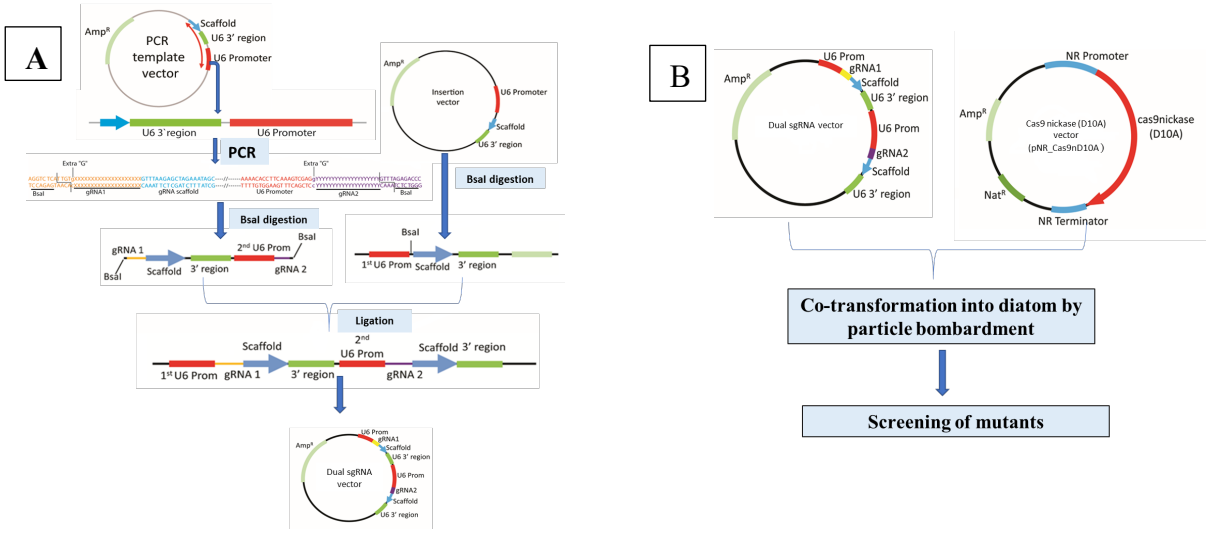

Figure 3(A): Strategy for construction of recombinant vector with dual sgRNAs. Step1: Amplification of the dual sgRNA cassette from the PCR template vector by PCR with specific gRNAs with BsaI restriction site. Step2: Digestion of the insertion vector and PCR amplified dual sgRNA cassette by BsaI. Step3: Ligation of the insertion vector and dual sgRNA cassette.

Figure 3(B): A schematic depiction of two vector expression system of dual sgRNAs and Cas9 nickase (D10A). A cas9 nickase D10A contains a point mutation that converts a GAC codon to GCC in the amino acid sequence coding for the *diacas9* gene. The fragment encoding Cas9 nickase (D10A) was cloned into the pNR vector to produce pNR_Cas9nD10A vector. The dual sgRNA vector is the expression vector of the two sgRNA driven by the U6 promoter.

The clones were confirmed by sequencing. The vector harboring the gRNA pair along with the nickase vector were co-transformed into the *T. pseudonana* cells ($O.D._{730}$ - 0.1 to 0.2) by particle bombardment method. The cells were then screened on plates containing F/2 medium (containing NH_4^+ as nitrogen source) with nourseothricin. The colonies that appeared on the plates were further streaked on F/2 medium (containing NO_3^- as nitrogen source).

The total number of colonies obtained were 38 (for *T.pseudonana*). No colonies were observed after transformation of *P. tricornutum*, which reflects failure of the transformation process or the lethality of the knock out mutant. The maximum number of colonies were observed in Tpphyshell2 (16 colonies) followed by Tppyshell3 (12 colonies) and Tpslc4-2 and Tppyshell1 (5 colonies each). No colonies were observed in Tpslc4-3 transformed plates.

An active transport of carbon from the cytoplasm into the chloroplast is the main driver of the diatom CCM. Therefore, the identification of the chloroplast located bicarbonate transporters is crucial step for the understanding of a diatomic CCM. The PtSLC4-6 and PtSLC4-7 genes have been reported to be constitutively expressed under high and low CO_2 conditions suggesting that these transporters regulate DIC flow from the cytosol to the plastid regardless of ambient CO_2 concentrations (Nakajima et al., 2013). If this is the case, then a knock out mutant of these transporters may be lethal. And if the activity of either of them is perturbed, the organism may still survive as the other may carry out the transport function. It is also expected that the loss of the chloroplast pump may result in higher cytoplasmic DIC which may increase the CO_2 leakage.
The effect of mutagenesis on the pyshell proteins is an unknown facet of CCM research. The role of these protein in carbon sequestration and the reason for the existence of the protein in close vicinity of other CCM components in the pyrenoids is still a fascinating question.

Deadline : Within two months after finishing the research period.

Sabbatical leave with grant: Submit this report to President with confirmation by the dean of school you belong to.

※Postdoctoral fellow is required to submit this report with confirmation by the dean of graduate school before the end of employment period.

Where to submit : Organization for Research and Development and Outreach (NUC)

◆We put this report on the web of KGU. If there is any problem about it because of difficulties on your research, please let us know.