

Title: Additional conjugation systems for inter-domain plasmid transfer to the diatom *Phaeodactylum tricornutum*

Authors:

Anthony K. Kang, Vincent A. Bielinski, Tayah M. Bolt, Christopher L. Dupont[#], Philip D. Weyman

Affiliations:

J. Craig Venter Institute, 4120 Capricorn Lane, La Jolla, CA 92037.

Correspondence:

#: Christopher L. Dupont, cdupont@jcvj.org, 858-200-1886

Keywords:

Conjugation, diatoms, *Phaeodactylum tricornutum*, RSF1010, *Agrobacterium tumefaciens*, Ti plasmid

Abstract:

Bacterial conjugation utilizes a type IV secretion system and a DNA transfer mechanism to deliver DNA from one cell to another. Conjugative partners are conventionally confined to the prokaryotic domain. In a prominent exception, *Agrobacterium tumefaciens* type IV secretion-mediated transfer of DNA to plant cells can result in subsequent chromosomal integration. Recently, we demonstrated interdomain conjugation from *Escherichia coli* to the diatom *Phaeodactylum tricornutum* with the subsequent maintenance of an episome at chromosomal copy numbers if it contains diatom centromeres or centromere-like elements. The genes involved in the conjugation process can be separated into those encoding the type IV secretion system, also called the mating pair formation (MPF) genes, and genes involved in DNA processing called the mobilization (MOB) genes. Various protein families compose each class of conjugation genes, including common MOB types F, P, and Q and MPF types F, P, and T. The conjugative transfer from *E. coli* to *P. tricornutum* was demonstrated with a vector expressing MOB_P and MTF_P. Here we show that the MOB_P system can be deleted and complemented with a MOB_Q system in *E. coli*-diatom conjugations with subsequent episomal maintenance. Utilization of both MOB_P and MOB_Q systems results in substantially higher efficiencies in *E. coli*-diatom conjugation. Finally, we demonstrate conjugative gene transfer between *P. tricornutum* and *A. tumefaciens* expressing a MPF_T, the first demonstration of this system in diatoms, resulting in episomal maintenance or chromosomal integration, depending on the ex-conjugant. The promiscuity of MOB and MTF systems permitting prokaryote to diatom conjugative DNA transfer suggest major environmental and evolutionary importance of this process. The increased efficiency of dual MOB systems immediately improves genetic engineering in diatoms and has interesting basic cellular biology implications.

Introduction:

Diatoms are eukaryotic photosynthetic microorganisms with great ecological and biotechnological importance. Better understanding of how they perform vital ecosystem functions and how they can be further improved to become economically viable bioproduction platforms will require extensive genetic interrogation. In the last three years, the genetic toolset of diatoms has been greatly improved through the development of targeted gene editing (1-3), artificial chromosomes (4), and improved methods for DNA introduction based on conjugation from *E. coli* (5, 6). In addition to accelerating genetic manipulations, this interdomain conjugation may have important evolutionary implications.

While many examples of horizontal gene transfer from bacteria to eukaryotes exist (7, 8), the mechanisms facilitating these transfers often remain unknown. While conjugation has typically thought to be reserved for intraprokaryotic gene transfers (i.e. 'bacterial mating'), numerous recent studies describe prokaryotes and conjugation systems capable of transferring DNA to a variety of eukaryotes (9). Conjugative systems consist of the mating-pair formation (MPF) genes to establish the cell-to-cell connection, and a set of genes to process the transferred DNA, called the DNA transfer and replication (DTR) or mobilization (MOB) genes (10). The MPF consists of a type IV secretion system and the DTR features a single stranded nuclease called a relaxase that nicks the DNA at a specific sequence called the origin of transfer (OriT) and becomes covalently linked to the 5' end. The relaxases have been classified into various mobilization groups (of which F, P, Q are frequently encountered examples)

according to their phylogenetic relationships (11) and similar approaches for the MPF genes have also allowed phylogenetic clustering (10).

At least three conjugation systems have been used for interdomain conjugation.

Transfers from *E. coli* to yeast (*Saccharomyces cerevisiae*) have been successfully accomplished with the F-plasmid (MOB_F, MPF_F) and IncP-based RP4 systems (MOB_P, MPF_P) (12, 13). Transfers to yeast using the *Agrobacterium tumefaciens* Ti-plasmid (MOB_Q, MPF_T) have been reported (14). *A. tumefaciens* is perhaps the most well-known mediator of interdomain DNA transfers, transferring pathogenesis-related genes to a wide variety of plants during the course of various plant diseases such as crown gall. Interdomain conjugation has also been reported for mammalian and fungal cells using RP4-based systems (Sikorski et al., 1990; Waters, 2001; Schröder et al., 2011)(15-17).

Plasmids such as F, RP4/RK2, and Ti contain both the MPF and the DTR functions and are self-transmissible or conjugative (Smillie et al., 2010)(10). Other plasmids such as RSF1010 contain only the MOB or DTR and are known as mobilizable. RSF1010 is known to have several features that make it broad host range among bacteria (Meyer, 2009; Jain & Srivastava, 2013) and capable of being transferred into plant cells from *A. tumefaciens* (Buchanan-Wollaston, Passiatore & Cannon, 1987). Among these features that allow for its extraordinary host range is a relaxase (MobA) that contains a dedicated DNA primase for synthesis of the second strand after transfer (Meyer, 2009). This feature may improve interdomain conjugative transfer, but direct comparison of plasmid RSF1010 and RP4-based systems has never been reported.

Currently, conjugation to diatoms is achieved using *E. coli* and derivatives of conjugative plasmid RP4 with a type IV secretion system (MPF_P) combined with a MOB_P relaxase(5, 6). It remains an open question as to whether this plasmid transfer system is simply fortuitous or whether additional conjugation systems and different bacterial donors can transfer DNA to diatoms. The possibility that multiple conjugation systems could transfer DNA to diatoms would justify a more systematic investigation into the ecological and evolutionary importance of bacterial-diatom gene transfer. Such studies may also identify systems with efficiency improvements over the current RP4 and *E. coli* system. Improving conjugation efficiencies to $\sim 10^{-2}$ from the current 4×10^{-4} (6) would have great biotechnological implications allowing for practical screening of very large libraries of expression variants, mutants, and would permit other combinatorial strategies.

As a first step toward investigating whether other conjugation systems could deliver DNA to diatoms, we tested an alternative relaxase (MOB_Q-based) and an alternative type IV secretion system (*A. tumefaciens*-based). A plasmid encoding a MOB_Q relaxase was able to complement a relaxase mutant in the RP4/RK2 plasmid to successfully transfer DNA to diatoms. We also present data demonstrating that the MPF_T-based system of *A. tumefaciens* was capable of transferring plasmids to diatoms.

Materials and Methods:

Microbial Strains and growth conditions:

For all experiments, *Escherichia coli* strain Epi300 and *Phaeodactylum tricornutum* strain CCMP 632 were used. *E. coli* was cultured at 37 °C in lysogeny broth (LB) in liquid form or supplemented with 15 g L⁻¹ agar. *P. tricornutum* was grown in L1 medium (18) at 20 °C under

cool white fluorescent lights ($50 \mu\text{E m}^{-2} \text{s}^{-1}$). For *E. coli*, antibiotics were used at the following concentrations: ampicillin (Ap, $100 \mu\text{g mL}^{-1}$), tetracycline (Tc, $10 \mu\text{g mL}^{-1}$), chloramphenicol (Cm, $25 \mu\text{g mL}^{-1}$), spectinomycin (Sp, $40 \mu\text{g mL}^{-1}$), gentamicin (Gm, $20 \mu\text{g mL}^{-1}$), kanamycin (Km, $25 \mu\text{g mL}^{-1}$). *P. tricornutum* lines were selected and maintained on phleomycin $20 \mu\text{g mL}^{-1}$.

95 *Agrobacterium tumefaciens* strain LBA4404 was cultured in AB* media (19) and grown on plates or in broth at 28°C . Antibiotics for selection were used as follows: kanamycin ($100 \mu\text{g mL}^{-1}$), and cefotaxime ($100 \mu\text{g mL}^{-1}$) to eliminate *A. tumefaciens* culture from the diatom ex-conjugant colonies.

100 *Plasmid construction and conjugation experiments:*

Cargo plasmids were designed and assembled to assess the viability of IncQ-based conjugation in *P. tricornutum*. The pTA-MOB transfer plasmid (20) maintained in *E. coli* strain Epi300 was modified to insertionally knock out the relaxase-encoding *tral* gene resulting in plasmid pTA-MOB-*tral::kan_r*. The gene *tral* was knocked out using Lambda Red (21),
105 subsequently disrupting the relaxosomal TraH subdomain of Tral. To do this, the kanamycin resistance gene from plasmid pACYC177 was amplified using primers Tral-KO-2 and Tral-KO-3. These primers contained 30-40 bp of sequence homology to the *tral* region. After introducing the Lambda Red plasmid pKD46 into the *E. coli* strain containing pTA-MOB, the cells were electroportated with the PCR product encoding the kanamycin resistance cassette.

110 Because pTA-MOB is maintained as a low but multiple copy plasmid in *E. coli*, the resulting Tral knockouts contained both mutant and wild-type forms of pTA-MOB in the cells following Lambda Red. To accommodate for this, pTA-MOB plasmids were extracted via miniprep and transformed back into *E. coli* Epi300 to select for the segregated pTA-MOB-*tral::kan_r* transformants.

115 Cargo plasmid pPt-Cargo1383a was constructed using the pRL1383a (22) vector derived from plasmid RSF1010 (23) that included the IncQ-MobA relaxase and transfer origin (*oriT*) to complement the knocked out IncP-Tral relaxase on pTA-MOB. A CenArsHis (containing the yeast-derived CenArsHis region that allows for centromeric replication in *P. tricornutum* (4, 6) and the ShBle cassette that permits growth in the presence of phleomycin
120 (24) were added to allow for episomal maintenance and selection of the cargo plasmid in *P. tricornutum*. To make plasmid pPt-Cargo1383a, the CenArsHis-ShBle cassette was amplified using primers 1383a-insert-1 and 1383a-insert-2 (Table 1) from template plasmid pPtPBR1. Following PCR, the DNA fragments were purified by PCR cleanup kit (QIAquick, Qiagen). Plasmid pRL1383 was digested with restriction enzymes EcoRI and SacI and purified by
125 ethanol precipitation. The linearized pRL1383 backbone and purified PCR products were assembled using Gibson assembly and transformed into *E. coli* strain Epi300. The construction of plasmid pPt-Cargo1383a was validated with colony PCR, restriction digest analysis, and Sanger DNA sequencing.

130 The functionality of the IncQ-based complementation of the *tral::kan_r* in *E. coli* was verified by a preliminary *E. coli*-to-*E. coli* conjugation experiment. Four donor *E. coli* strains with different plasmid pairs were prepared for conjugation: pTA-MOB-*tral::kan_r* + pPtPBR1 (negative control), pTA-MOB + pPtPBR1 (positive control), pTA-MOB-*tral::kan_r* + pPt-Cargo1383a (experimental pairing), and pTA-MOB + pPt-Cargo1383a (experimental pairing). Each of the donor strains was paired with a recipient *E. coli* strain containing a chloramphenicol resistance

marker on a non-mobilizable plasmid to test whether the cargo plasmid could be transferred to the recipient. To prepare cells for conjugation, 50 mL cultures of donor and recipient cells were grown with shaking (200 rpm) at 37 °C in appropriate antibiotic media before being spun down; pellets were washed several times with LB to minimize antibiotic residues in the pellets before being resuspended in 500 µL LB. For the conjugation, 50 µL of donor and 50 µL of recipient *E. coli* were mixed and co-cultured on LB agar plates for 1 hour at 37°C. Conjugated cells were then resuspended from plates into liquid LB medium, and serial dilutions of the cells were re-plated on 1) LB-Cm to count the total number recipient cells, 2) LB-SpCm to provide counts for conjugated recipients for pPt-Cargo1383a (SpR), and LB-TcCm to provide counts for conjugated recipients for pPTPBR1 (TetR). CFU counts were subsequently obtained following overnight incubation at 37°C.

Conjugation between *E. coli* donor cells and *P. tricornutum* recipients was then tested using the same four sets of donor cells with the protocol that was previously developed for IncP-based conjugation between *E. coli* and *P. tricornutum* (6). *P. tricornutum* colonies were patched onto L1-Phelo20 plates and were tested for the presence of bacteria by streaking on LB medium. Once verified to be free of donor bacterium, cargo plasmids were extracted from the patched cultures using a previously described protocol (6) and transformed into *E. coli* by electroporation. Several *E. coli* colonies from each original *P. tricornutum* exconjugant were grown for plasmid extraction. Extracted plasmids were then characterized with EcoRV and BamHI restriction digest analyses and compared to similarly digested controls.

To test the viability of Ti plasmid-based DNA conjugation into *P. tricornutum*, two plasmids were constructed: pPT-Agro_LB and pPT-Agro_No_LB. These plasmids were based on plant binary vector pBI121 that was augmented with the yeast CenArsHis cassette to provide centromeric maintenance in *P. tricornutum*, the ShBle cassette to provide phleomycin resistance in *P. tricornutum*, and a spectinomycin resistance cassette to aid in assembly. First, plasmid pPT-Agro was constructed by linearizing plasmid pBI121 with restriction enzyme HindIII and using Gibson assembly to insert the *P. tricornutum* ShBle cassette amplified using primers Pt-Agro-1 and Pt-Agro-2 (Table 1). To make plasmids pPT-Agro_LB and pPT-Agro_No_LB, a cassette containing the CenArsHis, ShBle, and spectinomycin resistance (SpR) cassettes was amplified with primers Pt-Agro-7nF + Pt-Agro-6nR for pPT-Agro_No_LB (to make fragment P3N) and PT-Agro-5F and Pt-Agro-6R for pPT-Agro_LB (to make fragment P7). For plasmid pPT-Agro_LB, the pPT-Agro plasmid was amplified as three fragments (omitting the previously added ShBle cassette in pPT-Agro). In the first fragment, primers PT-Agro-4F and Pt-Agro-3R were used to amplify fragment P3. In the second fragment, primers PT-Agro-3F and Pt-Agro-5R were used to amplify fragment P5. A four piece Gibson Assembly was used to construct pPT-Agro_LB (P7 + P9 + P5 + P3). The integrity of constructed pPT-Agro vectors were subsequently validated with colony PCR (primers Pt-Agro-6F and Pt-Agro-4R, Table 1) and restriction digest analysis.

To make plasmid pPT-Agro_No_LB, pPT-Agro was amplified as three fragments and the left border sequence, CaMV promoter, and *gus* sequences of pPT-Agro were deleted as well as the previously added ShBle cassette. The first was performed using primers Pt-Agro-6F and Pt-Agro-4R to make fragment P2N. The second fragment was amplified using primers Pt-Agro-4F and Pt-Agro-7nR to make fragment P1N. A three piece Gibson Assembly was used to construct pPT-Agro_No_LB (P1N + P2N + P3N). The integrity of constructed pPT-Agro vectors were

subsequently validated with colony PCR (primers Pt-AgroCheck1 and SeqTest4-R, Table 1) and restriction digest analysis.

Plasmids were transformed into *A. tumefaciens* via electroporation. To prepare *A. tumefaciens* donor strains for conjugation with *P. tricornutum*, transformed cells were incubated in 50 mL IM liquid media with acetosyringone ($200 \mu\text{g mL}^{-1}$, hereafter noted as AS-200) and Kanamycin $50 \mu\text{g mL}^{-1}$ overnight at 28°C shaking to $\text{OD}_{600}=0.6$ (14). In an alternative protocol, transformed *A. tumefaciens* was grown overnight in medium AB*i (19) containing Kanamycin $50 \mu\text{g mL}^{-1}$ and AS-200 to a similar density. Cells were washed by spinning down cultures and resuspending pellets in $500 \mu\text{L}$ IM + AS-200 media prior to conjugation; $200 \mu\text{L}$ of *A. tumefaciens* was mixed and co-cultured with $200 \mu\text{L}$ of prepared *P. tricornutum* recipients (5×10^8 cells mL^{-1} concentration) on plates that consisted of L1 : AB*i (AB* media with AS-200): 3% agar in a 2:1:1 ratio. Cocultures were incubated for 90 minutes at 28°C , then at 20°C for 2-3 days, before being resuspended in $800 \mu\text{L}$ of L1 and replated on L1 plates containing Phleo20 supplemented with Cefotaxime ($100 \mu\text{g mL}^{-1}$). Resulting *P. tricornutum* ex-conjugants were counted following a 7-8 day incubation at 20°C .

P. tricornutum colonies were then patched and tested for the presence of bacteria by plating on LB agar. Once the absence of bacteria from the *P. tricornutum* cultures was verified, plasmids from the pPT-Agro_LB and pPT-Agro_No_LB ex-conjugant patches were extracted, transformed into *E. coli* strain Epi300 by electroporation, and several *E. coli* colonies from each *P. tricornutum* colony were grown for plasmid extraction and restriction digest analysis with EcoRV.

Results:

Comparison between MOB_Q and MOB_P interdomain conjugation

We developed a dual plasmid system to test whether the plasmid RSF1010 was able to be mobilized to the diatom nucleus by conjugation. We first constructed an RSF1010-based cargo plasmid, pPt-Cargo1383a, that contained a bleomycin selectable marker for *P. tricornutum* (ShBle) and a centromeric maintenance region (6, 24). Mobilization of RSF1010 requires the MPF functions of a conjugative plasmid such as RP4/RK2 plasmid. We used a version of the RP4/RK2 plasmid called pTA-MOB that contained the Tra1 and Tra2 islands from RP4/RK2 but lacked its own oriT making it incapable of self-transfer (Strand et al., 2014). The Tra1 region of the RP4/RK2 conjugative plasmid contains the TraI relaxase for DNA mobilization. As such, testing the compatibility of the IncP MPF with mobilizable IncQ plasmids necessitates the knockout of the *traI* relaxase gene on the plasmid pTA-MOB. We constructed a *traI* knockout in plasmid pTA-MOB using Lambda Red recombineering (21). The TraH gene encodes a cytoplasmic relaxosomal domain that, in F conjugation, is localized to the outer membrane during conjugation (25, 26). TraH is maintained as a subdomain of TraI, and thus our knockout of TraI included the initiator methionine of TraH to avoid relaxosomal components of IncP from interfering with the function of the IncQ relaxase in this study.

We paired the cargo plasmid pPt-Cargo1383a (containing MOB_Q and OriT $_{\text{MOBQ}}$) with both the original pTA-MOB and our version with a insertionally inactivated *traI*. As a control, we also paired cargo plasmid pPtPBR1, containing only the OriT $_{\text{MOBP}}$ (6), with both pTA-MOB and pTA-MOB-*traI::kan_r*. As shown in Figure 1, we hypothesized that pPtPBR1 would transfer when paired with the original version of pTA-MOB, but pTA-MOB-*traI::kan_r* would prevent conjugative

transfer of pPtPBR1. Furthermore, because the RSF1010-based plasmid, pPt-Cargo1383a, encoded its own relaxase, it should transfer using both the original and the relaxase mutant versions of pTA-MOB. The pPtPBR1 control also allowed us to compare the efficiency of MOB_Q-based conjugation with that of the currently-used MOB_P-based conjugation.

We verified that the our proposed experimental system was functioning properly by first performing *E. coli-E. coli* conjugation experiments. As expected, the pPt-Cargo1383a plasmid transferred to an *E. coli* recipient using a donor containing either pTA-MOB and pTA-MOB-*tral::kan_r*, but the pPtPBR1 plasmid only transferred using the original pTA-MOB (data not shown).

We then set up analogous conjugation experiments with the diatom *P. tricornutum* (Fig. 1C). The efficiency of conjugation was quantified by counting CFUs of transconjugant *P. tricornutum*, from which the cargo plasmids were extracted and characterized by restriction digest analysis. We observed that the RSF1010-based pPt-Cargo1383a was able to transfer to *P. tricornutum* in both the original and *tral* mutant versions of pTA-MOB consistent with the MOB_Q relaxase successfully transferring the DNA to the eukaryotic nucleus (Fig. 1D). The control pPtPBR1 plasmid performed as expected yielding colonies with the original pTA-MOB but not pTA-MOB-*tral::kan_r*. The increase in conjugation efficiency of having both relaxases present in the pPt-Cargo1383a/pTA-MOB pairing was consistently observed to be between 4-6 fold over multiple repeated experiments (Fig. 1D).

Patches were tested by streaking on LB agar to verify that no donor *E. coli* was present in the diatom culture. To determine whether the transferred DNA was maintained as an episome in the recipient diatom, we performed plasmid purification on the diatom exconjugants. Isolated plasmids were transformed into *E. coli* to 1) verify that the plasmids were in closed circular form when isolated from the diatom and 2) obtain enough plasmid to perform restriction enzyme digest. All plasmids tested had the expected band pattern suggesting that plasmids delivered to the diatom nucleus by the RSF1010-based transfer system were maintained episomally (Fig. 1E).

A. tumefaciens-mediated transfer of episomes to *P. tricornutum*

We also tested whether plasmids could be delivered by *A. tumefaciens* to the *P. tricornutum* and if they were maintained as episomes after transfer. In plants, *A. tumefaciens* is primarily used as a vehicle for insertion of DNA into the native chromosomes, yet Bundock et al. (14) observed that the majority of DNA transferred to yeast by *A. tumefaciens* was maintained episomally when the 2µm replication origin was included on the shuttle plasmid. We created two vectors to test transfer and maintenance. The first was a simple modification of the binary vector pBI121 to add a *P. tricornutum* selectable marker (ShBle) and the centromeric replication sequence (CenArsHis) to create plasmid pPt-Agro_LB (Fig. 2A). In the second plasmid, we deleted the Left Border (LB) sequence along with much of the beta-glucuronidase gene found on plasmid pBI121 to create plasmid pPt-Agro_No_LB (Fig. 2D). The rationale for this was that the linear transfer of the T-DNA region is usually achieved by nicking at both the right and left border sequences; by removing the left border nick sites, we reasoned that the entire plasmid would more likely be transferred.

We were successful in obtaining *P. tricornutum* colonies from our conjugations with *A. tumefaciens* with both cargo plasmid versions. Two pPt-Agro_LB and three pPt-Agro_No_LB *P.*

tricornutum colonies were recovered following conjugation between *A. tumefaciens* and *P. tricornutum*; plasmid purification was performed on two of each of the pPt-Agro_LB and pPt-Agro_No_LB colonies, which were then transformed into *E. coli* on separate plates. The ability to isolate episomes from the *P. tricornutum* ex-conjugants that can be subsequently transformed into *A. tumefaciens* indicates that the plasmid was maintained as an episome in the diatom rather than inserted into the *P. tricornutum* chromosome. Both pPt-Agro_No_LB plates yielded *E. coli* colonies, while only one of the two pPt-Agro_LB plates yielded colonies. We digested plasmids extracted from four *E. coli* colonies from the one pPt-Agro_LB ex-conjugant that was transformable into *P. tricornutum* with the restriction enzyme EcoRV to gather insight on the sizes of the recircularized plasmids post-conjugation. Plasmids from three of the four *E. coli* colonies had the correct restriction band pattern suggesting that in general, the plasmids were transferred and maintained without alteration in *P. tricornutum* (Fig. 2B). Restriction enzyme digest of plasmids isolated from two *E. coli* colonies from each of the three pPt-Agro_No_LB colonies that yielded plasmids resulted in only one of the six plasmids having the correct band pattern. This suggests that transfer and maintenance of the plasmids lacking the left border has lower fidelity.

In a second experiment, eight pPt-Agro_LB and five pPt-Agro_No_LB *P. tricornutum* colonies were obtained following conjugation from *A. tumefaciens*. Of these colonies, we performed plasmid purification and transformation into *E. coli* on five pPt-Agro_LB and three pPt-Agro_No_LB *P. tricornutum* patched colonies; only one ex-conjugant colony of each plasmid group yielded *E. coli* colonies.

Discussion:

We found that additional conjugation systems based on RSF1010 (MOB_Q) and *A. tumefaciens* (MPF_T) could transfer DNA to *P. tricornutum*. While RSF1010-based plasmids have been used previously to transfer to plant cells, they were used with an *A. tumefaciens* system (27). To our knowledge, this is the first time MOB_Q and MPF_P have been used with *E. coli* to perform interdomain conjugative DNA transfer.

We hypothesized that plasmid pPtCargo1383a, based on RSF1010, would have an increased conjugation efficiency because the relaxase has an extra primase domain that is not found on the RP4/RK2 relaxase used previously for interdomain plasmid transfer (28). If synthesis of the second strand in the nucleus is a limiting factor for successful establishment of the plasmid, then the activity of the primase on the protein-DNA complex that is transported to the diatom nucleus may improve the efficiency of episome establishment and therefore improve overall conjugation efficiency as measured by the number of resulting colonies. However, we observed no increase in conjugation efficiency with the RSF1010 relaxase. Thus, the inclusion of the primase domain on the relaxase was not sufficient to improve episome establishment, and may require the coordinated activities of other cellular enzymes including DNA polymerase for any beneficial effects to be measured.

We consistently observed an unexpected improvement in DNA transfer when both RSF1010 MOB_Q and RP4/RK2 MOB_P relaxases were present on separate vectors. This may be a result of unexpected effects of the *tral* relaxase mutant on other genes of RP4/RK2 involved in conjugation. Alternatively, the improvement may suggest that the activity of the relaxase is limiting either in nicking the DNA in the donor bacterium or successfully transferring to the

diatom nucleus; simply having more activity through the combination of the two enzymes may improve the overall rate of successful DNA transfer.

We also observed that *A. tumefaciens* could perform interdomain conjugation with *P. tricornutum* resulting in establishment of episomes in the recipient when the cargo plasmid contained a centromeric maintenance sequence. Agrobacterium has been used extensively to transfer DNA to eukaryotic algae in the green algal lineage with subsequent chromosomal integration(29-32). In contrast, to our knowledge, the only other example beyond our results of using *A. tumefaciens* to deliver DNA maintained as an episome in the recipient was with yeast(14).

During conjugation from *E. coli* to diatoms, it remains unclear how the DNA-protein complex is transported to the recipient nucleus once in the eukaryotic cell. In the case of *A. tumefaciens*, the mechanisms have been well established and nuclear localization is known for the VirD2 and VirE2 proteins of *A. tumefaciens* Ti-plasmid (33-35). Neither the RSF1010 MOB_Q relaxase (MobA) nor the RP4/RK2 MOB_P relaxase (Tral) are known to contain predicted nuclear localization signals. High scoring NSL signals are also absent from the complete set of predicted proteins from the RSF1010 and RP4/RK2 plasmids (Table 2). The relaxases from RSF1010 and RP4/RK2 had predicted NLS scores of 0.48 and 0.60, respectively. With this score, there is ~70% chance that the proteins are nuclear localized. In contrast, the VirD2 relaxase from the *A. tumefaciens* Ti plasmid has a score of 0.93, consistent with its known localization to the nucleus. Brameier et al. suggest that NucPred scores of over 0.80 correctly predict nuclear localization 93% of the time. The intermediate scores of RSF1010 and RP4/RK2 leave open the question of whether they have some intrinsic nuclear localization feature.

Conclusions:

We have observed interdomain conjugation between bacteria and the diatom *P. tricornutum* for two additional conjugation systems. First, the plasmid RSF1010 is able to be mobilized from *E. coli* using the RP4/RK2 type IV secretion system. Second, a cargo plasmid based on pBI121 was able to be mobilized from *A. tumefaciens* using the Ti plasmid. In both cases, episomes could be isolated from the *P. tricornutum* recipient. This finding broadens the scope of interdomain conjugation, suggesting that it may be occurring in a more diverse prokaryote-eukaryote interactions with possible evolutionary implications. Further, it highlights a compatibility between the different MPF and DTR systems.

Figures:

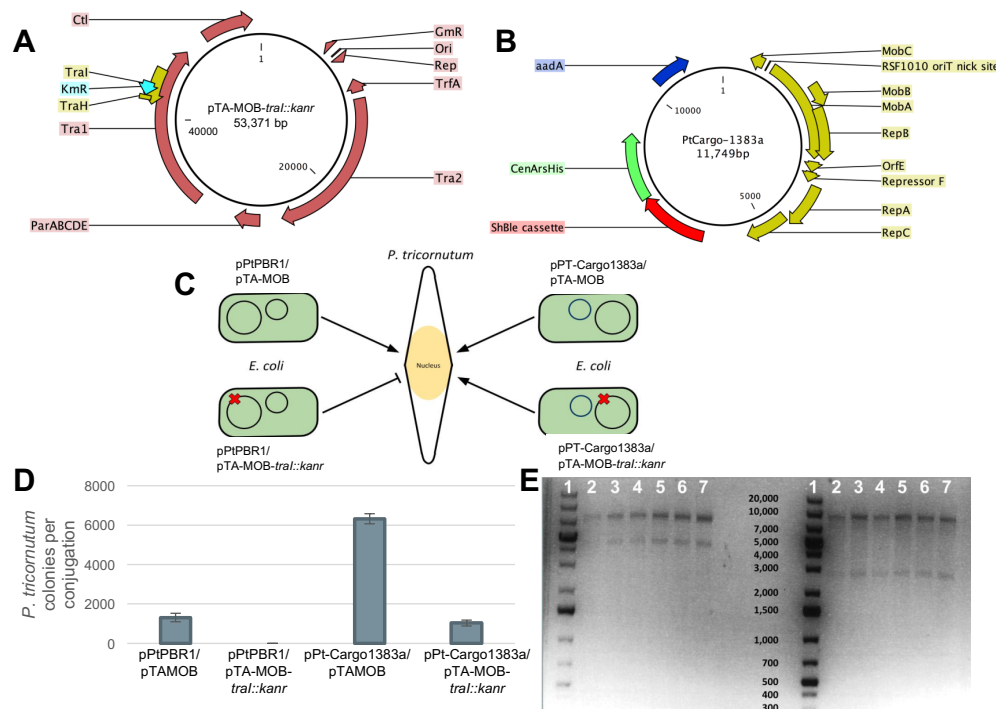


Figure 1: IncQ-based conjugation into *P. tricornutum*. **A and B.** Plasmid maps of the relaxase mutant pTA-MOB-tral::kanr and the RSF1010-based cargo plasmid, pPt-Cargo1383a. **C.** Hypothesized results for the various cargo-conjugative plasmid pairs. **D.** Results from *E. coli* conjugation to *P. tricornutum*. Colony counts per conjugation from three independent conjugations per plasmid system are reported. **E.** Digests of extracted plasmids. From each of three resulting *P. tricornutum* colonies, plasmids were extracted and transformed into *E. coli*. Two *E. coli* colonies from each of the original three *P. tricornutum* colonies were cultured. Plasmids were extracted and digested; lanes 2-3, 4-5, and 6-7 represent such *E. coli* colony pairs. The left gel displays results from EcoRV digest (4.2 kb and 7.4 kb expected band sizes), while the rightmost gel displays those of the BamHI digest (2.6 kb and 9.1 kb expected band sizes). Matching band patterns on all six samples to expected band sizes for both digests suggests post-conjugation episomal maintenance of pPt-Cargo1383a in *P. tricornutum*.

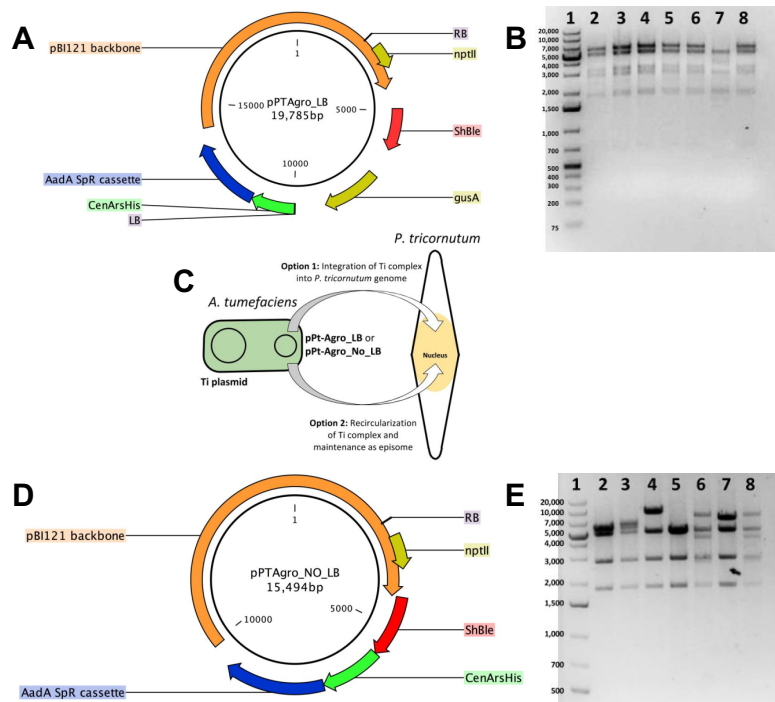


Figure 2: *A. tumefaciens* Ti-based conjugation into *P. tricornutum*. **A.** Plasmid map of pPTAgro_LB. **B.** Following extraction of conjugated plasmids from *P. tricornutum* and transformation of plasmid extracts into *E. coli*, plasmids were extracted and digested to verify post-conjugation episomal maintenance. Lanes 2-4 display the positive control of the original pPT-Agro_LB used previously to transfect the donor *E. coli* cells. Lanes 5-8 contain pPT-Agro_LB extracts of four different *E. coli* colonies from plasmids extracted from the *P. tricornutum* colony #1. Similarity in band pattern in lanes 5, 6, and 8 to the positive controls (lanes 2-4) suggest complete post-conjugation episomal maintenance of the pPT-Agro_LB plasmid in *P. tricornutum*. **C.** Diagram of experimental design for *A. tumefaciens* cargo plasmid transfer to *P. tricornutum*. **D.** Plasmid map of pPTAgro_No_LB. **E.** EcoRV restriction digest of pT-Agro_No_LB plasmids extracted from *E. coli*. Lane 2 displays the positive control of an original miniprep of pPT-Agro_No_LB used previously to transfect the donor *A. tumefaciens* cells. Lanes 3-5 contain pPT-Agro_No_LB extracts of three different *E. coli* colonies from plasmids extracted from *P. tricornutum* colony #2. Lanes 6-8 contain pPT-Agro_No_LB extracts of three different *E. coli* colonies from plasmids extracted from *P. tricornutum* colony #3. While band pattern in lane 3 matches that of the positive control in lane 2, wide variance in the size of the largest fragment prevails in the samples of lanes 3-8.

Tables:

400 **Table 1:** Primers used in this study.

Primer Name	Primer sequence (5'-3')
Tral-KO-3	GTTCTGGCTGCGGCAGGTCGAATGCCTCCATCGCCGCGCCAGTGTTACAA CCAATTAACC
Tral-KO-2	GGGCCATGCGGCCGTGACGGACAACATCACGAAGAAGGCGATTTATTCAA CAAAGCCACG
Tral-KO-scrn-3	ATGCGAACCACACCGAGTTG
Tral-KO-scrn-2	CAAGATCAACCGCGAGTACC
1383a-insert-1	CTTGCATGCCTGCAGGTCGACTCTAGGGGTACCGAGCTCGACCATGATTA CGCCAAGCTC
1383a-insert-2	TATGCATCCGCGGGCCCGGGTTACGAGATCGATGGTCAAGTCCAGACTCC TGTGTAAAC
pTA-Cargo-Scrn1	GTGGAGGCCATCAAACCACG
pTA-Cargo-Scrn2	CCAGTCCCTCCCAGCTGATG
4F	CGCACCCCCGAACACGAG
3R	AGCTAACTCACATTAATTGCGTTGCGCTC
3F	CCGCGCGTTGGCCGATTCATTAA
5F	GCGTCAATTTGTTTACACCACAATATATCCTGCCACACTAGCTCGTCTTCAA TCACGTGC
5R	GCACGTGATTGAAGACGAGCTAGTGTGGCAGGATATATTGTGGTGTAAACA AATTGACGC
6R	CCGAGCTGCCGGTCGGGGAGCTGTTGGCTGGCTGGTGTAAATAAAAAA GGGGACCTCT
6F	AATTAATTGGGGACCCTAGAGGTCCCCTTTTTTATTTTAACACCAGCCAGCC AACAGCTC
4R	TCGGGGCATTACGGACTTCATG
4F	CGCACCCCCGAACACGAG
7nR	AGGGTTAATTTTCGAGCTTGGCGTAATCATGGTCCCGATCTAGTAACATAGA TGACACCGC

7nF	CGCGGTGTCATCTATGTTACTAGATCGGGACCATGATTACGCCAAGCTCGA AATTAACCC
6nR	CGGTCGGGGAGCTGTTGGCTGGCTGGTGTAAATAAAAAAGGGGACCTC TAGGGTCCCC
Pt-AgroCheck1	CAATCAGCTGTTGCCCCGTCTCACTG
SeqTest4-R	CGTGTGTGGTCTTCTACACAGAC
Pt-Agro-1	CAGGAAACAGCTATGACCATGATTACGCCACTAACAGGATTAGTGCAATTCGAGTTGA AT
Pt-Agro-2	AGGCTAATCTGGGGACCTGCAGGCATGCAAGCTTGAAGACGAGCTAGTGTTATTCCT GAC

Table 2: Nuclear localization prediction scores for proteins sequences from RSF1010, the Tra proteins from RP4/RK2, and the Vir proteins from plasmid Ti (Genbank accession numbers in parentheses). The top 10 ranking sequences from each plasmid are listed (relaxase proteins are noted). NuPred predictions were performed according to (Brameier et al., 2007).

RSF1010 (M28829.1)		RP4/RK2 (BN000925.1)		A. tumefaciens Ti plasmid (AF242881)	
Seq-ID	NucPred-score	Seq-ID	NucPred-score	Seq-ID	NucPred-score
replication_protein_B	0.53	traI (relaxase)	0.60	virD2 (relaxase)	0.92
replication_protein_C	0.50	traK	0.41	virD5	0.89
mobilization_protein_A (relaxase)	0.48	traE	0.38	virE3	0.88
mobilization_protein_C	0.44	traC1	0.34	virB11	0.62
AAA26452.1	0.35	traC2	0.32	virA	0.60

repressor_protein_F	0.30	traH	0.16	virE2	0.58
mobilization_protein_B	0.17	traG	0.13	virD4	0.52
replication_protein_A	0.13	traB	0.07	virB10	0.51
Sm_resistance_protein_A	0.07	traA	0.05	virB1	0.49
Sm_resistance_protein_B	0.07	traJ	0.03	virC2	0.39

410

Acknowledgements: This work was supported by a Gordon and Betty Moore Foundation (GBMF5007 to P.D.W and C.L.D)

415

1. Daboussi F, Leduc S, Maréchal A, Dubois G, Guyot V, Perez-Michaut C, Amato A, Falcatore A, Juillerat A, Beurdeley M, Voytas DF, Cavarec L, Duchateau P. Genome engineering empowers the diatom *Phaeodactylum tricornutum* for biotechnology. *Nature communications*. 2014;5(May):3831-. doi: 10.1038/ncomms4831.
2. Nymark M, Sharma AK, Sparstad T, Bones AM, Winge P. A CRISPR/Cas9 system adapted for gene editing in marine algae. *Scientific reports*. 2016;6(April):24951-. doi: 10.1038/srep24951.
3. Weyman PD, Beeri K, Lefebvre SC, Rivera J, McCarthy JK, Heuberger AL, Peers G, Allen AE, Dupont CL. Inactivation of *Phaeodactylum tricornutum* urease gene using transcription activator-like effector nuclease-based targeted mutagenesis. *Plant biotechnology journal*. 2015;13(4):460-70. doi: 10.1111/pbi.12254.
4. Diner RE, Noddings CM, Lian NC, Kang AK, McQuaid JB, Jablanovic J, Espinoza JL, Nguyen NA, Anzelmatti MA, Jansson J, Bielinski VA, Karas BJ, Dupont CL, Allen AE, Weyman PD. Diatom centromeres suggest a mechanism for nuclear DNA acquisition. *Proceedings of the National Academy of Sciences*. 2017;114(29):E6015-E24. doi: 10.1073/pnas.1700764114.
5. Diner RE, Bielinski VA, Dupont C, Allen AE, Weyman PD. Refinement of the Diatom Episome Maintenance Sequence and Improvement of Conjugation-based DNA Delivery Methods. *Frontiers in Bioengineering and Biotechnology*. 2016;4(August):65-. doi: 10.3389/FBIOE.2016.00065.
6. Karas BJ, Diner RE, Lefebvre SC, McQuaid J, Phillips APR, Noddings CM, Brunson JK, Valas RE, Deerinck TJ, Jablanovic J, Gillard JTF, Beeri K, Ellisman MH, Glass JI, Hutchison Iii Ca, Smith HO, Venter JC, Allen AE, Dupont CL, Weyman PD. Designer diatom episomes delivered by bacterial conjugation. *Nature Communications*. 2015;6:6925-. doi: 10.1038/ncomms7925.
7. Dunning Hotopp JC, Clark ME, Oliveira DCSG, Foster JM, Fischer P, Muñoz Torres MC, Giebel JD, Kumar N, Ishmael N, Wang S, Ingram J, Nene RV, Shepard J, Tomkins J, Richards S, Spiro DJ, Ghedin E, Slatko BE, Tettelin H, Werren JH. Widespread lateral gene transfer from intracellular bacteria to multicellular eukaryotes. *Science (New York, NY)*. 2007;317(5845):1753-6. doi: 10.1126/science.1142490.
8. Keeling PJ, Palmer JD. Horizontal gene transfer in eukaryotic evolution. *Nature reviews Genetics*. 2008;9(8):605-18. doi: 10.1038/nrg2386.
9. Lacroix B, Citovsky V. Transfer of DNA from Bacteria to Eukaryotes. *mBio*. 2016;7(4):1-9. doi: 10.1128/mBio.00863-16.
10. Smillie C, Garcillan-Barcia MP, Francia MV, Rocha EPC, de la Cruz F. Mobility of Plasmids. *Microbiology and Molecular Biology Reviews*. 2010;74(3):434-52. doi: 10.1128/MMBR.00020-10.
11. Garcillan-Barcia M, Francia M, de la Cruz F. The diversity of conjugative relaxases and its application in plasmid classification. *FEMS Microbiology Reviews*. 2009;33(3):657-87. doi: 10.1111/j.1574-6976.2009.00168.x.
12. Bates S, Cashmore AM, Wilkins BM. IncP Plasmids Are Unusually Effective in Mediating Conjugation of *Escherichia coli* and *Saccharomyces cerevisiae*: Involvement of the Tra2 Mating System. *J Bacteriol*. 1998;180(24):6538-43.

460

13. Heinemann JA, Sprague Jr GF. Bacterial conjugative plasmids mobilize DNA transfer between bacteria and yeast. *Nature*. 1989;340(6230):205-9. Epub 07/20. doi: 10.1038/340205a0.
14. Bundock P, den Dulk-Ras a, Beijersbergen a, Hooykaas PJ. Trans-kingdom T-DNA
465 transfer from *Agrobacterium tumefaciens* to *Saccharomyces cerevisiae*. *The EMBO journal*. 1995;14(13):3206-14.
15. Schröder G, Schuelein R, Quebatte M, Dehio C. Conjugative DNA transfer into human cells by the VirB/VirD4 type IV secretion system of the bacterial pathogen *Bartonella henselae*. *Proceedings of the National Academy of Sciences of the United States of America*. 2011;108(35):14643-8. doi: 10.1073/pnas.1019074108.
- 470 16. Sikorski RS, Michaud W, Levin HL, Boeke JD, Hieter P. Trans-kingdom promiscuity. *Nature*. 1990;345:581-2.
17. Waters VL. Conjugation between bacterial and mammalian cells. *Nature genetics*. 2001;29(4):375-6. doi: 10.1038/ng779.
- 475 18. Guillard RRL, Hargraves PE. *Stichochrysis immobilis* is a diatom, not a chrysophyte. *Phycologia*. 1993;32(3):234-6. doi: 10.2216/i0031-8884-32-3-234.1.
19. Wise Aa, Liu Z, Binns AN. Culture and maintenance of *Agrobacterium* strains. *Methods in molecular biology* (Clifton, NJ). 2006;343(4):3-13. doi: 10.1385/1-59745-130-4:3.
20. Strand TA, Lale R, Degnes KF, Lando M, Valla S. A New and Improved Host-
480 Independent Plasmid System for RK2-Based Conjugal Transfer. *PloS one*. 2014;9(3):e90372-e. doi: 10.1371/journal.pone.0090372.
21. Datsenko KA, Wanner BL. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proceedings of the National Academy of Sciences of the United States of America*. 2000;97(12):6640-5. doi: 10.1073/pnas.120163297.
- 485 22. Wolk CP, Fan Q, Zhou R, Huang G, Lechno-Yossef S, Kuritz T, Wojciuch E. Paired cloning vectors for complementation of mutations in the cyanobacterium *Anabaena* sp. strain PCC 7120. *Archives of microbiology*. 2007;188(6):551-63. doi: 10.1007/s00203-007-0276-z.
23. Meyer R. Replication and conjugative mobilization of broad host-range IncQ plasmids. *Plasmid*. 2009;62(2):57-70. doi: 10.1016/j.plasmid.2009.05.001.
- 490 24. Falciatore A, Casotti R, Leblanc C, Abrescia C, Bowler C. Transformation of Nonselectable Reporter Genes in Marine Diatoms. *Mar Biotechnol*. 1999;1(3):239-51.
25. Grahn AM, Haase J, Bamford DH, Lanka E. Components of the RP4 conjugative transfer apparatus form an envelope structure bridging inner and outer membranes of donor cells: Implications for related macromolecule transport systems. *Journal of Bacteriology*. 2000;182(6):1564-74. doi: 10.1128/JB.182.6.1564-1574.2000.
- 495 26. Arutyunov D, Arenson B, Manchak J, Frost LS. F plasmid TraF and TraH are components of an outer membrane complex involved in conjugation. *Journal of Bacteriology*. 2010;192(6):1730-4. doi: 10.1128/JB.00726-09.
27. Buchanan-Wollaston V, Passiatore JE, Cannon F. The mob and oriT mobilization
500 functions of a bacterial plasmid promote its transfer to plants. *Nature*. 1987;328(6126):172-5.
28. Meyer R. Mapping Type IV Secretion Signals on the Primase Encoded by the Broad-Host-Range Plasmid R1162 (RSF1010). *Journal of Bacteriology*. 2015;197(20):3245-54. doi: 10.1128/JB.00443-15.
29. Cha T-S, Chen C-F, Yee W, Aziz A, Loh S-H. Cinnamic acid, coumarin and vanillin:
505 Alternative phenolic compounds for efficient *Agrobacterium*-mediated transformation of the unicellular green alga, *Nannochloropsis* sp. *Journal of microbiological methods*. 2011;84(3):430-4. doi: 10.1016/j.mimet.2011.01.005.
30. Kathiresan S, Chandrashekar a, Ravishankar Ga, Sarada R. *Agrobacterium* -Mediated Transformation in the Green Alga *Haematococcus Pluvialis* (Chlorophyceae, Volvocales).
510 *Journal of Phycology*. 2009;45(3):642-9. doi: 10.1111/j.1529-8817.2009.00688.x.

31. Kumar SV, Misquitta RW, Reddy VS, Rao BJ, Rajam MV. Genetic transformation of the green alga--*Chlamydomonas reinhardtii* by *Agrobacterium tumefaciens*. *Plant Science*. 2004;166(3):731-8.
- 515 32. Prasad B, Vadakedath N, Jeong H-j, General T, Cho M-G, Lein W. *Agrobacterium tumefaciens*-mediated genetic transformation of haptophytes (*Isochrysis* species). *Applied Microbiology and Biotechnology*. 2014;98(20):8629-39. doi: 10.1007/s00253-014-5900-7.
33. Ballas N, Citovsky V. Nuclear localization signal binding protein from *Arabidopsis* mediates nuclear import of *Agrobacterium* VirD2 protein. *Proceedings of the National Academy of Sciences of the United States of America*. 1997;94(20):10723-8. doi: 10.1073/pnas.94.20.10723.
- 520 34. Citovsky V, Zupan J, Warnick D, Zambryski P. Nuclear localization of *Agrobacterium* VirE2 protein in plant cells. *Science*. 1992;256(5065):1802-5. doi: 10.1126/science.1615325.
35. Rossi L, Hohn B, Tinland B. The VirD2 protein of *Agrobacterium tumefaciens* carries nuclear localization signals important for transfer of T-DNA to plant. *Molecular & general genetics* : MGG. 1993;239(3):345-53.
- 525