

1 **1 Targeted expression of nuclear transgenes in *Chlamydomonas reinhardtii* with a**
2 **2 versatile, modular vector toolkit**

3 **3 Author names and affiliations:** Kyle J. Lauersen, Olaf Kruse and Jan H. Mussgnug*
4 4 Bielefeld University, Faculty of Biology, Center for Biotechnology (CeBiTec),
5 5 Universitätsstrasse 27, 33615, Bielefeld, Germany.

6 ***Corresponding Author:** Jan H. Mussgnug, jan.mussgnug@uni-bielefeld.de

7 **7 Present/Permanent address:** Bielefeld University, Faculty of Biology, Center for
8 8 Biotechnology (CeBiTec), Universitätsstrasse 27, 33615 Bielefeld, Germany. Phone:
9 9 +49 521 106-12260, Fax: +49 521 106-12290

10

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

1 **Abstract**

2 We present a versatile vector toolkit for nuclear transgene expression in the model
3 green microalga *Chlamydomonas reinhardtii*. The vector was designed in a modular
4 fashion which allows quick replacement of regulatory elements and genes of interest.
5 The current toolkit comprises two antibiotic resistance markers (paromomycin and
6 hygromycin B), five codon optimized light-emission reporters, including the *Gaussia*
7 *princeps* luciferase, as well as bright cyan, green, yellow, and red fluorescent protein
8 variants. The system has demonstrated robust functional flexibility with signal options
9 to target the protein of interest to the cytoplasm, the nucleus, cellular microbodies, the
10 chloroplast, mitochondria, or via the endoplasmic reticulum-Golgi apparatus secretory
11 pathway into the culture medium. Successful fluorescent reporter protein fusion to *C.*
12 *reinhardtii* Rubisco small subunit 1 was accomplished with this system. Localization
13 of the fluorescently tagged protein was observed in the chloroplast pyrenoid via live-
14 cell fluorescence microscopy, the first report of heterologous protein localization to
15 this cellular structure. The functionalities of the vector toolkit, the individual modular
16 elements, as well as several combinations thereof are demonstrated in this manuscript.
17 Due to its strategic design, this vector system can quickly be adapted to individual
18 tasks, and should, therefore, be of great use to address specific scientific questions
19 requiring nuclear recombinant protein expression in *C. reinhardtii*.

20
21 **Keywords:**

22 Microalga, *Chlamydomonas reinhardtii*, recombinant proteins, fluorescent proteins,
23 luciferase, protein secretion

1 **Introduction**

2 Microalgae are a heterogeneous group of organisms, including photosynthetic pro-
3 and eukaryotes which are sources of biotechnologically relevant products (Hallmann
4 2007; Wijffels et al. 2013). Growth and inexpensive cultivation with sunlight energy
5 in simple mineral salt solutions make microalgae attractive as sustainable sources of a
6 vast array of natural bio-products with many application potentials. Ease of handling
7 and relatively rapid turnaround time from gene of interest to scale-up compared to
8 other photosynthetic systems are additional advantages of these unicellular production
9 organisms (León-Bañares et al. 2004). Microalgal biotechnology has seen a boom in
10 interest over the last 20 years, especially with regards to the potential of microalgal
11 biomass as third generation biofuel feedstock (Stephens et al. 2010). Currently,
12 microalgae are used as protein rich feed for agri- and aquaculture, sources of pigments
13 and antioxidants for cosmetics and food, sources of oils with variable fatty acid
14 compositions (reviewed in (Hallmann 2007; Draaisma et al. 2013; Wijffels et al.
15 2013)), and, more recently, as hosts for recombinant proteins (RP) for molecular
16 farming and proposed as hosts for metabolic engineering (Shao and Bock 2008;
17 Eichler-Stahlberg et al. 2009; Neupert et al. 2009; Cordero et al. 2011; Lauersen et al.
18 2013b; Rasala et al. 2014).

19 Transgenes can be expressed in microalgae from nuclear, chloroplast, or
20 mitochondrial genomes (Kindle 1990; Bateman and Purton 2000; Remacle et al.
21 2006) and since the host cells can combine prokaryotic and eukaryotic expression
22 features, they offer some unique possibilities for the production of RPs, especially
23 when their functions depend on specific post-translational modifications or complex
24 protein folding (Rasala and Mayfield 2014). Genetic manipulation has been discussed
25 for several decades as the means to achieve some of the ambitious goals of microalgal
26 technologies, including the competitive production of lipids for sustainable solar-
27 based liquid fuels, or the synthesis of complex chemicals which do not occur
28 naturally, by metabolic engineering (Vieler et al. 2012; Bogen et al. 2013; Liu et al.
29 2013).

30 Although several algal species have been transformed, including but not limited to
31 *Nannochloropsis sp.* (Kilian et al. 2011; Radakovits et al. 2012; Vieler et al. 2012),
32 *Phaeodactylum tricorutum* (Apt et al. 1996; Miyahara et al. 2014), and *Dunaliella*
33 *salina* (Feng et al. 2014a; Feng et al. 2014b), the alga with the most highly developed
34 molecular toolkit is the unicellular Chlorophyte *Chlamydomonas reinhardtii*. This

1 model microalga has served as a valuable tool for the study of photosynthesis for over
2 60 years, and has recently developed as a platform to model recombinant algal
3 technologies (Schwarz et al. 2007; Shao and Bock 2008; Eichler-Stahlberg et al.
4 2009; Neupert et al. 2009; Rasala et al. 2010; Chen and Melis 2013; Lauersen et al.
5 2013b; Rasala and Mayfield 2014). It is well documented that RP production from *C.*
6 *reinhardtii* chloroplasts can be economically viable, a feature owing to RP titers of
7 several percent total soluble protein (TSP) and reliable targeted genetic manipulation
8 by homologous recombination (Bateman and Purton 2000; Rasala et al. 2011; Chen
9 and Melis 2013). However, this strategy of gene expression is limited to products
10 which can accumulate as soluble protein within the stroma of the plastid, and which
11 do not require post-translational modifications other than disulphide bridging (Rasala
12 and Mayfield 2014).

13 Expression of RPs from the nuclear genome, by contrast, results in significantly lower
14 titers of RP, but allows for an array of alternative applications for the recombinant
15 product, including specific subcellular targeting, RP secretion, post-translational
16 modifications including glycosylation, and plasma membrane localization. The
17 capacity for subcellular targeting in the eukaryotic cell provides the possibility of
18 accessing specialized substrates within cellular compartments, a valuable trait for
19 metabolic engineering and *in vivo* bio-product manipulation.

20 Transformation of the nuclear genome of *C. reinhardtii* was first demonstrated in
21 1989 (Kindle et al. 1989; Kindle 1990), and for many years was limited to auxotrophy
22 complementation with genomic fragments for phenotype complementation (Rochaix
23 1995). Various antibiotic selection markers have since been demonstrated as effective
24 for use in nuclear transformation, including the aminoglycoside (3')
25 phosphotransferases *aphVII* and *aphVIII* from *Streptomyces hygroscopicus* and *S.*
26 *rimosus* (hygromycin B and paromomycin resistance, respectively), as well as the
27 bleomycin-antibiotic-family binding *Sh ble* from *Streptoalloteichus hindustanus*
28 (Lumbreras et al. 1998; Sizova et al. 2001; Berthold et al. 2002). These selection
29 markers were used directly from their bacterial origins without sequence
30 modification, owing to their high GC contents that match the third position GC bias of
31 the *C. reinhardtii* nuclear genome.

32 Heterologous reporters have also developed considerably in the last ten years, owing
33 in a large part to the rapid decrease of cost and ease of gene synthesis technologies, a
34 requirement to adapt codon usage of target sequences. Luciferases from *Renilla*

1 *reniformis* (sea pansy) and *Gaussia princeps* (a mesopelagic copepod) have been
2 codon optimized and successfully expressed from the nuclear genome of *C.*
3 *reinhardtii* (Fuhrmann et al. 2004; Ruecker et al. 2008; Shao and Bock 2008). In
4 addition to the green fluorescent protein (GFP) (Fuhrmann et al. 1999), a rapid
5 expansion of the spectral palette of available reporters has recently been achieved
6 (Rasala et al. 2013). Robust fluorescent reporter proteins provide many advantages for
7 analysis of subcellular targeting peptides and *in vivo* protein localization through
8 fusion to protein targets. Codon optimized GFP (CrGFP or cGFP) was demonstrated
9 to accumulate in the nucleus as a *Sh ble*-fusion (Fuhrmann et al. 1999; Rasala et al.
10 2012), and fusion of this reporter with either cNAPL or RAA4 demonstrated
11 chloroplast localization (Glanz et al. 2006; Glanz et al. 2012). In a recent publication,
12 mCerulean and mCherry reporter proteins were successfully targeted to the
13 chloroplast, nucleus, endoplasmic reticulum, or mitochondria (Rasala et al. 2014). In
14 addition, a different report discussed the targeting of the green fluorescent protein to
15 peroxisome-like microbody structures in *C. reinhardtii* (Hayashi and Shinozaki
16 2012).

17 A large amount of data now exists surrounding the regulation of nuclear expressed
18 transgenes, for example, the *Sh ble* gene and *Renilla* luciferase were both used to
19 demonstrate the enhancer element effect of introns in heterologous sequences on
20 protein expression levels (Lumbreras et al. 1998; Eichler-Stahlberg et al. 2009). In
21 addition, several endogenous promoters and synthetic fusion promoters have
22 demonstrated reliable gene expression (Sizova et al. 1996; Fischer and Rochaix 2001;
23 Berthold et al. 2002; Neupert et al. 2009). Very recently, successful expression of
24 multiple proteins from a single open reading frame was accomplished via inclusion of
25 viral 2A peptides (Rasala et al. 2012). However, most previous molecular tools were
26 exclusively constructed for specific nuclear target genes and the vector designs did
27 not incorporate options for efficient adaptation, allowing more versatile use.

28 Here, we describe the development of the novel pOptimized vector system. This
29 system was designed *in silico* and built completely *de novo* to be a modular,
30 standardized, and expandable vector toolkit for general use in microalgal research
31 with *C. reinhardtii*. The system employs a modular genetic element approach for
32 nuclear gene expression, codon optimized variants of four fluorescent proteins, one
33 luciferase, and modified versions of two antibiotic resistance markers are presented as
34 the foundation of this expandable system.

1 **Materials and Methods**

2 **Cultivation conditions and *Chlamydomonas reinhardtii* strains**

3 *C. reinhardtii* UVM4 cultures (graciously provided by Prof. Dr. Ralph Bock) were
4 routinely grown in TAP media with 150 μ E light intensity in shake flasks or on
5 TAP agar plates. UVM4 is a ultraviolet light derived mutant of CC-4350 (cw15 arg7-
6 8 mt+ [Matagne 302]) which was co-transformed with the emetine resistance cassette
7 CRY1 and ARG7, and demonstrated expression of nuclear transgenes with high
8 efficiency (Neupert et al. 2009). CC-4350 is available from the Chlamydomonas
9 Resource Center (<http://chlamycollection.org>). Transformations were performed with
10 glass bead agitation as previously described (Kindle 1990), with only 6 h recovery of
11 cells in liquid TAP rather than 18 h (Lumbreras et al. 1998). Positive transformants
12 were recovered on TAP agar plates containing respective antibiotics at 10 mg L⁻¹ with
13 150 μ E light intensity, and maintained on TAP agar by colony stamping. Routine
14 cultivation of transformants was conducted in liquid TAP media in shake flasks at
15 light intensities indicated above.

16 **Plasmid construction, codon optimization, promoter analysis, and antibiotic 17 resistance genes**

18 The pOptimized vector (pOpt) *Chlamydomonas* expression unit was designed *in silico*
19 and built *de novo* by oligonucleotide annealing gene synthesis into the
20 pBluescript II KS (+) *Escherichia coli* vector (Genscript, USA). The hygromycin B
21 resistance gene (*aphVII*, NCBI accession: CAF31839.1) was modified from that of
22 pHyg4 (Berthold et al. 2002), to remove the redundant restriction sites *EcoRI*
23 (GA(A/G)TTC), *BsiWI* (C(G/C)TACG), as well as *AatII* (GACGT(C/G)), and
24 synthesized as part of the first pOptimized vector (pOpt_gLuc_Hyg) between
25 5' *HindIII* and 3' *XhoI* sites. This vector was synthesized containing the synthetic
26 *Gaussia princeps* luciferase (sequence modifications described below) within a
27 second gene expression cassette.

28 The paromomycin resistance gene (*aphVIII*) region was amplified using vector pJR38
29 as template (Neupert et al. 2009). First, the region containing the expression cassette
30 from the HSP70A promoter until the TGA of *aphVIII* was amplified from pJR38 and
31 cloned into the pOpt_vector backbone from *MluI* to *XhoI* to demonstrate that the
32 downstream region was not important for antibiotic resistance activity. Second, the
33 region directly after the *MscI* site of pJR38 until the *aphVIII* stop codon was amplified

1 with primers containing *HindIII-XhoI* sites and cloned into these sites to replace the
2 hygromycin B resistance gene of pOpt_gLuc_Hyg to build the vector
3 pOpt_gLuc_Paro. These two vectors served as the templates for all variant vectors
4 described in this work.

5 The amino acid (aa) sequences for synthetic FP variants mCerulean3 (cyan, variant of
6 PDB: 4EN1_A as described in (Markwardt et al. 2011)), mVenus (yellow, (Kremers
7 et al. 2006), NCBI Accession: AAZ65844.1), mRuby2 and Clover (red and green
8 variants, AFR60232.1 and AFR60231.1, respectively (Lam et al. 2012)), and gLuc
9 were codon optimized as previously described (Verhaegen and Christopoulos 2002;
10 Ruecker et al. 2008; Lauersen et al. 2013a). All optimized sequences were modified
11 with synonymous codons to remove any redundant restriction sites within the
12 pOptimized vector concept (Table S1). The 329 bp intron 2 (i2) of the Rubisco small
13 subunit 2 (RBCS2) gene of *C. reinhardtii* (NCBI Accession: X04472.1) (Eichler-
14 Stahlberg et al. 2009) was integrated into each reporter to maintain a similar
15 CG/.../GT integration site near the middle of each sequence. Where necessary,
16 synonymous codons were used to modify the sequences to create this motif. Each
17 reporter was synthesized with 5' *NdeI*, *BglIII*, and *AatII* restriction sites as well as 3'
18 *EcoRV* and *EcoRI* sites and cloned into either pOpt_gLuc_Hyg or pOpt_gLuc_Paro
19 between *NdeI* and *EcoRI* to create the template vectors.

20 The four fluorescent reporter sequences above were also codon optimized as single
21 genes for expression in *E. coli* (Genscript, USA, Sequences and NCBI accession
22 numbers provided in Supplemental Data file), and cloned into the pET24a(+)
23 expression vector (Novagen) between *NdeI* and *XhoI* sites, followed by
24 transformation into chemically competent *E. coli* KRX cells (Promega). *E. coli*
25 expressed recombinant proteins were induced overnight with 1 mM IPTG in lysogeny
26 broth (LB media), followed by freeze-thaw lysis, and Strep-Tactin chromatography
27 routinely following manufacturer's protocols (IBA Life Sciences).

28 Targeting of reporters for secretion from *C. reinhardtii* was achieved by 5' fusion of
29 the *cCA* secretion signal between *NdeI* and *BglIII* within each pOptimized vector as
30 previously described (Lauersen et al. 2013a; Lauersen et al. 2013b). Confirmation of
31 microbody targeting was achieved by 3' addition of the three PTS1-like sequences,
32 presented in Table S2, between the *EcoRV* and *EcoRI* restriction sites of the
33 pOpt_mVenus_Paro vector. The minimal targeting sequence of the Simian Virus
34 nuclear localization sequence (SV40 NLS) was built by oligonucleotide annealing of

1 the respective DNA sequence, also presented in Table S2. Construction of both the 36
2 aa *C. reinhardtii* photosystem I reaction center subunit II (PsaD) and 46 aa
3 mitochondrial ATP synthase subunit A (AtpA) N-terminal targeting peptide DNA
4 sequences was achieved by annealing equimolar ratios of the four primers for each
5 construct, listed in Table S2, in polymerase chain reactions (PCR) followed by
6 restriction digest and ligation between *NdeI* and *BglIII* sites within the
7 pOpt_mVenus_Paro vector.

8 Localization of the ribulose-1,5-bisphosphate carboxylase small-subunit 1 (RBCS1)
9 with mVenus fluorescence tagging was achieved by *de novo* synthesis of the RBCS1
10 mRNA sequence (Genscript, USA) (Goldschmidt-Clermont and Rahire 1986; Genkov
11 et al. 2010), followed by cloning of the sequence between *NdeI* and *BglIII* in the
12 pOpt_mVenus_Paro vector, resulting in a C-terminal fusion of mVenus to the RBCS1
13 protein sequence. Truncations of RBCS1 N-terminal region were performed by PCR
14 amplification of the 5' region of the RBCS1 sequence with the primers listed in
15 Table S2.

16 All cloning in this work was performed with Fermentas Fastdigest restriction enzymes
17 following manufacturer's protocols, alkaline phosphatase and ligation reactions were
18 performed with the Rapid DNA Dephos & Ligation Kit (Roche). All PCRs were
19 performed with Q5 High Fidelity polymerase with GC enhancer solution (New
20 England Biolabs) following manufacturer's protocols and cloning was conducted as
21 described above using primers listed in Table S2. After each cloning step, vector
22 sequences were confirmed by sequencing (Sequencing Core Facility, CeBiTec,
23 Bielefeld University, Germany).

24 The template vector sequences and feature annotation for pOpt_X_Paro and
25 pOpt_X_Hyg, where X is any of the five reporters described above, have been
26 deposited to NCBI (Access. No. can be found in Table S3). Reporter vectors,
27 including *cCA* secretion variants, for both antibiotic resistance markers, as well as the
28 four *E. coli* reporter expression strains, have been made available at the
29 Chlamydomonas Resource Center (<http://chlamycollection.org>), including supporting
30 vector sequence information. All vectors for *C. reinhardtii* transformation created in
31 this work are listed in Table S4, variants not deposited to the Resource Center can be
32 generated with the oligonucleotides listed in Table S2 as described above.

1 **Bioluminescence**

2 Bioluminescence was analyzed as previously described for mutant identification at
3 plate level (Lauersen et al. 2013a). Recombinant Gaussia luciferase was detected in
4 Western or dot-blot analysis with an anti-Gaussia luciferase antibody (New England
5 Biolabs, MA, USA) using secreted gLuc standard from *Kluyveromyces lactis* as
6 previously described (Lauersen et al. 2013a).

7 **Fluorescent protein analysis**

8 Positive fluorescent protein (FP) expression was detected from plate level *C.*
9 *reinhardtii* cultures using a Leica-Binocular MZFLIII system equipped with filters for
10 CFP, GFP, YFP, and DsRED to detect mCerulean3, Clover, mVenus, and mRuby2,
11 respectively. Fluorescence imaging was carried out using a confocal laser-scanning
12 microscope with respective filter sets and excitation and emission wavelengths
13 previously described for each reporter (Table S5) (LSM780, Carl Zeiss GmbH,
14 Germany). mCerulean3 was detected by excitation with 458 nm laser and emission
15 reading between 460-490 nm, Clover was detected by excitation with 488 nm laser
16 and detection between 500-530 nm, mVenus was detected with excitation at 514 nm
17 and emission between 520-550 nm, mRuby2 was detected with excitation at 561 nm
18 and emission between 590-620 nm. Chlorophyll fluorescence was observed
19 independently with stimulation at 488 nm and emission from 650-700 nm. Live cells
20 were allowed to settle in microtiter plates, and 8 µl was pipetted onto a glass slide
21 prior to microscopy analysis.

22 Supernatants containing fluorescent proteins were harvested from stationary phase
23 cultures by centrifugation for 3 min at 3000xg in 50 mL Falcon tubes, followed by
24 0.2 µm micro filtration, and concentration using 10 kDa MWCO centrifugal filter
25 units (Millipore) to a final concentration of 10X for each FP. Concentrated media was
26 visualized through Leica-Binocular filters and captured with a digital camera through
27 the auxiliary viewing lens of the microscope. Proteins were analyzed by either
28 Western- or dot-blot analysis using a GFP Rabbit IgG Antibody Fraction HRP
29 Conjugate (Life Technologies, Germany). Green fluorescent protein (GFP) purified
30 from *E. coli* was used as a standard (graciously provided by Prof. Dr. Karl Friehs).

31

1 Results

2 Vector design, codon optimization, and expression of recombinant reporters in 3 *C. reinhardtii*

4 The general vector design is depicted in Fig. 1. In order to develop a standardized,
5 modular, and easily adjustable gene expression system for *C. reinhardtii*, two separate
6 expression cassettes were included in the vector, one for the expression of an
7 antibiotic resistance marker and the other for the expression of the gene of interest
8 (GOI). The heat shock 70A-Rubisco small subunit 2 fusion promoter with Rubisco
9 small subunit intron 1 (collectively HSP70A-RBCS2-i1) and RBCS2 3' untranslated
10 region (3' UTR) were chosen as regulatory elements for both cassettes, because these
11 elements resulted in efficient transgene expression in previous work (Lauersen et al.
12 2013a). To achieve an overall modular design, each regulatory element was separated
13 by a unique restriction enzyme recognition sequence (Table S1). In addition, a
14 multiple cloning site for future modifications was added between the two gene
15 expression cassettes.

16 Two antibiotic resistance genes, *aphVII* and *aphVIII*, were used in this study, and both
17 conferred growth of positive transformants on medium containing the antibiotics
18 hygromycin B (10 mg L⁻¹) or paromomycin (10 mg L⁻¹), respectively (Fig. 2a,c). For
19 paromomycin resistance, we determined that a DNA fragment containing only a
20 839 bp region up to the TGA stop codon, shorter than the 904 bp *aphVIII* fragment
21 described before (Sizova et al. 2001), was sufficient to confer antibiotic resistance
22 (Fig. 2b,c). The *aphVII* gene for hygromycin B resistance was modified slightly in its
23 nucleotide sequence with synonymous codons to remove unwanted restriction enzyme
24 recognition sequences (Fig. S1), which did not result in a loss of antibiotic resistance
25 activity for selection of positive transformants. Both constructs resulted in
26 transformation efficiencies of ~1x10³ colonies per µg vector for either resistance
27 cassette under optimal transformation conditions.

28 Since the expression cassette for the GOI was separated from the antibiotic resistance
29 marker, antibiotic resistance does not necessarily guarantee efficient expression of the
30 GOI. Light emitting reporter genes (gLuc, mCerulean3, Clover, mVenus, and
31 mRuby2) were therefore cloned into the GOI expression cassette after codon
32 optimization and removal of interfering restriction enzyme cut sites to allow fast
33 evaluation of GOI transgene expression efficiencies. Again, a modular design was

1 chosen and complementary restriction sites were placed on the 5' and 3' ends so that
2 the GOI reporter can be easily replaced in the expression vector backbone and used in
3 fusion protein concepts (Fig. 2d). Each reporter protein was constructed to contain the
4 second intron of RBCS2 (i2) in the middle of its DNA sequence as it has been
5 previously shown that introns from this gene in their physiological order result in
6 increased transgene expression (Eichler-Stahlberg et al. 2009). As a result, each
7 reporter exhibited expression that was detectable at the plate level (Fig. 3), with single
8 gene products detected at the protein level by Western blot (Fig. S2a-c), indicating no
9 alternative splicing or protein degradation. Rates of false positive transformants were
10 determined in three independent transformations of vectors pOpt_mVenus_Hyg and
11 pOpt_mVenus_Paro. For both vectors it was determined that $30\% \pm 5\%$ of the
12 antibiotic resistant transformants showed detectable levels of YFP fluorescence on
13 agar plates.

14 Reporter activities were detectable in colonies at the plate level (Fig. 3a,b,c). In case
15 the 21 aa carbonic anhydrase 1 secretion signal (cCA, Fig. 2e) was cloned N-terminal
16 to the reporter proteins, these were efficiently secreted, and were also detectable at the
17 plate level (Fig. 3a,c(vi-ix)) or in culture supernatant (Fig. 3d). Here, secreted reporter
18 proteins accumulated up to $\sim 0.7 \text{ mg L}^{-1}$ in standard cultivation conditions (Fig.
19 S3a,b). The unique spectral properties of each reporter protein were visible under
20 fluorescent protein filters in 10X concentrated media samples without further affinity-
21 based purification being necessary (Fig. 3d(i-v)). In case no targeting signal was used,
22 the recombinant reporter proteins were also strongly expressed and accumulated in
23 the cytoplasm to up to 0.25% of TSP (Fig. 3b,c(ii-v), Fig. S3c). For all fluorescence
24 reporter constructs, it was possible to visualize the respective intracellular,
25 cytoplasmic signals by laser scanning confocal microscopy with respective excitation
26 and emission wavelengths (Fig. 4a-d). However, for intracellular expression strains,
27 the mVenus signal proved to be superior to the other markers tested, since this
28 reporter exhibits minimal spectral interference with the chlorophyll pigment
29 background fluorescence as well as a bright signal in confocal microscopy (Fig. 3c(i),
30 Fig. 4b). For these reasons, mVenus was chosen as a reporter for subsequent
31 subcellular targeting studies.

1 Subcellular targeting and fluorescent protein tagging

2 Expression of the reporters in the cytoplasm presented a pattern of fluorescence which
3 filled the shape of the cell and was clearly distinguishable from the red chlorophyll
4 background fluorescence signal (Fig. 4a-d). Therefore, cytoplasmic localization
5 provided a signal pattern that was used for comparison of constructs targeted to
6 various cellular compartments. In our setup, the parental strain demonstrated typical
7 red chlorophyll fluorescence background signal but no emission signal in the yellow
8 range (Fig. 5a). The chlorophyll background fluorescence was used to orient the cells
9 (red signal in all confocal microscopy pictures), and the mVenus fluorescence signal
10 was clearly distinguishable in all cases (Fig. 5b-j). Expressed mVenus reporter
11 protein, without targeting signals or fusions, accumulated in the cell cytosol (Fig. 5b).
12 We then used mVenus fusion constructs to successfully target the reporter to the
13 ER/secretion pathway (Fig. 5c), the chloroplast pyrenoid structure (Fig. 5d),
14 intracellular microbodies (with three different signals, Fig. 5e-g), the nucleus
15 (Fig. 5h), the chloroplast (Fig. 5i), and mitochondria (Fig. 5j). The corresponding
16 individual emission channels for each construct are presented in Fig. S4a-j.

17 For microbody targeting, we tested three separate potential signal peptides with our
18 mVenus containing pOptimized plasmid. These included the C-terminal, 10 aa,
19 proposed type 1 peroxisomal targeting sequence from pumpkin malate synthase
20 (PkPTS1, IHHPRELSRL*, vector Fig. 2f), the C-terminal, 10 aa sequence from *C.*
21 *reinhardtii* malate synthase (MSPTS1, HIVTKTPSRM*, vector Fig. 2g), as well as a
22 novel, seven aa sequence. The latter is shorter than the malate synthase sequences, but
23 contains similar amino acids in a randomly shuffled sequence order (rPTS1,
24 SWRRSLI*, vector Fig. 2h). Hayashi and Shinozaki (2012) demonstrated that C-
25 terminal addition of these PTS1 sequences to GFP resulted in accumulation of this
26 reporter in small spots within the cell. We were able to confirm these results, since the
27 mVenus reporter was detected in cellular compartments resembling spherical
28 microbodies in confocal laser scanning microscopy analyses (Fig. 5e,f). This
29 localization pattern was also true for the novel seven aa sequence, indicating
30 functional exchangeability of amino acids (Fig. 5g). The signals were located
31 throughout the cell close to the chloroplast and were determined to be not located in
32 mitochondria by co-staining with mitotracker and 3-dimensional modeling of Z-stack
33 confocal microscope imaging (Fig. S5a,b and Movie S1).

1 The monopartite simian virus 40 large T-antigen nuclear localization signal
2 (SV40 NLS) was first described in 1984 as a sequence responsible for protein import
3 into eukaryotic nuclei (Kalderon et al. 1984). In order to test the greater applicability
4 of reporter localization via peptide transport signals, we cloned nucleotides coding for
5 the 7 aa SV40 NLS (PKKKRKV*, vector Fig. 2i) onto the C-terminus of the mVenus
6 reporter and expressed this construct from the nucleus of *C. reinhardtii*. Colonies
7 exhibiting yellow fluorescence at plate level were analyzed by confocal microscopy to
8 elucidate the subcellular localization of fluorescence signals. SV40 NLS-tagged
9 mVenus accumulated in the nucleus of *C. reinhardtii* cells (Fig. 5h) with fluorescence
10 patterns resembling those recently described from similar mCerulean constructs
11 containing a 2X repeat of this sequence (Rasala et al. 2014), indicating successful
12 nuclear localization of the reporter protein.

13 To demonstrate the versatility of the pOptimized vector system for full-length protein-
14 FP tagging and *in vivo* localization, we chose the Rubisco small subunit 1 (RBCS1),
15 which had previously been shown to be amenable to cDNA based expression,
16 modification with plant homologues, and intron rearrangements while maintaining
17 proper expression (Genkov et al. 2010). The cDNA sequence of RBCS1 was cloned
18 5' to the mVenus reporter, so that the fluorescent protein was fused to the C-terminus
19 of the target protein (Fig. 2l). In this orientation, intron 1 (from the fusion promoter)
20 and intron 2 (in the mVenus reporter) of RBCS2 surround the sequence of RBCS1. It
21 was expected that fluorescence signals would accumulate in the chloroplast of
22 transformant cells. Interestingly, the fluorescence signal was localized at the pyrenoid
23 structure of the cells within the chloroplast (Fig. 5d), which potentially indicated that
24 the fusion protein was incorporated into the Rubisco multi-subunit enzyme complex.
25 The full-length fusion protein product was detected at the predicted molecular mass of
26 ~51 kDa by Western blot against the mVenus protein (Fig. S6a).

27 In order to investigate whether the plastid targeting amino acid sequence of RBCS1
28 could be used for plastid localization of reporter proteins on its own, constructs
29 containing the 45 aa N-terminal targeting peptide, as well as a series of four 10 aa
30 extensions into the RBCS1 sequence were analyzed for their ability to transport
31 mVenus to the chloroplast (vectors Fig. 2m-q). For construct T1 (truncation 1), with
32 no additional amino acids past the cleavage peptide, mVenus fluorescence could be
33 detected in the chloroplast (Fig. S6b) as was observed with the PsaD targeting peptide
34 (Fig. 5i, vector Fig. 2j). However, T2-5, which contained 10-40 aa downstream of the

1 RBCS1 cleavage peptide, resulted in highly unstable fluorescence and poor imaging
2 during confocal analyses as well as the detection of only degradation products of the
3 mVenus in Western blotting analyses (Fig. S6b and not shown). Fluorescence signals
4 from chloroplast localized mVenus were difficult to image as the signals were spread
5 throughout the stroma, exhibiting a horseshoe shaped pattern of diffuse fluorescence
6 signal around the darker cytoplasmic space (Fig. 5i, 6d, S6b). It was therefore difficult
7 to focus on one cellular structure to obtain a crisp image in comparison to the strong
8 signal obtained from mitochondrial localized AtpA_mVenus (Fig. 5j, vector Fig. 2k).
9 The pOptimized system presents the option of multiple construct transformations
10 using two separate proteins in GOI expression cassettes of vectors containing either
11 paromomycin or hygromycin B resistance cassettes. For each of the subcellular
12 localized constructs, we also transformed these into cytosolic mCerulean3 expressing
13 strains generated with hygromycin B selection using the pOpt_mCerulean3_Hyg
14 vector. As shown in Fig. 6, the cytoplasmic signals of the mCerulean3 reporter
15 contrasted clearly with the mVenus signals, which were directed to cytoplasm,
16 microbodies, pyrenoid, or chloroplast, respectively (Fig. 6a-d). For the SV40 NLS
17 tagged mVenus construct, we transformed this into a strain previously transformed
18 with the pOpt_cCA_mCerulean3_Hyg vector, resulting in a faint mCerulean3 signal
19 observed from the secretory pathway radiating outwards from the mVenus signal
20 localized in the nucleus (Fig. 6e). Individual imaging layers and combined channel
21 overlays for these constructs as well as comparisons of control signals derived from
22 the parental strain are presented in Fig. S7a-c, Fig. S8a,b, Fig. S9a,b, and Fig. S10a,b.
23 To obtain pure recombinant reporter proteins for signal comparisons, each fluorescent
24 reporter was also expressed in *E. coli* and purified using standard Strep-Tactin
25 chromatography. Each reporter demonstrated appropriate spectral properties (Fig.
26 S11a) and migrated in SDS PAGE gels to appropriate positions related to their
27 predicted molecular masses (Fig. S11b).

1 **Discussion**

2 Heterologous gene expression systems for eukaryotic microalgae are important tools
3 for recombinant protein production and are also the prerequisite to adapt cellular
4 pathways for metabolic engineering. However, even for the most established model
5 microalgal species, *C. reinhardtii*, the availability of efficient and versatile vector
6 systems for nuclear transgene expression is still very limited.

7 Previously, vector pSI103 was constructed by combining the gene expression cassette
8 of the *Sh ble* resistance gene (Lumbreras et al. 1998), containing the RBCS2
9 promoter, with the HSP70A promoter, RBCS2 intron 1 (HSP70A-RBCS2-i1), and the
10 *aphVIII* gene to confer resistance to paromomycin (Sizova et al. 2001). This antibiotic
11 resistance cassette was used as the basis of subsequent vector pJR38 (Neupert et al.
12 2009), which expanded expression capabilities by using the *Chlamydomonas* PsaD
13 promoter and 3' UTR to express a codon optimized green fluorescent protein. pJR38
14 was then used as a backbone to create the vector pcCAGLucLpIBP (Lauersen et al.
15 2013a; Lauersen et al. 2013b) which was used to express a *Gaussia* luciferase-*Lolium*
16 *perenne* ice binding protein fusion targeted for secretion from *C. reinhardtii*.

17 Given the success of the HSP70A-RBCS2-i1 fusion promoter for the reliable
18 expression of the *gLuc-LpIBP* fusion in our laboratory, we chose to expand on the
19 concept of the synthetically derived pcCAGLucLpIBP vector (Lauersen et al. 2013a)
20 to establish a platform for standardized vector design and future nuclear gene
21 expression studies from *Chlamydomonas*. The heterogeneity of available vectors for
22 nuclear gene expression currently published for *C. reinhardtii* is the result of
23 independent development of expression systems in multiple research groups over the
24 past decades of research with this species. Our goal with the pOptimized vector
25 system was to create a platform for standardization and expansion of gene expression
26 possibilities, while making the system accessible for general research of this model
27 green alga. The pOptimized vector allows multiple cloning of any element through
28 digestion with strategically designed restriction enzyme recognition sequences
29 (Fig. 1).

30 In order to achieve a system which entirely lacked redundant restriction sites, the gene
31 of interest and antibiotic expression cassettes were re-designed *in silico*. In addition, a
32 series of reporters, including a luciferase, four fluorescent proteins, and two antibiotic
33 resistance markers were also built *de novo* to fit into this vector concept (Fig. 1, 2).

1 Each reporter was tested for reliable expression from vectors conferring resistance to
2 either antibiotic and demonstrated multiple subcellular localizations (Fig. 3-6).
3 Although the work presented here has been conducted with the mutant strain UVM4,
4 which is generally prone to high recombinant gene expression (Neupert et al. 2009),
5 we have observed the reporter constructs to be reliable in other *C. reinhardtii* strains
6 as well, including CC 406, and CC 1883 (not shown).
7 As this vector system is expandable, we suggest that future synthetic reporter
8 sequences could be designed to fit within the concept, avoiding redundant restriction
9 sites in order to allow the greatest flexibility. We also propose a nomenclature
10 structure for variants of the pOptimized vectors, that each element different from the
11 templates provided be separated by an underscore. For example, the paromomycin
12 resistant cytosolic mVenus and secreted (*c*CA tagged) mVenus vectors are given the
13 names pOpt_mVenus_Paro and pOpt_cCA_mVenus_Paro, respectively, see Table S4
14 for other examples.
15 The availability of protein standards is important for quantification of recombinant
16 protein production. A recombinant secreted *g*Luc is available commercially, however,
17 no commercially available standards for the specific fluorescent variants are available.
18 Therefore, we also constructed *E. coli* codon optimized variants of these four
19 reporters, adapted to contain the additional amino acids present from the restriction
20 sites of the pOptimized open reading frame as well as the StrepII tag for ease of
21 isolation and use as reliable standards for future investigations (Fig. S11). It should be
22 noted that the 3' UTR of the GOI expression cassette of the pOptimized vectors
23 contain a sequence for an optional StrepII tag addition, which may function as a
24 Western blotting target when specific antibodies are not available. We chose these
25 fluorescent reporters as bright, monomeric, photostable representatives of each class
26 of FP which can also function as FRET pairs for future investigations using this
27 vector concept (Kremers et al. 2006; Markwardt et al. 2011; Lam et al. 2012).
28 The bright nature of the reporter proteins allowed practical screening using
29 fluorescence or bioluminescence signals at the plate level (Fig. 3a-c). Although there
30 is a strong variability in transformant reporter expression throughout a mutant
31 population (e.g. Fig. 3a,b), the transformants with the strongest protein expression
32 could be easily identified at plate level, avoiding tedious scale up and blotting.
33 Cytosolic expression of mVenus was robust with accumulation up to ~0.25% TSP
34 (Fig. S3c) indicating the reliability of the vector concept for reporter protein

1 expression. For constructs with subcellular localization of mVenus, colonies could
2 also be readily screened at plate level (not shown), owing to a lack of background
3 fluorescence from chlorophyll at respective excitation and emission wavelengths
4 (Fig. 3c).

5 It has been recently described that nuclear, chloroplast, and mitochondrial localization
6 of fluorescent reporters can be reliably accomplished with a 2X repeat of the SV40
7 NLS, the PsaD N-terminal peptide, and AtpA N-terminal targeting peptide
8 respectively (Rasala et al. 2014). Here, we expand on this concept with demonstration
9 of microbody targeting with two known PTS1-like sequences, confirming recently
10 published results (Hayashi and Shinozaki 2012) as well as presenting a novel, smaller,
11 synthetic PTS1-like signal with similar targeting activity (Fig. 5e-g, and 6b). The
12 reliability of targeting to these structures with multiple signals will likely aid in future
13 proteomic analyses of these microbodies and their constituent components. It is
14 unclear at this time what potential biotechnological relevance these cellular structures
15 have, however, as isolated subcellular compartments, they may present interesting
16 applications for target protein sequestration.

17 A single copy of the SV40 NLS is able to reliably target mVenus to the nucleus in our
18 analyses (Fig. 5h). Since this sequence must be exposed for its proper recognition
19 (Kalderon et al. 1984), it is understandable why a 2X repeat may be of value for some
20 protein expression targets, as has been recently demonstrated (Rasala et al. 2014).

21 We demonstrate reliable tagging of the Rubisco small subunit 1 with mVenus and that
22 this protein fusion accumulates in the pyrenoid (Fig. 5d, 6c). To our knowledge, this
23 is the first demonstration of reporter protein localization in the pyrenoid of *C.*
24 *reinhardtii*. Although the 45 aa N-terminal targeting peptide of RBCS1 could localize
25 mVenus into the chloroplast stroma, extensions of the RBCS1 N-terminal plastid
26 targeting sequence resulted in unstable mVenus, both the fluorescent signal and
27 protein quantity as analyzed by Western blot were unusable for analysis from these
28 transformants (Fig. S6b and not shown). This instability likely indicates that further
29 RBCS1 sequence elements were required for functional assembly with the larger
30 Rubisco complex and subsequent pyrenoid localization. It is likely that these fusions
31 were recognized as misfolded RBCS1 and degraded within the stroma. It was
32 proposed by Genkov et al. (2010) that *C. reinhardtii* Rubisco subunits may contain an
33 as-of-yet unknown pyrenoid localization signal, as hybrid Rubisco small subunits
34 containing plant homologues did not result in pyrenoid formation in complemented

1 photosynthetic *C. reinhardtii* mutants, these mutants also exhibited reduced
2 photosynthetic growth (Genkov et al. 2010). It is still unclear whether the RBCS1
3 contains such a localization signal or, rather, its incorporation into the larger Rubisco
4 complex and secondary transport is responsible for this localization. Further
5 truncation experiments and reporter-Rubisco fusions with the pOptimized vector will
6 likely contribute to elucidating this unique feature. Expression of the cDNA sequence
7 alone for RBCS1 was likely assisted by the RBCS2 introns 1 and 2 surrounding its
8 sequence in the pOpt_mVenus_Paro vector, demonstrating the value of the use of
9 RBCS2i2 within the reporter sequences.

10 Each subcellular localized construct was contrasted with a separately localized
11 reporter protein by a second transformation with pOpt_mCerulean3_Hyg or *cCA*
12 variant vector (Fig. 6). The availability of a secondary antibiotic selection marker
13 allows for a second GOI to be transformed and independently screened for expression
14 from a strain already known to express another reporter. Since reporters expressed
15 from either antibiotic resistance vector, resulted in ~30% of colonies with
16 fluorescence signals detectable at the plate level, we found that secondary
17 transformation, rather than co-transformation, allowed practical pre-screening of
18 strong reporter expression, followed by a subsequent screen for secondary reporter
19 expression. It has been recently published that the incorporation of viral 2A self-
20 cleaving peptides allows multiple gene expression from a single GOI expression
21 cassette (Rasala et al. 2012; Rasala et al. 2014). Recovery on media containing a
22 member of the bleomycin antibiotic family resulted in selection for those
23 transformants which express high titers of target GOI. This strategy is effective for
24 expression of target proteins which may be otherwise difficult to detect, as surviving
25 colonies likely express the target protein. It has also been shown that multiple gene
26 traits can be combined through mating of heterologous marker expressing strains to
27 yield progeny expressing up to four separately targeted reporters (Rasala et al. 2014).
28 The pOptimized vector concept naturally complements these expression strategies and
29 presents multiple new options for combined GOI expression. In this work, multiple
30 gene traits were combined without mating, rather by double transformation and
31 selection mediated by separate antibiotic resistance markers. The ability to use
32 multiple different antibiotic selection methods, as well as separate markers for gene
33 expression, expands the possibilities of complex metabolic engineering from *C.*
34 *reinhardtii*. Indeed, 2A peptide-bleomycin resistance fusions, devoid of redundant

1 restriction sites, will likely be among the first new genetic elements designed for the
2 pOptimized system.

3 In summary, with this work we introduce a new and versatile vector toolkit which can
4 be of great use for projects involving heterologous gene expression in the model green
5 microalga *C. reinhardtii*. All reporter sequences and basic vectors have been
6 deposited at the publicly accessible Chlamydomonas Resource Center
7 (<http://chlamycollection.org/>) and, therefore, will be freely available for future
8 research projects.

9

1 **Acknowledgements**

2 The authors would like to acknowledge the CLIB Graduate Cluster Industrial
3 Biotechnology (Federal Ministry of Science & Technology North Rhine Westphalia,
4 Germany (to K.J.L.)), The authors would also like to express thanks to Martina
5 Lummer, Prof. Dr. Thorsten Seidel, Prof. Dr. Karsten Niehaus, Dr. Darius Widera for
6 assistance with fluorescence protein analysis and microscopy, Prof. Dr. Ralph Bock
7 for strain UVM4, as well as to Jan Schwarzhans and Prof. Dr. Karl Friehs for
8 providing the GFP standards. The authors declare that they have no conflict of
9 interest.

10

11 **Conflict of Interest**

12 The Authors declare that they have no conflict of interest.

1 References

- 2
3 2 Apt KE, Kroth-Pancic PG, Grossman AR (1996) Stable nuclear transformation of the
4 3 diatom *Phaeodactylum tricorutum*. Mol Gen Genet 252(5):572–579
5
6 4 Bateman JM, Purton S (2000) Tools for chloroplast transformation in
7 5 *Chlamydomonas*: expression vectors and a new dominant selectable marker. Mol
8 6 Genet Genomics 263:404–410
9
10 7 Berthold P, Schmitt R, Mages W (2002) An engineered *Streptomyces hygroscopicus*
11 8 aph 7" gene mediates dominant resistance against hygromycin B in
12 9 *Chlamydomonas reinhardtii*. Protist 153(4):401–412
13
14 10 Bogen C, Klassen V, Wichmann J, La Russa M, Doebbe A, Grundmann M, Uronen P,
15 11 Kruse O, Mussnug JH (2013) Identification of *Monoraphidium contortum* as a
16 12 promising species for liquid biofuel production. Bioresour Technol 133:622–626
17
18 13 Chen H-C, Melis A (2013) Marker-free genetic engineering of the chloroplast in the
19 14 green microalga *Chlamydomonas reinhardtii*. Plant Biotechnol J 11(7):818–828
20
21 15 Cordero BF, Couso I, León R, Rodríguez H, Vargas MA (2011) Enhancement of
22 16 carotenoids biosynthesis in *Chlamydomonas reinhardtii* by nuclear
23 17 transformation using a phytoene synthase gene isolated from *Chlorella*
24 18 *zofingiensis*. Appl Microbiol Biotechnol 91:341–51
25
26 19 Draaisma RB, Wijffels RH, Slegers PME, Brentner LB, Roy A, Barbosa MJ (2013)
27 20 Food commodities from microalgae. Curr Opin Biotechnol 24(2):169–177
28
29 21 Eichler-Stahlberg A, Weisheit W, Ruecker O, Heitzer M (2009) Strategies to facilitate
30 22 transgene expression in *Chlamydomonas reinhardtii*. Planta 229(4):873–883
31
32 23 Feng S, Feng W, Zhao L, Gu H, Li Q, Shi K, Guo S, Zhang N (2014a) Preparation of
33 24 transgenic *Dunaliella salina* for immunization against white spot syndrome virus
34 25 in crayfish. Arch Virol 159(3):519–525
35
36 26 Feng S, Li X, Xu Z, Qi J (2014b) *Dunaliella salina* as a novel host for the production
37 27 of recombinant proteins. Appl Microbiol Biotechnol 98(10):4293–300
38
39 28 Fischer N, Rochaix J-D (2001) The flanking regions of PsaD drive efficient gene
40 29 expression in the nucleus of the green alga *Chlamydomonas reinhardtii*. Mol
41 30 Genet Genomics 265(5):888–894
42
43 31 Fuhrmann M, Hausherr A, Ferbitz L, Schödl T, Heitzer M, Hegemann P (2004)
44 32 Monitoring dynamic expression of nuclear genes in *Chlamydomonas reinhardtii*
45 33 by using a synthetic luciferase reporter gene. Plant Mol Biol 55(6):869–881
46
47 34 Fuhrmann M, Oertel W, Hegemann P (1999) A synthetic gene coding for the green
48 35 fluorescent protein (GFP) is a versatile reporter in *Chlamydomonas reinhardtii*.
49 36 Plant J 19(3):353–361
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

- 1 Genkov T, Meyer M, Griffiths H, Spreitzer RJ (2010) Functional hybrid rubisco
2 enzymes with plant small subunits and algal large subunits: engineered rbcS
3 cDNA for expression in *Chlamydomonas*. J Biol Chem 285(26):19833–19841
4
5 4 Glanz S, Bunse A, Wimbert A, Balczun C, Kück U (2006) A nucleosome assembly
6 protein-like polypeptide binds to chloroplast group II intron RNA in
7 *Chlamydomonas reinhardtii*. Nucleic Acids Res 34(18):5337–51
8
9 7 Glanz S, Jacobs J, Kock V, Mishra A, Kück U (2012) Raa4 is a trans-splicing factor
10 that specifically binds chloroplast tscA intron RNA. Plant J 69(3):421–31
11
12 9 Goldschmidt-Clermont M, Rahire M (1986) Sequence, evolution and differential
13 expression of the two genes encoding variant small subunits of ribulose
14 bisphosphate carboxylase/oxygenase in *Chlamydomonas reinhardtii*. J Mol Biol
15 191(3):421–432
16
17 13 Hallmann A (2007) Algal transgenics and biotechnology. Transgenic Plant J 1(1):81–
18 98
19
20 15 Hayashi Y, Shinozaki A (2012) Visualization of microbodies in *Chlamydomonas*
21 *reinhardtii*. J Plant Res 125(4):579–586
22
23 17 Kalderon D, Roberts BL, Richardson WD, Smith AE (1984) A short amino acid
24 sequence able to specify nuclear location. Cell 39(2):499–509
25
26 19 Kilian O, Benemann CSE, Niyogi KK, Vick B (2011) High-efficiency homologous
27 recombination in the oil-producing alga *Nannochloropsis* sp. Proc Natl Acad Sci
28 U S A 108:21265–21269
29
30 22 Kindle KL (1990) High-frequency nuclear transformation of *Chlamydomonas*
31 *reinhardtii*. Proc Natl Acad Sci USA 87(3):1228–1232
32
33 24 Kindle KL, Schnell RA, Fernández E, Lefebvre PA (1989) Stable nuclear
34 transformation of *Chlamydomonas* using the *Chlamydomonas* gene for nitrate
35 reductase. J Cell Biol 109(6):2589–2601
36
37 27 Kremers G-J, Goedhart J, van Munster EB, Gadella TWJ (2006) Cyan and yellow
38 super fluorescent proteins with improved brightness, protein folding, and FRET
39 Förster radius. Biochemistry 45(21):6570–6580
40
41 30 Lam AJ, St-Pierre F, Gong Y, Marshall JD, Cranfill PJ, Baird MA, McKeown MR,
42 Wiedenmann J, Davidson MW, Schnitzer MJ, Tsien RY, Lin MZ (2012)
43 Improving FRET dynamic range with bright green and red fluorescent proteins.
44 Nat Methods 9(10):1005–1012
45
46 34 Lauersen KJ, Berger H, Mussgnug JH, Kruse O (2013a) Efficient recombinant protein
47 production and secretion from nuclear transgenes in *Chlamydomonas reinhardtii*.
48 J Biotechnol 167(2):101–110
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

- 1 Lauersen KJ, Vanderveer TL, Berger H, Kaluza I, Mussnug JH, Walker VK, Kruse
2 O (2013b) Ice recrystallization inhibition mediated by a nuclear-expressed and -
3 secreted recombinant ice-binding protein in the microalga *Chlamydomonas*
4 *reinhardtii*. *Appl Microbiol Biotechnol* 97(22):9763–9772
- 5
6 León-Bañares R, González-Ballester D, Galván A, Fernández E (2004) Transgenic
7 microalgae as green cell-factories. *Trends Biotechnol* 22(1):45–52
- 8
9
10 Liu B, Vieler A, Li C, Jones a D, Benning C (2013) Triacylglycerol profiling of
11 microalgae *Chlamydomonas reinhardtii* and *Nannochloropsis oceanica*.
12 *Bioresour Technol* 146:310–316
- 13
14
15 Lumbreras V, Stevens R. D, Purton S (1998) Efficient foreign gene expression in
16 *Chlamydomonas reinhardtii* mediated by an endogenous intron. *Plant J*
17 14(4):441–447
- 18
19
20 Markwardt ML, Kremers G-J, Kraft CA, Ray K, Cranfill PJC, Wilson KA, Day RN,
21 Wachter RM, Davidson MW, Rizzo MA (2011) An improved cerulean
22 fluorescent protein with enhanced brightness and reduced reversible
23 photoswitching. *PLoS One* 6(3):e17896
- 24
25
26 Miyahara M, Aoi M, Inoue-Kashino N, Kashino Y, Ifuku K (2014) Highly efficient
27 transformation of the diatom *Phaeodactylum tricorutum* by multi-pulse
28 electroporation. *Biosci Biotechnol Biochem* 77(4):874–876
- 29
30
31 Neupert J, Karcher D, Bock R (2009) Generation of *Chlamydomonas* strains that
32 efficiently express nuclear transgenes. *Plant J* 57(6):1140–1150
- 33
34
35 Radakovits R, Jinkerson RE, Fuerstenberg SI, Tae H, Settlege RE, Boore JL,
36 Posewitz MC (2012) Draft genome sequence and genetic transformation of the
37 oleaginous alga *Nannochloropsis gaditana*. *Nat Commun* 3:686
- 38
39
40 Rasala BA, Barrera DJ, Ng J, Plucinak TM, Rosenberg JN, Weeks DP, Oyler GA,
41 Peterson TC, Haerizadeh F, Mayfield SP (2010) Production of therapeutic
42 proteins in algae, analysis of expression of seven human proteins in the
43 chloroplast of *Chlamydomonas reinhardtii*. *Plant Biotechnol J* 8(6):719–733
- 44
45
46 Rasala BA, Barrera DJ, Ng J, Plucinak TM, Rosenberg JN, Weeks DP, Oyler GA,
47 Peterson TC, Haerizadeh F, Mayfield SP (2013) Expanding the spectral palette
48 of fluorescent proteins for the green microalga *Chlamydomonas reinhardtii*.
49 *Plant J* 74(4):545–56
- 50
51
52 Rasala BA, Chao S-S, Pier M, Barrera DJ, Mayfield SP (2014) Enhanced genetic
53 tools for engineering multigene traits into green algae. *PLoS One* 9(4):e94028
- 54
55
56 Rasala BA, Lee PA, Shen Z, Briggs SP, Mendez M, Mayfield SP (2012) Robust
57 expression and secretion of Xylanase1 in *Chlamydomonas reinhardtii* by fusion
58 to a selection gene and processing with the FMDV 2A peptide. *PLoS One*
59 7(8):e43349
- 60
61
62
63
64
65

- 1 Rasala BA, Mayfield SP (2014) Photosynthetic biomanufacturing in green algae;
2 production of recombinant proteins for industrial, nutritional, and medical uses.
3 Photosynth Res. doi: 10.1007/s11120-014-9994-7
- 4 Rasala BA, Muto M, Sullivan J, Mayfield SP (2011) Improved heterologous protein
5 expression in the chloroplast of *Chlamydomonas reinhardtii* through promoter
6 and 5' untranslated region optimization. Plant Biotechnol J 9(6):674–683
- 7 Remacle C, Cardol P, Coosemans N, Gaisne M, Bonnefoy N (2006) High-efficiency
8 biolistic transformation of *Chlamydomonas* mitochondria can be used to insert
9 mutations in complex I genes. Proc Natl Acad Sci U S A 103(12):4771–4776
- 10 Rochaix JD (1995) *Chlamydomonas reinhardtii* as the photosynthetic yeast. Annu
11 Rev Genet 29:209–230
- 12 Ruecker O, Zillner K, Groebner-Ferreira R, Heitzer M (2008) Gaussia-luciferase as a
13 sensitive reporter gene for monitoring promoter activity in the nucleus of the
14 green alga *Chlamydomonas reinhardtii*. Mol Genet Genomics 280(2):153–162
- 15 Schroda M, Blöcker D, Beck CF (2000) The HSP70A promoter as a tool for the
16 improved expression of transgenes in *Chlamydomonas*. Plant J 21(2):121–131
- 17 Schwarz C, Elles I, Kortmann J, Piotrowski M, Nickelsen J (2007) Synthesis of the
18 D2 protein of photosystem II in *Chlamydomonas* is controlled by a high
19 molecular mass complex containing the RNA stabilization factor Nac2 and the
20 translational activator RBP40. Plant Cell 19(11):3627–3639
- 21 Shao N, Bock R (2008) A codon-optimized luciferase from *Gaussia princeps*
22 facilitates the in vivo monitoring of gene expression in the model alga
23 *Chlamydomonas reinhardtii*. Curr Genet 53(6):381–388
- 24 Sizova I, Fuhrmann M, Hegemann P (2001) A *Streptomyces rimosus aphVIII* gene
25 coding for a new type phosphotransferase provides stable antibiotic resistance to
26 *Chlamydomonas reinhardtii*. Gene 277(1-2):221–229
- 27 Sizova IA, Lapina T V., Frolova ON, Alexandrova NN, Akopiants KE, Danilenko VN
28 (1996) Stable nuclear transformation of *Chlamydomonas reinhardtii* with a
29 *Streptomyces rimosus* gene as the selective marker. Gene 181(1-2):13–18
- 30 Stephens E, Ross IL, Mussgnug JH, Wagner LD, Borowitzka MA, Posten C, Kruse
31 O, Hankamer B (2010) Future prospects of microalgal biofuel production
32 systems. Trends Plant Sci 15(10):554–564
- 33 Verhaegen M, Christopoulos TK (2002) Recombinant Gaussia luciferase.
34 Overexpression, purification, and analytical application of a bioluminescent
35 reporter for DNA hybridization. Anal Chem 74(17):4378–4385
- 36 Vieler A, Wu G, Tsai C-H, Bullard B, Cornish AJ, Harvey C, Reza I-B, Thornburg C,
37 Achawanantakun R, Buehl CJ, Campbell MS, Cavalier D, Childs KL, Clark TJ,
38 Deshpande R, Erickson E, Armenia Ferguson A, Handee W, Kong Q, Li X, Liu

1 B, Lundback S, Peng C, Roston R L, Sanjaya, Simpson JP, Terbush A,
2 Warakanont J, Zäuner S, Farre EM, Hegg EL, Jiang N, Kuo M-H, Lu Y, Niyogi
3 KK, Ohlrogge J, Osteryoung KW, Shachar-Hill Y, Sears BB, Sun Y, Takahashi
4 H, Yandell M, Shiu S-H, Benning C (2012) Genome, functional gene annotation,
5 and nuclear transformation of the heterokont oleaginous alga *Nannochloropsis*
6 *oceanica* CCMP1779. PLoS Genet 8(11):e1003064

7
8
9 7 Wijffels RH, Kruse O, Hellingwerf KJ (2013) Potential of industrial biotechnology
10 8 with cyanobacteria and eukaryotic microalgae. Curr Opin Biotechnol 24(3):405–
11 9 413

12
13 10
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

1 **Figure Captions**

2 **Fig. 1** The pOptimized vector concept, markers, and reporters. Each element of the
3 vector was designed to be surrounded by unique restriction digest recognition
4 sequences, allowing simple removal and opening of the vector backbone for future
5 marker, reporter, or regulatory element investigations. Codon optimized and RBCS2i2
6 (i2) containing markers were generated to match restriction sites of the vector
7 backbone. Multiple digestion sites on either side of the reporter allow for rapid
8 development of fusion protein targets and targeting peptides. Paromomycin and
9 hygromycin B resistance markers were also developed as single unit modules. H: *C.*
10 *reinhardtii* heat shock 70A promoter, R: Rubisco small subunit 2 (RBCS2) promoter,
11 i/i2: RBCS2 intron 1 or 2, 3': RBCS2 3' untranslated region, StrepII-TAA: StrepII
12 affinity tag and stop codon (WSHPQFEK*), MCS: multiple cloning site.

13
14 **Fig. 2** Schematic overview of all vectors used in this study. (a) The marker expression
15 cassette with hygromycin B resistance gene. (b) The initial paromomycin resistance
16 marker cassette, including promoters and 839 bp of *aphVIII* from vector pJR38. The
17 original 67 bp downstream of the TGA stop codon was not required for efficient
18 antibiotic resistance. (c) The pOptimized paromomycin resistance marker was
19 developed by amplification of the upstream region of *aphVIII* (λ : from original *MscI*
20 restriction insertion site in pJR38 to the start codon) until its TGA stop codon,
21 allowing insertion of this modular genetic element into the pOptimized vector
22 between *HindIII* and *XhoI*. (d) Depiction of element orientation of the cytosolic
23 expression reporters. (e) The secretion vectors, maintain the *C. reinhardtii* carbonic
24 anhydrase 1 (*cCA*) secretion signal between *NdeI* and *BglII* for efficient secretion of
25 all reporters. (f-h) The microbody targeting vectors, including the 3' addition of the
26 nucleotides encoding for the pumpkin and *C. reinhardtii* malate synthase PTS1
27 signals (PkPTS1 and MSPTS1), or a smaller, novel PTS1-like sequence (rPTS1),
28 respectively. The signal sequences were cloned between the *EcoRV* and *EcoRI*
29 restriction sites on the marker sequence coding for mVenus. (i) The nuclear
30 localization vector, including one repeat of the SV40 nuclear localization signal. (j,k)
31 N-terminal fusion of the 36 aa targeting peptide of *C. reinhardtii* PsaD and 46 aa
32 targeting peptide of *C. reinhardtii* mitochondrial AtpA to mVenus for chloroplast or
33 mitochondrial targeting, respectively. (l) Fusion construct of RBCS1 cDNA to

1 mVenus. (m-q) Truncations of various lengths of the RBCS1 N-terminal coding
2 sequence, arrows indicate the cleavage site of the predicted chloroplast targeting
3 peptide. Each construct contains sequences encoding for 10 aa past the previous
4 cleavage peptide (T1-5).

5
6
7
8
9 **Fig. 3** Plate-level detection of the intracellular and extracellular expression of the five
10 light-emission reporters used in this work. Strains expressing cytoplasmic and
11 secreted reporters (Fig. 2 vector constructs d and e) are shown. (a, b) gLuc reporter
12 proteins were detected by bioluminescence assay. Depending on the presence or
13 absence of the *cCA* secretion signal, gLuc dependent signals were detected as halos
14 outside of the cells (a) or within the cell colony (b). (c) (i) Fluorescence background
15 signals of parental strain with CFP, YFP, GFP, and DsRed filter settings, respectively.
16 (ii-v) mCerulean3, mVenus, Clover, and mRuby2 expression from respective
17 cytoplasmic pOptimized vectors in transformant colonies under the same filter
18 settings as in (i). (vi-ix) mCerulean3, mVenus, Clover, and mRuby2 expression from
19 respective pOptimized expression and secretion vectors under the same filter settings
20 as in (i). Fusion of with the *cCA* secretion signal resulted in reporter accumulation
21 outside of the cell colonies. (d) Secreted fluorescent protein detection in 10X
22 concentrated culture media from the *cCA* secretion strains (c, vi-ix) without further
23 purification. (i) Transmission light without excitation and emission filters is shown,
24 followed by the same filter settings as in (i) .

25
26
27
28
29 **Fig. 4** Detection of the intracellular fluorescence of reporter proteins accumulated in
30 the cytoplasm by scanning confocal laser microscopy. Transformants generated from
31 fluorescent reporter constructs described in Fig. 2d are depicted. All reporters,
32 mCerulean3, mVenus, Clover, and mRuby2 (a-d, respectively), demonstrated reliable
33 expression and were clearly visualized with respective excitation and emission
34 wavelengths (see table S5). The pattern of chlorophyll background fluorescence are
35 shown as red signals in the left column, specific fluorescence signals are shown in the
36 middle column as blue, bright yellow, green or orange signals, respectively, and an
37 overlay of the two signals is shown in the right column. Scale bars represent 5 μm .

38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65 **Fig. 5** Subcellular localization of mVenus reporter with peptide fusions. Chlorophyll
background fluorescence is used to orient the signals in the cells. (a) Parental

1 untransformed strain is imaged with the same laser intensities, demonstrating no
2 background signal in the yellow channel, is presented as a control. (b) Cytosolic
3 mVenus localization (vector construct depicted in Fig. 2d). (c) Remaining mVenus
4 signal from the secretory pathway of *cCA_mVenus* protein secretion strain (vector
5 construct as Fig. 2e). (d) RBCS1_mVenus fusion localized the YFP signal to the
6 chloroplast pyrenoid, (vector construct Fig. 2l). Vectors with pumpkin (e) and *C.*
7 *reinhardtii* (f) malate synthase PTS1 signals, as well as the novel PTS1-like signal (g)
8 in C-terminal fusion to mVenus exhibited fluorescent signal in spherical microbodies
9 throughout the cell (vector constructs Fig. 2f-h, respectively). (h) The mVenus_SV40
10 NLS construct (vector Fig. 2i) accumulated in the nucleus of *C. reinhardtii*. (i) 36 aa
11 PsaD N-terminal targeting peptide_mVenus fusion was detected as a diffuse signal
12 within the chloroplast stroma (vector Fig. 2j). (j) The 46 aa N-terminal AtpA targeting
13 peptide_mVenus fusion accumulated in the mitochondrial network and displays
14 reticular, strong fluorescence signals (vector Fig. 2k). Scale bars represent 5 μ m.

16 **Fig. 6** Double transformations using separate resistance markers demonstrating
17 expression and clear contrast of separate reporter localizations. Cell lines transformed
18 with pOpt_mCerulean3_Hyg expressing mCerulean3 (cyan) in the cytoplasm were
19 subsequently transformed with mVenus (yellow) reporter constructs. Chlorophyll
20 fluorescence is shown as red background signals. (a) Secondary transformation of
21 mCerulean3 expressing strain with pOpt_mVenus_Paro resulted in simultaneous
22 expression of both reporters in the cytoplasm and overlapping fluorescence signals.
23 (b) Spherical mVenus containing microbodies contrasts against cytosolic mCerulean3
24 and chlorophyll background signal (mVenus vector construct from Fig. 2h). (c)
25 RBCS1_mVenus pyrenoid localization contrasts against chlorophyll background
26 signal and cytosolic mCerulean3 (mVenus vector construct Fig. 2l). (d) mVenus
27 accumulation in the chloroplast stroma (construct Fig. 2j) contrasts against
28 chlorophyll background fluorescence and mCerulean3 spread throughout the
29 cytoplasm. (e) mVenus_SV40NLS signal (construct Fig. 2i) contrasts against the
30 comparably weak residual pOpt_cCA_mCerulean3_Hyg transformed strain and the
31 signal derived from mCerulean3 within the secretory pathway (construct as in Fig.
32 2e). Scale bars represent 5 μ m.

Figure1
[Click here to download high resolution image](#)

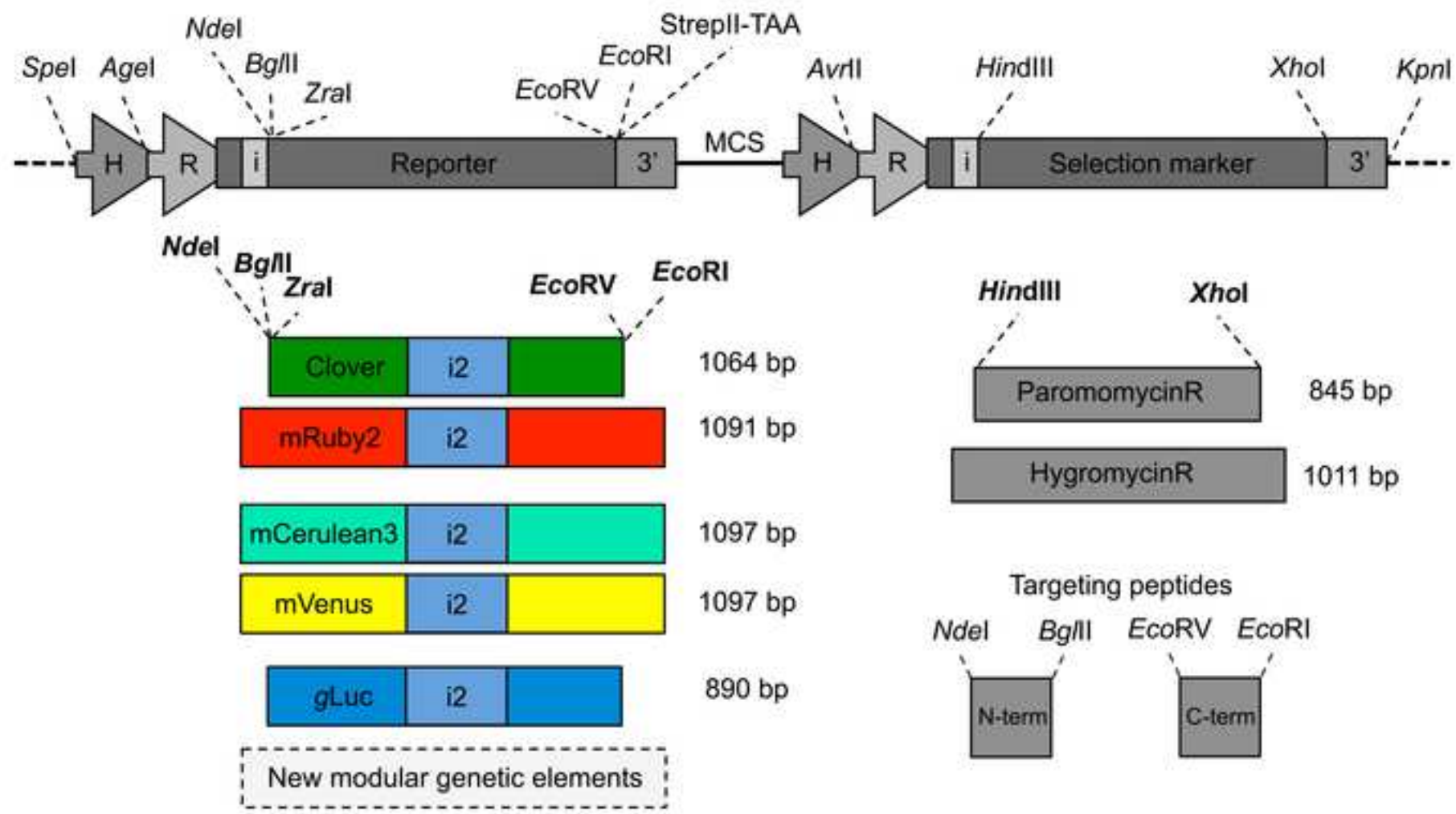


Figure 2

[Click here to download high resolution image](#)

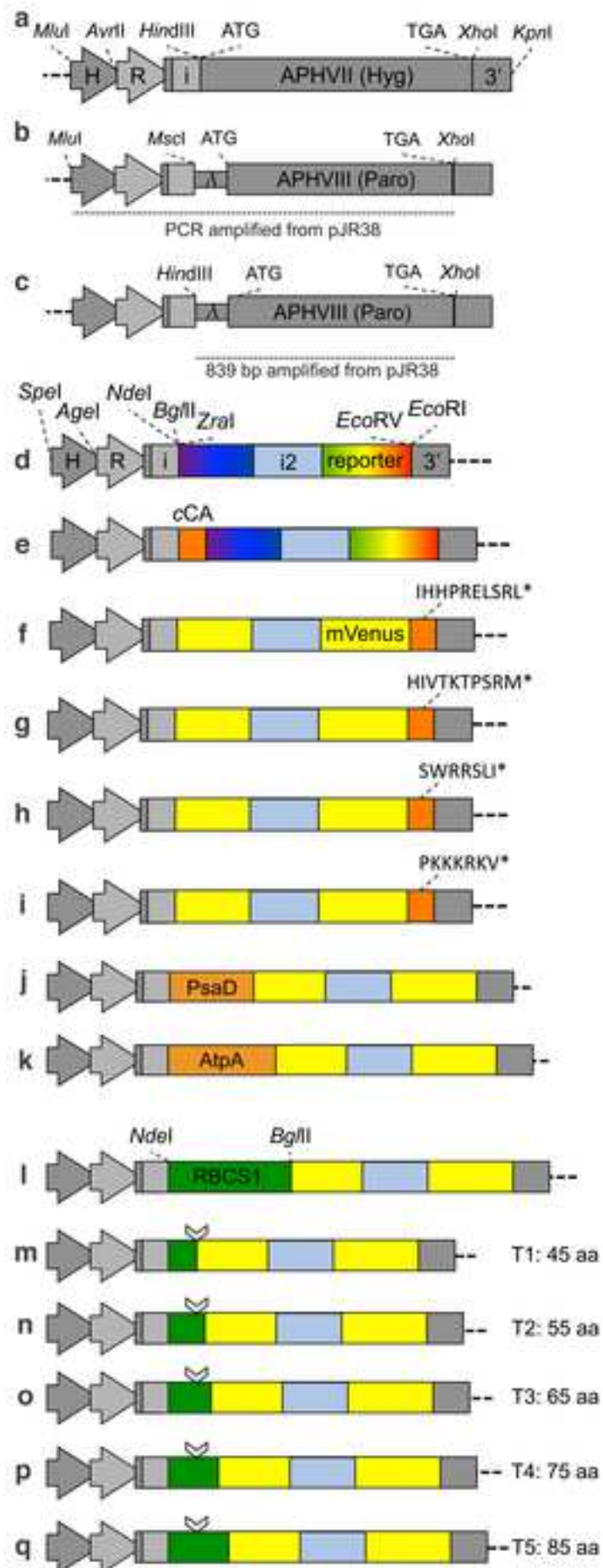


Figure 3
[Click here to download high resolution image](#)

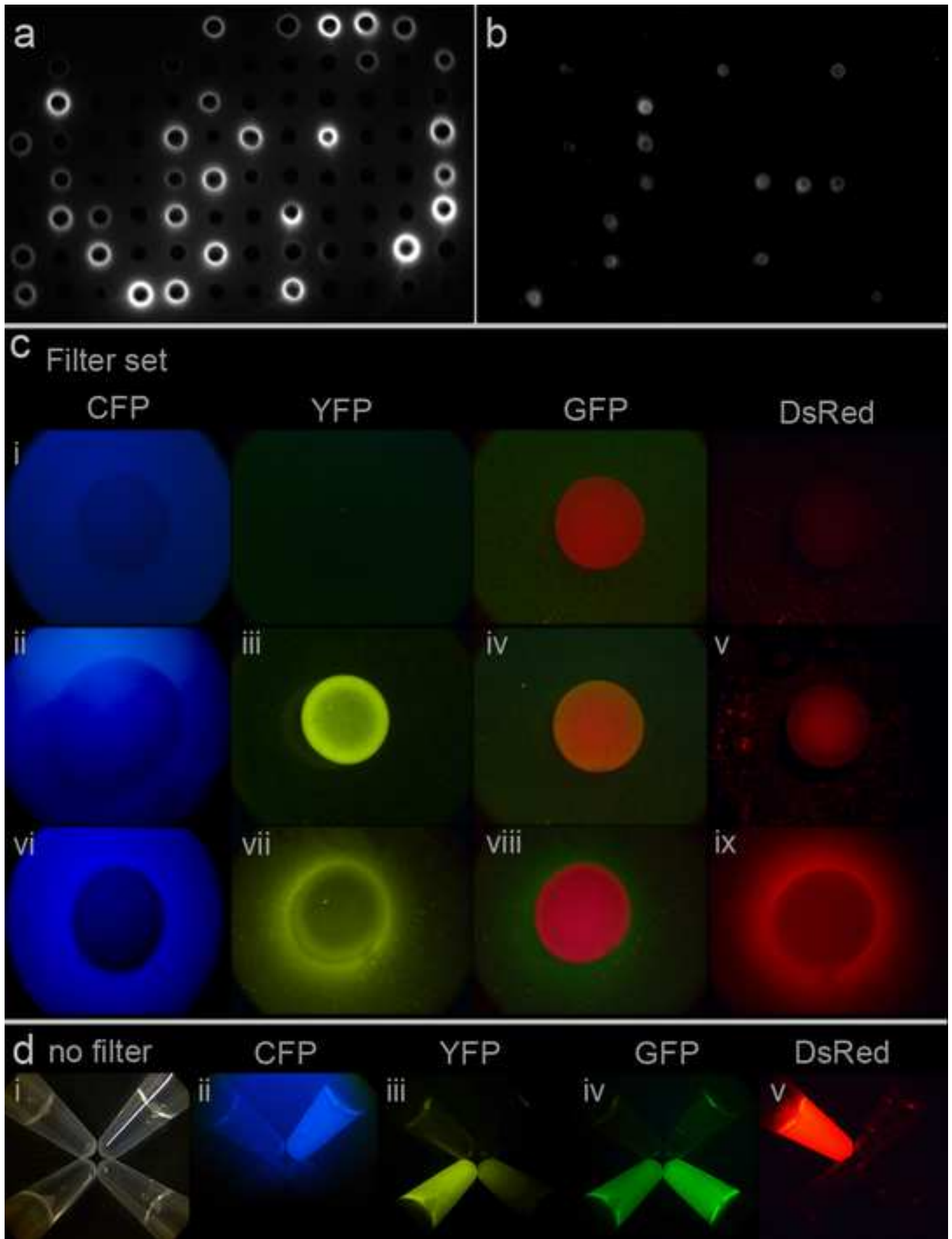


Figure4

[Click here to download high resolution image](#)

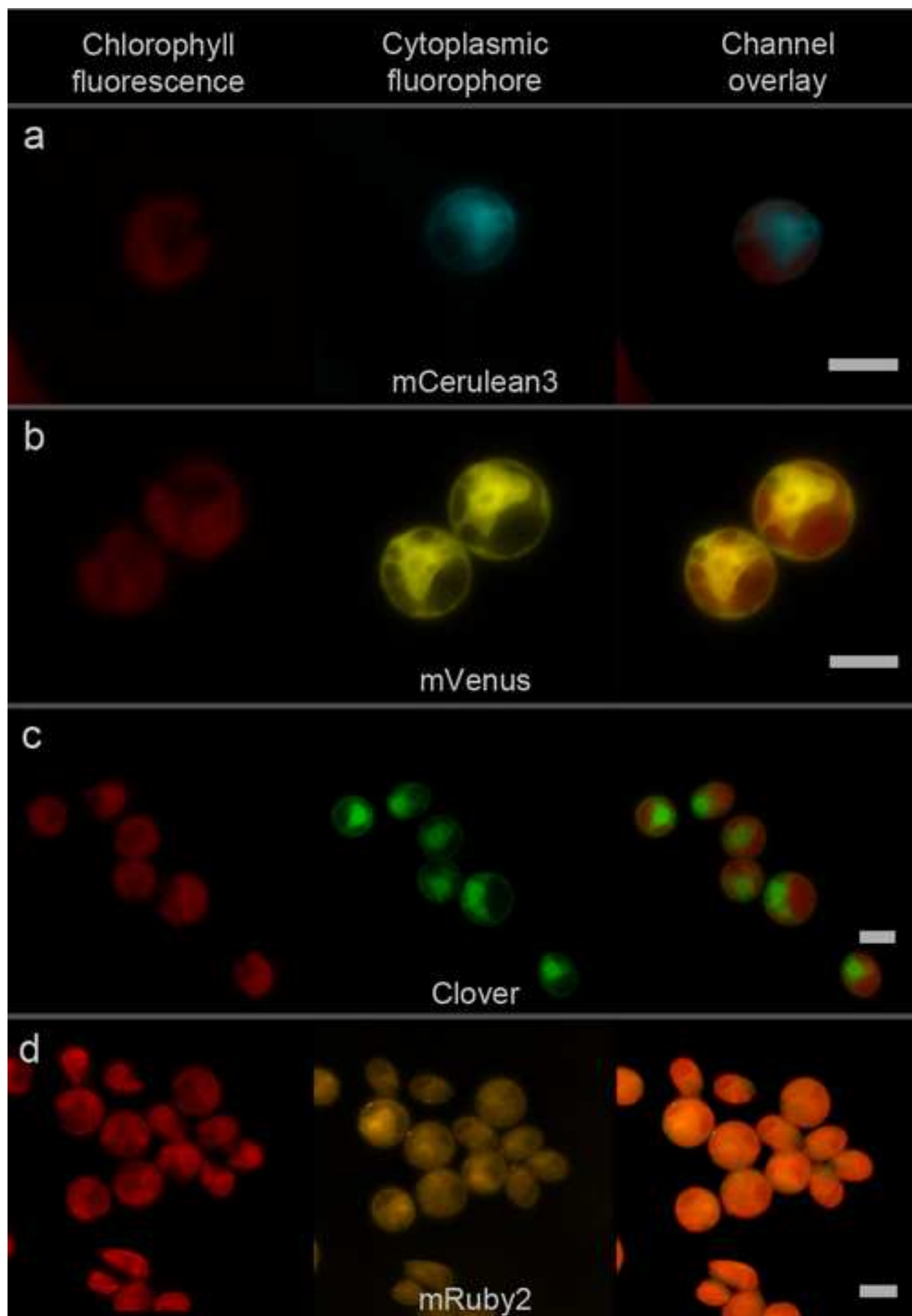


Figure5
[Click here to download high resolution image](#)

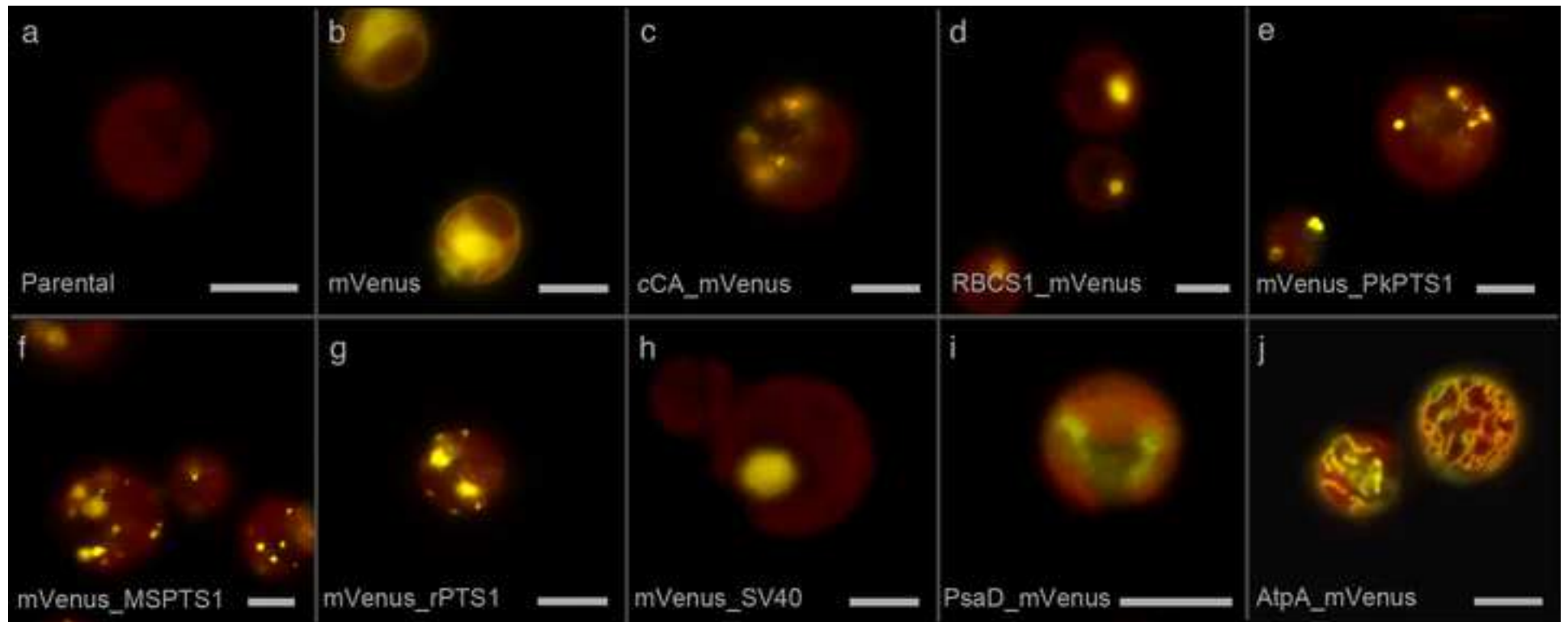
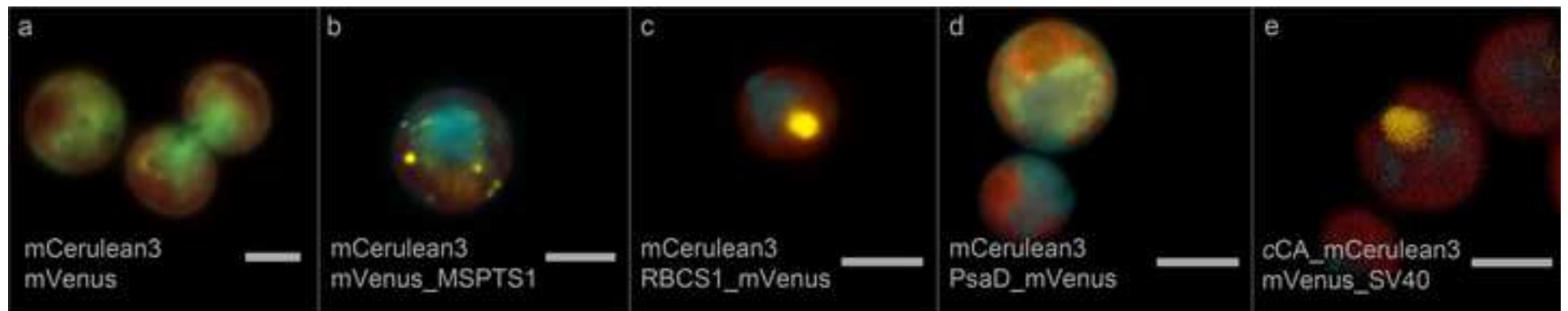


Figure6

[Click here to download high resolution image](#)



Supplementary Material

[Click here to download Supplementary Material: Online Resource Supporting Information.pdf](#)

Supplementary MovieS1

[Click here to download Supplementary Material: Movie S1.mpg](#)