the plant journal

SEB Society for Experimental Biology

The Plant Journal (2014) **78,** 1073–1083 doi: 10.1111/tpj.12530

TECHNICAL ADVANCE

Efficient gene targeting and removal of foreign DNA by homologous recombination in the picoeukaryote *Ostreococcus*

Jean-Claude Lozano[1,2,†], Philippe Schatt[1,2,†], Hugo Botebol[1,2], Valérie Vergé[1,3], Emmanuel Lesuisse[4], Stéphane Blain[1,2], Isabelle A. Carré[5] and François-Yves Bouget[1,2,*]

[1]*Sorbonne Universités, UPMC Univ Paris 06, UMR 7621, Laboratoire d'Océanographie Microbienne, Observatoire Océanologique, F-66650 Banyuls/mer, France,*

[2]*CNRS, UMR 7621, Laboratoire d'Océanographie Microbienne, Observatoire Océanologique, F-66650 Banyuls/mer, France,*

[3]*Unité Mixte de Service, UMS2348, F-66651 Banyuls/Mer, France,*

[4]*Lab Mitochondria Metals and Oxidative Stress, Institut Jacques Monod, CNRS-Université Paris Diderot, Paris, France, and*

[5]*School of Life Sciences, University of Warwick, Coventry, UK*

Received 18 July 2013; revised 14 February 2014; accepted 31 March 2014; published online 3 April 2014.

*For correspondence (e-mail fy.bouget@obs-banyuls.fr).

[†]These authors contributed equally.

SUMMARY

With fewer than 8000 genes and a minimalist cellular organization, the green picoalga *Ostreococcus tauri* is one of the simplest photosynthetic eukaryotes. *Ostreococcus tauri* contains many plant-specific genes but exhibits a very low gene redundancy. The haploid genome is extremely dense with few repeated sequences and rare transposons. Thanks to the implementation of genetic transformation and vectors for inducible overexpression/knockdown this picoeukaryotic alga has emerged in recent years as a model organism for functional genomics analyses and systems biology. Here we report the development of an efficient gene targeting technique which we use to knock out the *nitrate reductase* and *ferritin* genes and to knock in a luciferase reporter in frame to the ferritin native protein. Furthermore, we show that the frequency of insertion by homologous recombination is greatly enhanced when the transgene is designed to replace an existing genomic insertion. We propose that a natural mechanism based on homologous recombination may operate to remove inserted DNA sequences from the genome.

Keywords: *Ostreococcus*, homologous recombination, gene targeting, microalgae, genetic transformation, technical advance.

INTRODUCTION

Over the past 20 years the extensive development of genetic transformation technologies in model organisms has revealed different mechanisms of transgene integration. Gene targeting by homologous recombination (HR) is the method of choice for deleting a gene, introducing a selected mutation or fusing a tag to a protein. In bacteria and in a few eukaryotic model organisms, such as the yeast *Saccharomyces cerevisiae*, HR occurs preferentially over random integration of homologous sequences; however for most eukaryotes, transgene integration occurs almost exclusively in a random fashion.

Random insertion (RI) of transgenes and targeted insertion by HR are based on distinct mechanisms of DNA repair. Random insertion is based on non-homologous end joining (NHEJ) repair of double-strand breaks (DSBs) in DNA (Heyer *et al.*, 2010). Targeted insertion is mediated by the recombinase Rad51, which catalyzes the exchange of DNA strands between damaged DNA and intact, homologous DNA sequences (Shinohara *et al.*, 1992). This mechanism is conserved between bacteria and eukaryotes (Heyer *et al.*, 2010). The Spo11 protein mediates an additional mechanism of HR specific to eukaryotes, which allows the formation of crossovers between homologous chromosome pairs at meiosis to ensure their proper segregation (Keeney *et al.*, 1997). This is of crucial importance in allowing exchanges of

DNA between homologous chromosomes and the shuffling of genetic information.

Gene targeting by HR remains difficult in most photosynthetic eukaryotes, with the exception of mosses (Schaefer and Zryd, 1997). Most attempts to knock out genes by HR in algae and higher plants were designed to disrupt the *nitrate reductase* gene (*NR*). This strategy greatly simplified the identification of HR events, as knock-out mutants grew on ammonium but failed to grow in the presence of nitrate. Gene knock-out by HR in algae was first reported 20 years ago in the chlorophyte *Chlamydomonas*, but gene targeting by HR remains very difficult in this model organism (Sodeinde and Kindle, 1993; Zorin *et al.*, 2009). In contrast, efficient gene targeting has been reported for two unicellular algae: the thermophilic acidophilic red alga *Cyanidioschizon merolae* and the heterokontophyte *Nannochloropsis* sp. (Minoda *et al.*, 2004; Kilian *et al.*, 2011).

Ostreococcus tauri (Prasinophyceae) has recently emerged as model marine organism. This tiny unicellular alga has a minimalist organization that allows approaches such as whole cell imaging by electron cryotomography (Gan *et al.*, 2011). Its haploid genome is compact and gene dense, with very little genetic redundancy (Derelle *et al.*, 2006). Functional genetic analyses in this organism are facilitated by an efficient genetic transformation protocol. This enabled, for example, the study of phenotypes caused by constitutive or inducible overexpression or knockdown of gene expression, or the monitoring of gene activity by *in vivo* imaging of luciferase reporter constructs (Corellou *et al.*, 2009; Moulager *et al.*, 2010; Djouani-Tahri *et al.*, 2011b). These tools were used in our research to analyze the genetic circuits and light signaling pathways of the circadian clock (Heijde *et al.*, 2010; Monnier *et al.*, 2010; Djouani-Tahri *et al.*, 2011a; Pfeuty *et al.*, 2012) and to demonstrate the presence of a non-transcriptional circadian clock shared with human red blood cells (O'Neill *et al.*, 2011).

However, antisense silencing of gene expression proved difficult in *O. tauri.* In our experience, phenotypes were not always correlated with reduced transcript levels, and RNA interference (RNAi) approaches are unlikely to be successful because *O. tauri* lacks key components of the RNA-induced silencing complex such as Dicer and Argonaute. However, homologues of Rad51 and Spo11 were identified in this organism, suggesting the presence of HR mechanisms involved in DSB repair and meiosis (Derelle *et al.*, 2006). This suggested that targeted gene disruption by HR might be achievable.

In this paper we describe a method for efficient transgene insertion by HR. This was used to knock out two different genes, *NR* and *ferritin*, and in a knock-in experiment to insert a luciferase reporter gene in-frame at the *ferritin* locus. Interestingly the frequency of transgene insertion by HR varied between 1 and 100% depending on whether the construct was designed to replace wild-type (Wt) DNA sequences or to replace an existing transgene insertion. We propose that HR may normally operate in *Ostreococcus* to eliminate foreign DNA sequences from the genome.

RESULTS

Knock-out of the *NR* gene by homologous recombination

Ostreococcus tauri is normally able to grow on Artificial Seawater (ASW) containing either nitrate or ammonium as the sole source of nitrogen (nitrate-ASW and ammonium-ASW, respectively). However, loss of NR function in *NR knock-out* (NR^{KO}) lines should result in the loss of ability to grow on nitrate-ASW. This provided us with a simple and efficient method for identifying HR events.

Ostreococcus tauri is haploid, and *NR* is encoded by a single gene located on chromosome 10. We designed a disruption cassette containing about 2 kbp of sequence homologous to the *NR* locus. This was interrupted by the *KanMx* selection marker, which confers resistance to the antibiotic G418. The *Kan-Mx* marker was inserted either in a sense or an antisense orientation relative to the *NR* sequence to produce the *KanMx-s* and *KanMx-as* constructs (Figure 1a). This was used to determine whether the orientation of the *KanMx* gene relative to the *NR* gene affected the frequency of insertion by HR. These constructs were introduced into Wt cells using the standard electroporation protocol (Corellou *et al.*, 2009). Transformants were selected on semi-solid ammonium-ASW agarose plates containing 1 mg ml^{-1} G418. G418-resistant transformants were then tested for their ability to grow in nitrate-ASW. Clones that did not grow under these conditions identified putative HR events (Figure 1b).

Four G418-resistant *KanMx-s* transformants out of 193, and eight G418-resistant *KanMx-as* transformants out of 425 failed to grow in nitrate-ASW (Table 1). Cell growth of Wt and putative *NR* knock-out lines was monitored by flow cytometry (Figure 1c). Noticeably, Wt cells grew better in nitrate-ASW than in ammonium-ASW. Furthermore, *NR* knock-out lines reached higher concentrations in ammonium-ASW. These results suggest that *O. tauri* uses nitrate more efficiently than ammonium and that nitrate regulates the uptake and/or the assimilation of ammonium.

The first generations of *KanMx-as* and *KanMx-s* HR lines were denoted HR1 *KanMx-as* and HR1 *KanMx-s*, respectively. The nomenclature used to describe these lines and the lineages of subsequent HR lines are summarized in Figure 2(a). Polymerase chain reaction and Southern blot analysis were carried out to confirm the mechanism of transgene insertion in these putative HR lines (Figure 2, Table 1). The primers F1 and R1 were used to test for the presence of an intact *NR* locus. A 0.6 kbp fragment was amplified from Wt cells. The same fragment was amplified from all of the *KanMx-as* transformants that grew in

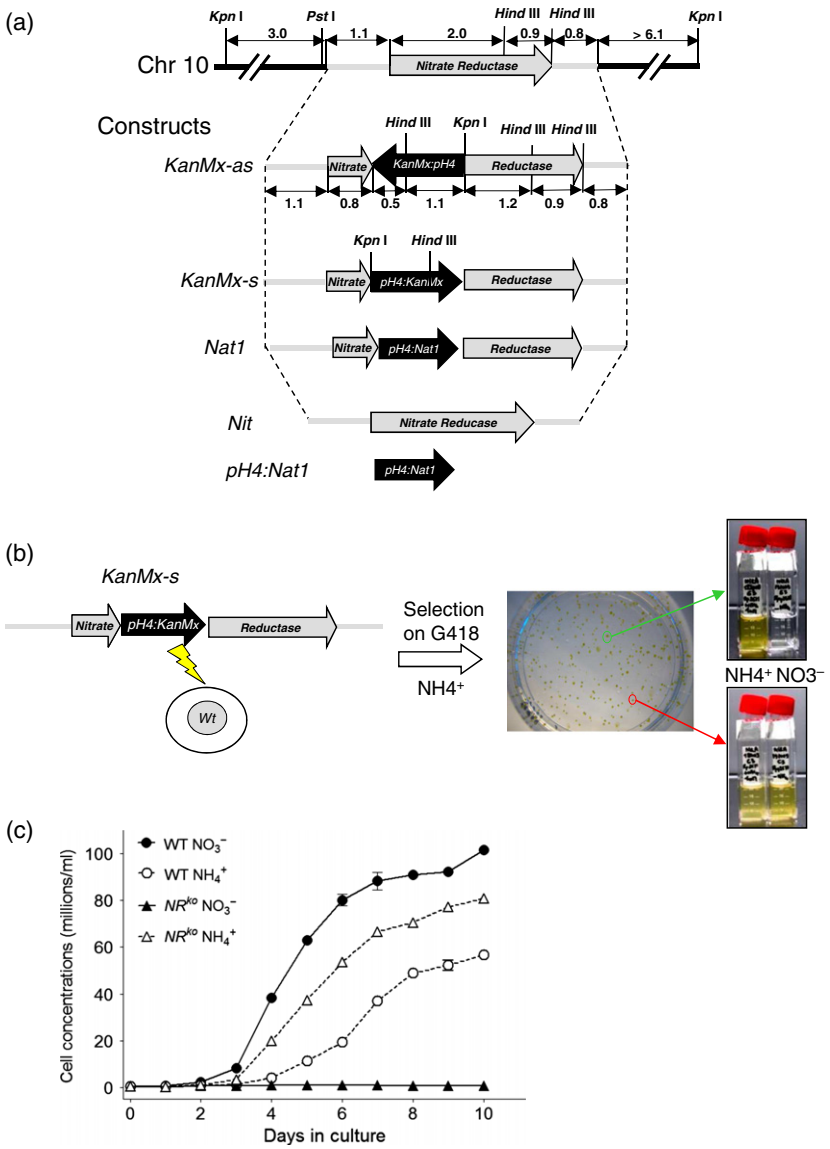

Figure 1. Strategy for the targeted disruption of the *nitrate reductase* (*NR*) gene in *Ostreococcus tauri*.

(a) Physical map of the *NR* locus and of the constructs used in the study. Homologous sequences are highlighted in grey and selection genes in black. The *pH4:Nat1* construct shares promoter and terminator sequences with the *pH4:KanMx* sequence of *KanMx-as* and *KanMx-s* constructs.

(b) Identification of a putative *NR* knock-out line. The *KanMx-as* construct was electroporated into *O. tauri* cells. Transformants resistant to G418 were selected in Artificial Seawater (ASW) agarose medium containing NH_4^+. Their growth was further tested in liquid ASW medium containing NO_3^- as the sole source of nitrogen. Most of the clones also grew on NO_3^- (bottom). However, those that grew on NH_4^+ but failed to grow on NO_3^- identified putative *NR* disruptants (top).

(c) Growth curves of wild type (Wt) and putative *NR* knock-out lines in NO_3^- and NH_4^+ ASW medium.

nitrate-ASW, confirming that they were the product of RI in the genome (Figure 2c). In contrast this fragment was not amplified from lines that failed to grow in nitrate-ASW, showing that the *NR* gene had been disrupted by insertion of the transgene. This was confirmed by the amplification of a 3-kbp product in all HR lines using a combination of primers, F2 in the selection gene and R2 in the *NR* sequence.

Southern blot analyses were carried out to check the number and pattern of transgene insertions in the HR lines. Genomic DNA was digested with *Kpn*I, which cleaved the insertion cassette immediately upstream of the *KanMx* resistance gene (Figure 2c). Hybridization of a *NR*-specific probe (P1) revealed a single band around 20 kbp in Wt and RI cells, corresponding to the intact *NR* locus. In contrast, a 6.5-kbp band was detected in HR lines, indicating the presence of the additional *Kpn*I site resulting from insertion of the *KanMx* gene at the *NR* locus. No additional band was detected, indicating that no additional insertions had occurred in the HR lines. A similar 6.5-kbp band was detected in HR lines using a probe specific to the *KanMx* transgene (P2) whereas a 15-kbp band was observed in the single RI line tested, corresponding to a RI of the transgene.

Similar experiments were performed in the lines resulting from the *KanMx-s* transformation (Figure S1 in Supporting Information). Polymerase chain reaction products were obtained at the expected size for HR, i.e. 3.2 and 2.2 kbp using (F3, R2) and (F1, R1) primer pairs, respectively. Hybridization of genomic DNA digested with *Kpn*I to the *NR* P1 probe revealed a band at a size (4.9 kbp) consistent with HR of the *KanMx-s* construct at the *NR* locus (Figure S1).

Table 1 Effect of construct orientation on the frequency of homologous recombination (HR) at the *NR* locus

	Transformants				
Construct (recipient Wt)	Selection	Growth on NH_4^+ (HR + RI)	No growth on NO_3^- (%)	HR lines confirmed by PCR analyses (no. of tested lines)	Name of HR lines
KanMx-as	G418	425	8[a] (about 2%)	8 (20)	HR1 *KanMx-as*
KanMx-s	G418	193	4[a] (about 2%)	4 (20)	HR1 *KanMx-s*

Wt, wild type; RI, random insertion.
[a]Lines were all analyzed by PCR.

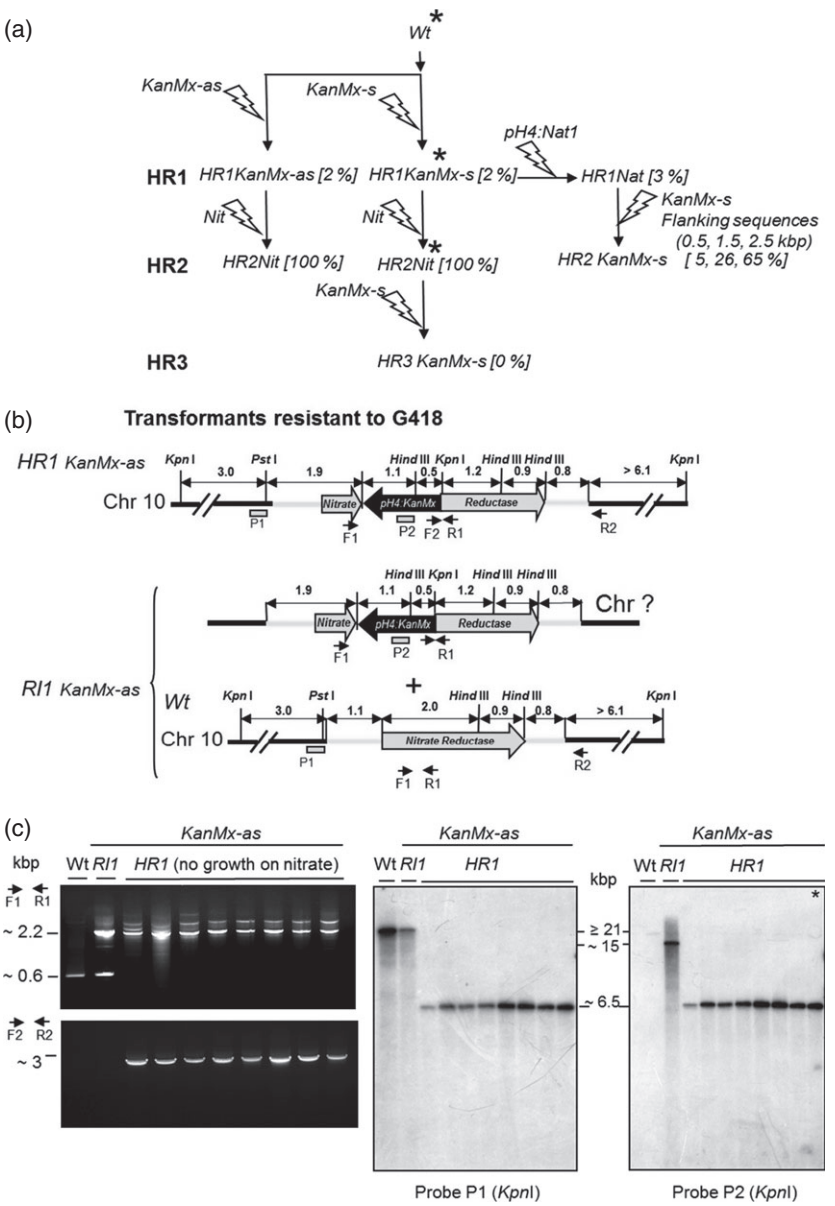

Figure 2. Targeted disruption of the nitrate reductase (*NR*) gene by homologous recombination (HR) using antisense *KanMx* constructs.
(a) Summary of the *NR* targeting experiments showing the history of each of the lines. Lines marked with asterisks were used for the chromatin immunoprecipitation experiments in Figure S4. The percentage of HR is given in brackets. The length of the homologous sequences used in Figure 4 is shown in parenthesis.
(b) Expected physical maps of homologous recombinants (*HR1KanMx-as*) and random insertion lines (*RI1KanMx-as*). The wild-type (Wt) *NR* locus is expected to remain intact if insertion takes place at a random genomic location.
(c) Analysis of *HR1KanMx-as* lines. Left: PCR analyses using either internal *NR* (F1,R1) or external and *KanMx* (F2,R2) primers. Right: Southern blot analysis. DNA was digested with *Kpn*I and hybridized to either *NR*- or *KanMx*-specific probes P1 and P2, respectively. All of the clones that failed to grow on nitrate had patterns of amplification and hybridization consistent with disruption of *NR* by HR.

In summary, the molecular characterization of HR events at the *NR* locus indicated that HR occurred at frequency of around 1–2%, regardless of orientation of the antibiotic resistance cassette (Table 1). No additional insertion took place through a random process in any of the HR lines analysed.

Complementation of the NR mutation by HR occurs at a very high frequency

In order to confirm that the lack of growth on nitrate-ASW was due to the disruption of the *NR* gene, we tested whether replacement of the disrupted copy of the *NR* gene in *HR1* lines with the Wt sequence restored a Wt phenotype. One *HR1KanMx-s* line and one *HR1KanMx-as* line were transformed with a 4.2-kbp *Nit* fragment comprising the full *NR* gene (see Figure 1). Transformants that had acquired a Wt copy of the *NR* gene were selected on nitrate-ASW medium. These transformants (denoted *HR2NitKanMx-s* or *HR2NitKanMx-as* depending on the parental HR1 *KanMx-s* or *KanMx-as* lines) were expected to arise either from replacement by HR of the disrupted copy of the *NR* gene or by RI of the *Nit* fragment elsewhere in the genome. The latter scenario should result in retention of the *KanMx* cassette at the *NR* locus and therefore resistance to the antibiotic G418. Surprisingly, none of the *HR2NitKanMx-s* and *HR2NitKanMx-as* transformants ($n = 54$ and 219, respectively) grew in the presence of G418 (Table 2). This suggested that the native *NR* locus had been restored in all of these transformants by HR. This was confirmed by PCR and Southern blot analyses as above (Figure 3).

The HR frequency is increased at loci containing foreign DNA

The very high frequency of HR in the complementation experiment contrasted with that of the original transformation experiments (100% compared with 1–2%). We first hypothesized that the frequency of HR might be affected by the different types of selection used in these experiments. In order to test this hypothesis, we generated *HR1Nat* lines that contained a *Nat1* insertion at the *NR* locus (see Figure S2). The *KanMx* and *Nat1* selection cassettes share the *histone H4* promoter and *Tef* terminator sequences (0.4 and 0.25 kbp long, respectively). We therefore transformed the *pH4:Nat1* construct into the *HR1KanMx-s* line in order to target the selection cassette (Figure S2a). Four out of 122 transformants showed a phenotype expected from HR, i.e. resistance to CloNat and sensitivity to G418 (Table 2). These were designated *HR1Nat* (Figure 1c). Polymerase chain reaction analysis using *PH4*- and *Tef*-specific primers confirmed that the *KanMx* marker had been replaced by *Nat1* in these lines. We were also able to replace the *Nat* selection cassette with *KanMx* in a subsequent experiment (Figure S2b). Transformation of a *HR1Nat* line using the *KanMx-s* construct to replace the *Nat1* resistance gene in *HR1Nat* line resulted in 39 transformants out of 60 becoming sensitive to CloNat (Table 2). The loss of the *Nat1* marker in these lines suggested that they were the product of HR, and this was confirmed in PCR analyses (Figure S2). The frequency of HR in this experiment was comparable with that obtained using the *Nit* construct in the complementation experiment above (approximately 65%). As acquisition of the *KanMx-s* and the *Nit* sequences is selected for using different mechanisms, this showed that the efficiency of HR was unrelated to the selection mechanism. However, the frequency of HR was much greater when transforming the *KanMx*-s construct into *H1RNat* than in wild-type cells (65% compared with 1–2%). This suggested that the frequency of HR was linked to the recipient line rather than to the specific construct used in the transformation experiment. We hypothesized that the interruption of the *NR* locus by a foreign sequence may increase the rate of HR, whether to restore the native DNA sequence or to replace one foreign DNA insertion

Table 2 Gene replacement at the *NR* locus. All transformants were analyzed by PCR

	Transformants								
	Homology length (kbp)								
Construct	5′	3′	Recipient strain	Selection	NO_3^- growth (%)	G418 growth (%)	CloNat growth (%)	PCR HR lines (% HR/RI)	Name of HR lines
Nit	1.9	2.5	*HR1 KanMx-s*	NO_3^-	219 (100%)	0	N/A	20 (20) (100%)	*HR2Nit KanMx-s*
Nit	1.9	2.5	*HR1 KanMx-as*	NO_3^-	54 (100%)	0	N/A	54 (54) (100%)	*HR2Nit KanMx-as*
pH4Nat	0.5	0.25	*HR1 KanMx-s*	CloNat	N/A	122	4 (3%)	4 (122) (3%)	*H1RNat*
KanMx-s	1.9	2.5	*HR1Nat*	G418	N/A	60	21 (35%)	39 (60) (65%)	*HR2 KanMx-s*
KanMx-s	1.9	2.5	*HR2Nit*	G418	0 (0%)	192	N/A	0 (192) (0%)	*HR3 KanMx-s*

HR, homologous recombination; RI, random insertion; N/A, not applicable.

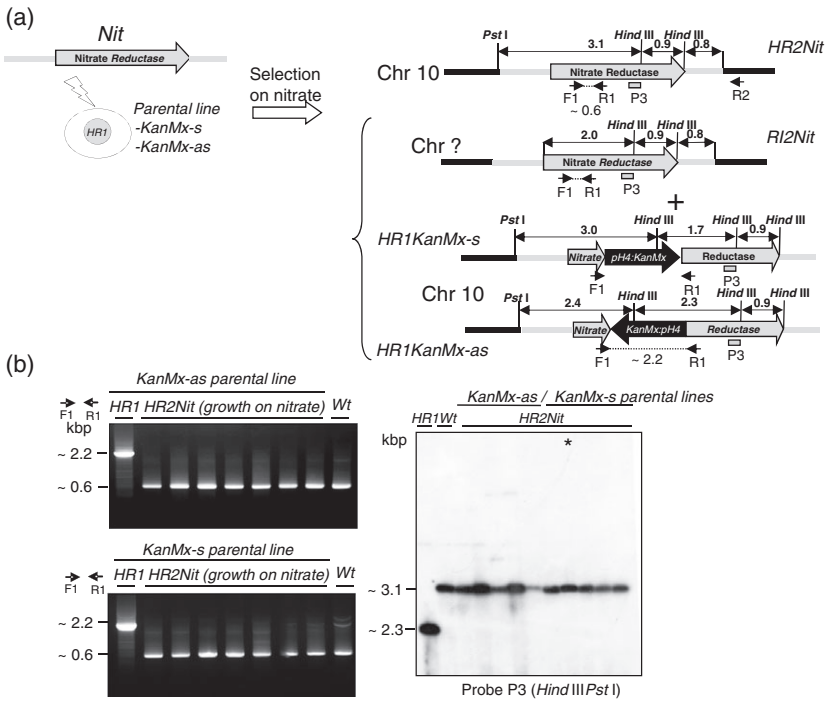

Figure 3. High-frequency transgene removal from the nitrate reductase (NR) locus by homologous recombination (HR).
HR1KanMx-s or *-as* cell lines were transformed using a wild-type (Wt) *NR* sequence fragment and selected on nitrate-Artificial Seawater (ASW).
(a) Expected physical maps of homologous recombinants (*HR2Nit*) and random insertion (RI) lines (*RINit*). Homologous recombination between the Wt transgene and the disrupted locus in the parental cell lines should lead to restoration of the Wt locus. However, RI of the *NR* sequence should introduce an additional copy of the *NR* locus elsewhere in the genome (indicated by Chr?).
(b) Analysis of transformants. Left: PCR analysis using the F1 and R1 primers. All transformants that grew on nitrate gave 0.6-kbp PCR bands typical of the Wt *NR* locus. Right: Southern blot analysis. Genomic DNA from transformants, from the HR1 recipient and from a Wt cell line was digested with *Hin*dIII and *Pst*I and hybridized to the *NR* probe P3. The line marked with an asterisk was used for the chromatin immunoprecipitation experiments in Figure S4.

with another. According to this hypothesis, HR frequencies should return to very low levels upon removal of the foreign DNA sequence. Thus, when *HR2Nit* cells (produced by complementation of the *HR1 NR* knock-out lines) were transformed using the *KanMx-s* construct, all of the 192 transformants were the product of RI rather than HR (Figure S3, Table 2). These results indicate that HR frequencies are low whether targeting the native or the restored *NR* gene, and are high when targeting a disrupted *NR* locus.

Importance of the length of homologous sequences

Homologous recombination occurred at frequencies ranging between 65 and 100% when targeting the disrupted *NR* locus using the *Nit* and *KanMx-s* constructs, which both contain 2.5 kbp of flanking sequences homologous to *NR*. In contrast the HR frequency was much lower (4%) using the *pH4:Nat1* construct (see Figure 1), which only comprises 0.25 and 0.5 kbp of sequence homologous to the *KanMx* cassette (Table 2). This suggests that the length of the homologous sequence may be a critical parameter in the efficiency of HR. This possibility was explored by systematically testing the effect of the length of homologous sequences on the efficiency of gene targeting. *HR1Nat* cells were transformed with a *KanMX-s* construct that contained *NR* flanking sequences of various lengths (0.5–2.5 kbp). The HR frequency with the full-length construct (approximately 65%) decreased to 26 and 5% when using homologous sequences of length 1 and 0.5 kbp, respectively (Figure 4). This result indicates that a length of homologous sequence >2.5 kbp is required for optimal gene targeting by HR.

Role of chromatin structure and transcription

There is evidence that chromatin packing regulates HR at immunoglobulin loci in *Drosophila* (Cummings *et al.*, 2007). Furthermore, histone acetyl transferase activity was required for HR in HeLa cells, suggesting a role for chromatin structure in this process (Kotian *et al.*, 2011). We therefore hypothesized that the dramatic increase in recombination frequencies upon insertion of the *KanMx* cassette in HR1 lines may be due to remodeling of the local chromatin. This effect would then be reversed upon removal of the *KanMx* cassette, leading to lower recombination rates in complemented, HR2 lines. We tested for changes in nucleosome density at the *NR* locus by quantifying the *in vivo* binding of histone H3 in chromatin immunoprecipitation (ChIP) experiments. In addition we tested for changes in histone H3 modifications that either open up the chromatin (H3 acetyl K9) or that result in a more compact structure (H3 trimethyl K4). Three different regions surrounding the insertion site were assayed, including the *NR* promoter (Region 2), the 3′ end of the *NR* coding region (Region 1) and Region 3, which falls outside the homologous sequence used in our *KanMx-s* construct. The CpG methylation of DNA was assayed by restriction enzyme digestion using the *Hpa*II/*Msp*I pair, followed by PCR analysis (Figure S4). No consistent reversible change was observed across any region that would support our hypothesis that one of the specific histone modifications

or DNA methylations tested could regulate the frequency of HR (Figure S4).

Integration of foreign DNA by HR interrupts the transcription of the *NR* gene. The effect of inhibition of transcription on the targeting efficiency was tested by growing Wt cells in ammonium-ASW prior to genetic transformation to repress the transcription of the *NR* gene. Polymerase chain reaction analysis using *NR*- and *KanMx*-specific primers (as described in Figure 2) revealed that all of the 192 transformants arose from RI of the *KanMx-s* transgene.

Knocking-out and tagging of the native *ferritin* gene

Experiments so far have focused on the *NR* locus, which is especially attractive because of the ease with which knock-out mutants can be identified. To determine whether the use of HR-mediated gene targeting could be generalized to other loci in *O. tauri*, we tested its application to the *ferritin* gene which does not have a selectable phenotype.

In *O. tauri* the ferritin is encoded by a single gene that is located on chromosome 2 (Figure 5). A similar approach to that described for *NR* was used to disrupt the *ferritin* gene. Wild-type cells were transformed with the gene interrupted by the *KanMx* selection marker. Ninety-six transformants were obtained after selection on G418 medium. These were subsequently screened by PCR to identify lines resulting from HR events (Figure 5a). Two HR lines (denoted *HRFe*) were identified by the production of a PCR band at 0.9 kbp using a *KanMx*-specific primer (FFt) and a reverse primer located in the *Ferritin* sequence (RFo). In addition, the 0.9-kbp band amplified from Wt cells using *ferritin* internal primers (FFe and RFe) was replaced by a band at 2.5 kbp indicating that the *KanMx* sequence had inserted into the endogenous *ferritin* locus. The frequency of recombination (2%) was in the range observed when targeting the Wt *NR* gene.

Gene targeting by HR is useful for generating gene knock-outs, but also for knock-in sequences such as epitope tags downstream of protein-coding sequences. Translational reporter fusions to the firefly luciferase reporter gene have been extensively used to monitor circadian changes in gene expression in *Ostreococcus* (Corellou *et al.*, 2009; Moulager *et al.*, 2010; Djouani-Tahri *et al.*, 2011b). We designed a knock-in construct to insert the luciferase sequence downstream of, and in frame with, the ferritin-coding region. The construct comprised the *luciferase-KanMx* flanked upstream by the ferritin-coding region (lacking a stop codon) and on the other side by downstream sequences of the *ferritin* gene (Figure 5b). A 1.2-kbp product indicative of HR was amplified from 10 out of 42 transformants using the *KanMx*-specific primer FH4 and the external primer RFi (24%). These 10 *HRFlu* lines also failed to yield a product at 0.27 kbp using the *ferritin*-specific primers FFo and RFo which was detected in Wt and RI lines (Figure 5b). Western blot analysis using an anti-luciferase antibody revealed a protein band at about 85 kDa in both *RIFlu* and *HRFLu* lines (Figure 5b), confirming the production of a *Ferritin-luc* fusion protein. These results indicate that HR can be used to generate not only knock-out but also knock-in recombinants in *Ostreococcus tauri*.

DISCUSSION

Efficient gene targeting by homologous recombination in *Ostreococcus tauri*

The ability to target DNA constructs to specific homologous regions of the genome provides a tool for functional genomics analysis that is available in only a few eukaryotic model systems and even fewer photosynthetic organisms.

Here we report the successful targeting of the *NR* locus in *O. tauri* in more than 10 independent experiments. We further demonstrate the knock-out of a non-selectable gene, *ferritin*, at a similar frequency. Southern blot experiments indicated that no illegitimate integration occurred in HR lines, suggesting that HR which occurs by DSB repair and RI (which occurs by NHEJ) are mutually exclusive in *O. tauri*. In contrast to previous reports in the moss *Physcomitrella patens* (Kamisugi *et al.*, 2006), we observed no

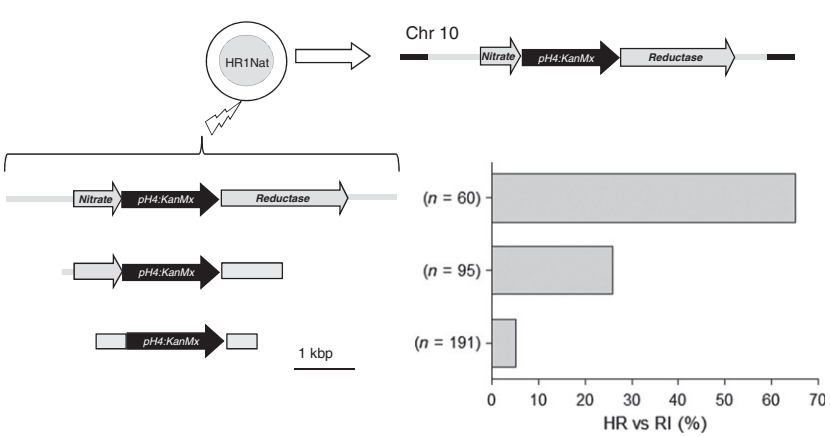

Figure 4. Effect of the length of homologous sequences on the efficiency of gene targeting. A cell line resistant to CloNat (*HR1Nat*) was transformed using constructs containing either 2.5, 1 or 0.5 kbp of sequence homologous to *nitrate reductase* (*NR*) on either side of the *KanMx-s* resistance cassette. The percentage of transgene insertion by homologous recombination (HR) is plotted on the *x*-axis (RI, random insertion).

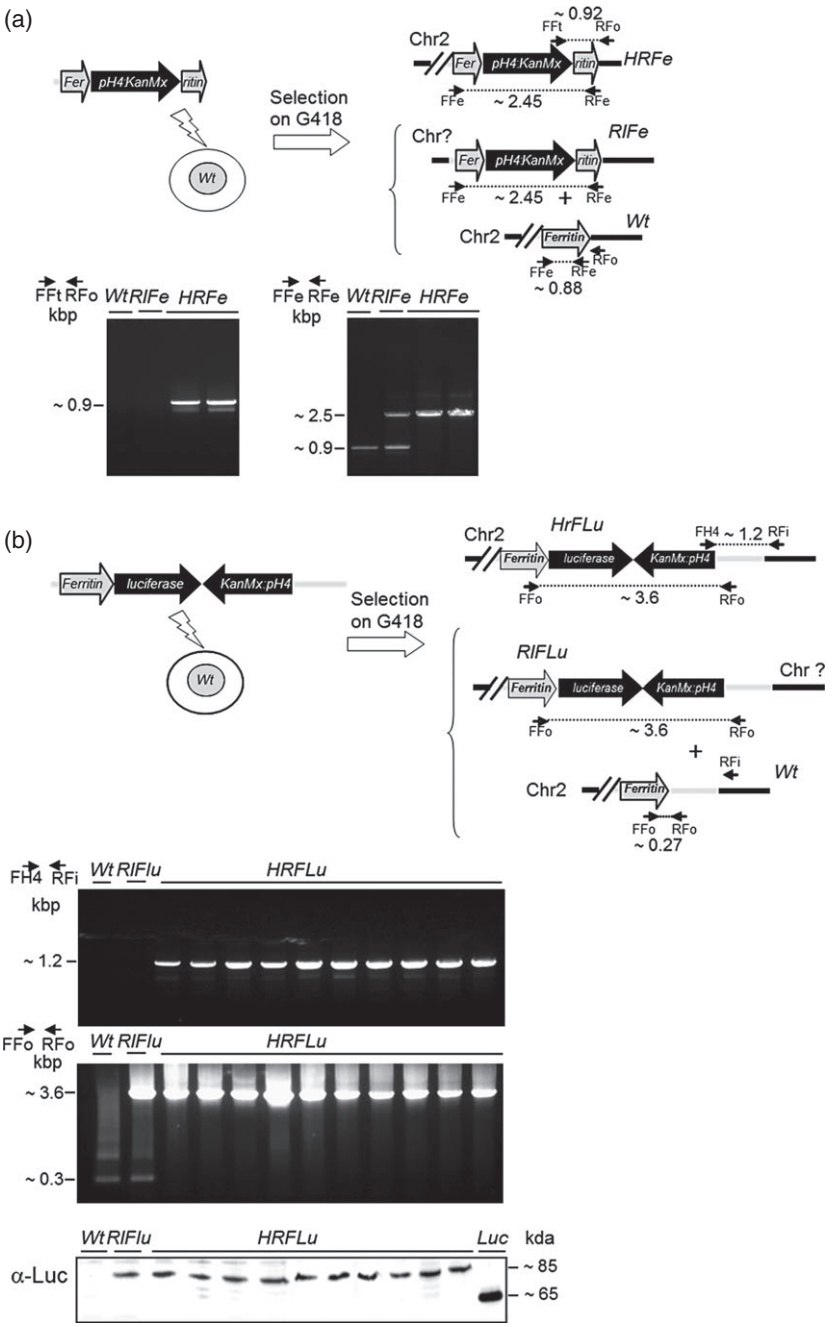

Figure 5. Targeted knock-out and knock-in of the *luciferase* reporter at the *ferritin* locus.
(a) Knock-out of the *ferritin* gene. Wild-type (Wt) cells were transformed using a *ferritin* sequence fragment disrupted by the *KanMx* selection marker. Transformants were identified by selection on G418. Homologous recombination (HR) was confirmed by obtaining a 0.9-kbp PCR product using the primers FFe and RFo, and a 2.5-kbp PCR product using the primers FFe and RFe. RIFe indicates *ferritin* random insertion (RI) lines and *HRFe*, homologous insertion lines.
(b) In-frame knock-in of the *luciferase* reporter. A construct in which the *luciferase* reporter gene was inserted in frame downstream of the *ferritin* open reading frame was electroporated into *Ostreococcus tauri* cells. Knock-in lines (*HRFlu*) were identified by PCR using either a combination of the selection marker FH4 and external primer RFi, or using the *ferritin*-specific primers FFo and RRFo. *RIFlu* indicates RI of the *ferritin-luc* construct. Bottom: Western blot analysis using an anti-luc (α-luc) antibody detected a protein at 85 kDa in *HRFlu* and *RiFlu* lines. *Luc* indicates recombinant luciferase protein (65 kDa), expressed in *O. tauri* under control of the *high affinity phosphate transporter* promoter.

concatenation of transgenes during HR. Taken together our results suggest that HR will be the method of choice for targeted gene disruption for the knock-in of transgenes in *Ostreococcus*.

Homologous recombination was observed at frequencies varying between 1 and 100% depending on the construct and the recipient strain. The length of homology was a key factor in determining the efficiency of gene targeting, but the mechanism of transgene selection had no effect. The GC content of the transgene was not important, as shown by the similar rates of recombination of the *KanMx* and *Nat1* sequences (40 and 60% GC, respectively) with the Wt gene.

The identification of HR events could be made even more effective by the use of a counter-selection marker to select against RI events. For example, the use of chlorate facilitated the selection of *NR* knock-out mutants in *Chlamydomonas* despite the very low frequencies of HR in this green microalga (Sodeinde and Kindle, 1993). However, our attempts to use chlorate as a counter-selection marker in *Ostreococcus* failed, all chlorate-resistant clones being false positives and able to grow on nitrate.

A memory of recombination events at the *nitrate reductase* locus?

To our surprise, gene integration occurred exclusively by HR when the native *NR* sequence was used to replace the disrupted *NR* sequence in *HR1KanMx* lines. Rates of HR were also high (65%) when a construct containing a *KanMx* selection marker was used to replace the *Nat1* sequence in *HR1Nat* lines. This contrasted with the much lower rates of HR when transforming Wt cells with the *KanMx* disruption cassette. The restored *NR* locus resulting from the complementation of the *NR* knock out by the Wt gene also displayed low rates of HR. On the basis of these observations, it is tempting to speculate that elevated rates of HR in HR1 *HR1KanMx* and *HR1Nat* lines result from the presence of insertions at the *NR* locus. In budding yeast, similarly, the insertion of the prokaryotic Tn3-beta lactamase into the genome resulted in a hotspot for meiotic recombination (Stapleton and Petes, 1991).

The genome of *O. tauri* has one of the highest gene densities known for a free-living eukaryote, and contains very few repeated and transposon-like sequences (Derelle *et al.*, 2006). This suggests that efficient mechanisms may operate to remove foreign DNA from the genome. One such mechanism may be based on HR. This would require the pairing of homologous DNA sequences. *Ostreococcus tauri* is haploid in culture, but pairing of homologous chromosomes may occur during mating. Comparison of *O. tauri* strains supports the existence of cryptic sex, although this remains to be demonstrated under laboratory conditions (Grimsley *et al.*, 2010). Whether HR takes place during sexual reproduction or not and how inserted DNA may be recognized as foreign and marked as a recombination hotspot is currently unknown.

We hypothesized that this may be mediated through changes in chromatin structure, because histone modifications such as H3K4Me can modulate the rate of recombination, and meiotic hotspots are usually associated with an open chromatin structure (Berchowitz and Hanlon, 2009; Borde *et al.*, 2009). We did not observe any obvious correlation between the abundance of H3 trimethyl K4, H3 acetyl K9, nucleosome density at the *NR* locus and the rate of HR in *Ostreococcus.* In addition, DNA CpG methylation did not promote HR at the *NR* locus. Transcription inhibition had a slightly positive effect on the efficiency of HR (0/192 in NH_4^+ versus 4/193 in NO_3^- Wt cultures) but these rates were much lower than those observed when targeting the *NR* disrupted locus (Figures 3 and S2).

Our results indicate that recombination frequencies vary with chromosomal location, which may reflect different states of the chromatin. For example, while the rate of HR to replace the ferritin open reading frame with a selection marker was only 1%, it was 25% for the knock-in construct which was targeted 500 bp downstream. This suggests that the recombination rate is dependent on the locus and the chromatin state, which can vary along chromosomes. At this stage, further investigations would be required to determine how inserted DNA sequences are efficiently removed from the *Ostreococcus* genome.

CONCLUSION

For most model organisms, like the plant *Arabidopsis thaliana*, large insertion collections of mutants are used to compensate for the lack of efficient gene targeting technologies. Our results indicate that in *O. tauri* gene targeting by HR can be used to knock out specific genes, or to fuse a protein using an epitope tag or a reporter gene. The whole process of gene targeting, from transformation of the transgene to the identification of haploid knock-out lines by PCR, takes about 3 weeks. In the future, this technology could be used to introduce mutations at selected positions in the DNA. When combined with the use of an inducible promoter (Djouani-Tahri *et al.*, 2011b), this should facilitate the controlled and inducible expression of recombinant proteins from specific loci. These tools promise a bright future for *Ostreococcus* as a 'green yeast' for functional genomics, systems biology and biotechnological developments.

EXPERIMENTAL PROCEDURES

Algal culture, genetic transformation and biological tests

Ostreococcus tauri strain 0TTH0595 was grown in flasks (Sarstedt, http://www.sarstedt.com/) or white 96-well microplates (Nunc, Perkin Elmer, http://www.perkinelmer.com/) under constant light at an intensity of 20 μmol quanta cm^{-2} sec^{-1}. Cells were grown in standard Keller medium, which contained natural seawater supplemented with trace metals and vitamins unless otherwise stated. Cell counting was performed by flow cytometry using a Cell Lab Quanta® SC MPL (Beckman Coulter, https://www.beckmancoulter.com/). Cells were fixed in 0.25% glutaraldehyde for 20 min before flow cytometric analysis.

Genetic transformation was carried out by electroporation as previously described (Corellou *et al.*, 2009). Stable transformant colonies were selected in semi-solid medium at 0.2% w/v agarose (low-melting-point agarose; Invitrogen, http://www.invitrogen.com/) in Keller medium supplemented with an appropriate antibiotic, G418 at 1 mg ml^{-1} (Merck Chemicals, http://www.merckmillipore.com) or CloNat (nourseothricin) at 2 mg ml^{-1} (WERNER BioAgents, http://www.webioage.de/). For the *NR* targeting experiments a Keller-based ASW (WERNER BioAgents) medium was used. This modified Keller medium contained 24.55 g L^{-1} NaCl, 0.75 g L^{-1} KCl, 4.07 g L^{-1} $MgCl_2$ $6H_2O$, 1.47 g L^{-1} $CaCl_2$ $2H_2O$, 6.04 g L^{-1} $MgSO_4$ $7H_2O$, 0.21 g L^{-1} $NaHCO_3$, 0.138 g L^{-1} NaH_2PO_4 and 0.75 g L^{-1} $NaNO_3$. NR^{KO} were selected on this medium supplemented with 0.534 g L^{-1} NH_4Cl. In *NR* complementation experiments, $NaNO_3$ was the only source of nitrogen.

NR^{KO} transformants were identified by their ability to grow on ASW lacking ammonium in microplates and confirmed by PCR tests (see below). Disruption of *Nat1* or *KanMx* genes was identified by the inability to grow on G418 at 1 mg ml^{-1} or on CloNat at 2 mg ml^{-1}. These tests were carried out with individual transformants grown on microplates.

Cloning strategy for gene targeting experiments

The PCR amplifications were performed on *O. tauri* genomic DNA using the Triple Master polymerase mix (Eppendorf, http://www.eppendorf.com/). Primer sequences are given in Table S1. The *Bgl*II and *Nco*I sites were added to the 5′ and 3′ ends of the 1102-bp *ferritin* PCR product to allow subsequent cloning. A subcloning step was performed in the pGEMT-easy vector (Promega, http://www.promega.com/). The *pH4::KanMx::Tef* selection cassette from the pOtluc vector (Corellou *et al.*, 2009) was introduced into the coding sequence of *NR* and *ferritin* by blunt-end ligation into *Afe*I (AGC/GCT) or *Nru*I (TCG/CGA) sites, respectively. The same strategy was used to introduce pH4::CloNat::Tef selection gene from the pOtox vector (Corellou *et al.*, 2009) into the *NR* sequence. The *ferritin-luc* knock-in construct was generated by cloning the full-length *ferritin* gene into the *Bgl*II–*Nco*I sites of the pOT-Luc vector, in frame with *luciferase*. A 1102-bp fragment corresponding to the 3′ end of the *ferritin* gene was then ligated into the *Sma*I site of pOtluc. The resulting construct (5462 bp) comprised the full-length *ferritin* gene fused in frame with *luciferase*, upstream of the *KanMx* selection gene. Prior to transformation, plasmids were digested with appropriate restriction enzymes and purified onto a NucleoSpin® Gel and PCR Clean-up kit (Macherey Nagel, http://www.mn-net.com/).

Molecular analysis of transformants

Ostreococcus transformants were analyzed by PCR using transgene-specific primers (see Table S1) and/or by Southern blot analysis. Genomic DNA was extracted using the DNeasy Plant Minikit (Qiagen, http://www.qiagen.com/). For DNA blots, 1 μg of DNA was digested with appropriate enzymes, migrated in a 0.8% 2-amino-2-(hydroxymethyl)-1,3-propanediol (TRIS)-acetate-EDTA (TAE) agarose gel and transferred onto Hybond N+ membrane (GE HEALTHCARE Life sciences, http://www.gelifesciences.com) as previously described (Corellou *et al.*, 2009). DNA probes were generated by PCR amplification using primers described in Table S1, except for the P3 probe corresponding to the *KanMx* sequence, which was excised from the pOtluc vector using *Hin*dIII and *Eco*RI. All probes were radiolabelled by random priming. Hybridization and washing of the membranes were performed as previously described (Corellou *et al.*, 2009).

Chromatin immunoprecipitation

Cultures were grown in constant light until late log-phase. Cells were fixed by the addition of formaldehyde to the medium to a concentration of 1% (v/v). Glycine was added after 10 min to a 125 mM concentration. After 5 min cells were washed twice with ice-cold PBS. Cell pellets were frozen in liquid nitrogen then stored at −80°C until extraction. In order to extract chromatin, cells were resuspended in extraction buffer (50 mM TRIS-HCl pH 8, 10 mM EDTA, 1% SDS) containing a protease inhibitor cocktail (Roche, http://www.roche.com/) prior to sonicating three times for 10 sec at 50-sec intervals using a Branson sonicator. The cell lysate was centrifuged at 10 000 ***g*** for 10 min and the supernatant containing the chromatin was frozen at −80°C. For ChIP analyses, 200 μl of chromatin was diluted to 2 ml using ChIP dilution buffer (167 mM NaCl, 16.7 mM TRIS-HCl pH 8, 1.2 mM EDTA, Triton X-100, 1 mM phenylmethylsufonyl fluoride and protease inhibitors). After pre-clearing with protein A Dynabeads (Invitrogen), samples were incubated overnight at 4°C with either anti-H3 (1:200), anti H3-Acetyl K9 or anti-H3 trimethyl K4 antibodies (Abcam, http://www.abcam.com/). The immunocomplexes were isolated by incubation with protein A Dynabeads for 2 h at 4°C. The beads were washed as described (Haring *et al.*, 2007) with the addition of three extra-high-salt buffer washes. The DNA to be analysed by quantitative PCR was eluted from the beads in the presence of 10% Chelex according to Nelson *et al.* (2006). Real-time PCR was carried out on a LightCycler 1.5 (Roche Diagnostics, http://www.roche.com/) with LightCycler DNA Master SYBR Green I (Roche Molecular Biochemicals) Putative target loci in immunoprecipitated samples were amplified using specific primers (Table S1). Results were analysed using the comparative critical threshold ($\Delta\Delta C_T$) method, quantified relative to the original input chromatin sample and presented as percentage of input DNA.

DNA methylation analysis

To analyze the DNA methylation status, 0.25 μg of genomic DNA was digested with *Msp*I and *Hpa*II restriction enzymes (Promega) overnight prior to PCR amplification of the regions of interest. These two enzymes have different sensitivities to CpG methylation. When the internal CpG in the 5′-CCGG-3′ tetranucleotide sequence is methylated, cleavage with *Hpa*II is blocked, but cleavage with *Msp*I is not affected. Polymerase chain reaction analysis was performed on DNA regions containing *Msp*I/*Hpa*II sites (see Table S1) so that cytosine demethylation of any of the CpG target sites would impair the amplification of *Hpa*II-cut DNA. *Msp*I-digested DNA was used as a negative control. A fragment lacking CpG *Msp*I/*Hpa*II sites was used as positive control (c).

Western blot analysis

Cells were harvested by centrifugation in conical bottles (10 000 ***g***, 4°C, 10 min), after addition of Pluronic-F68 (0.1%) to the medium. Pellets were frozen in liquid nitrogen and stored at −80° C until extraction. Cells were ground in lysis buffer (100 mM potassium phosphate pH 7.8, 1 mM EDTA, 1 mM DTT, 1% Triton® X-100, 10% glycerol) using a RNA tissue lyser.

Protein concentration was determined by the Bradford method (Sigma, http://www.sigmaaldrich.com/) and the same amount of protein was loaded in each well on a 10% SDS-polyacrylamide denaturing gel with 4× Laemmli buffer. Western blot analysis was performed as follows: the gel was liquid-transferred onto a nylon membrane (PVDF, Amersham Life Sciences). The membranes were blocked in TRIS buffer saline (TBS) containing 5% milk powder antibody for 1 h and then incubated with an anti-luciferase antibody (sc-74548, Santa Cruz Biotechnology, http://www.scbt.com/) at a 1/2000 dilution The membranes were washed three to six times in TBS containing 0.1% Tween 20 and the bound antibody was detected with a goat anti-mouse antibody (sc-2005). Immunodetection was performed using the ECL+ reagent (Amersham Life Science). Recombinant luciferase from *Photinus pyralis* (Sc-32896) was used as a positive control.

ACKNOWLEDGEMENTS

This work was supported by ASSEMBLE EU Project to FYB and IC and ANR Phytoiron to EL and FYB. Thanks to Charles White for helpful discussions and advice.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article.

Figure S1. Targeted disruption of the *NR* gene by homologous recombination using sense *KanMx-s* constructs.

Figure S2. Selection marker replacements by homologous recombination.

Figure S3. Targeting of the *NR* restored line.

Figure S4. Chromatin remodeling and DNA CpG methylation during successive homologous recombination events.

Table S1. List and sequences of oligonucleotides used in cloning and PCR experiments. The size (in bp) of the PCR products is indicated.

REFERENCES

Berchowitz, L. and Hanlon, S. (2009) A positive but complex association between meiotic double-strand break hotspots and open chromatin in Saccharomyces cerevisiae. *Genome Res.* **19**, 2245–2257.

Borde, V., Robine, N., Lin, W., Bonfils, S., Géli, V. and Nicolas, A. (2009) Histone H3 lysine 4 trimethylation marks meiotic recombination initiation sites. *EMBO J.* **28**, 99–111.

Corellou, F., Schwartz, C., Motta, J.-P., Djouani-Tahri, E.B., Sanchez, F. and Bouget, F.-Y. (2009) Clocks in the green lineage: comparative functional analysis of the circadian architecture of the picoeukaryote *Ostreococcus*. *Plant Cell*, **21**, 3436–3449.

Cummings, W.J., Yabuki, M., Ordinario, E.C., Bednarski, D.W., Quay, S. and Maizels, N. (2007) Chromatin structure regulates gene conversion. *PLoS Biol.* **5**, e246.

Derelle, E., Ferraz, C., Rombauts, S. *et al.* (2006) Genome analysis of the smallest free-living eukaryote Ostreococcus tauri unveils many unique features. *Proc. Natl Acad. Sci. USA* **103**, 11647–11652.

Djouani-Tahri, E.-B., Christie, J.M., Sanchez-Ferandin, S., Sanchez, F., Bouget, F.-Y. and Corellou, F. (2011a) A eukaryotic LOV-histidine kinase with circadian clock function in the picoalga Ostreococcus. *Plant J.* **65**, 578–588.

Djouani-Tahri, E.B., Sanchez, F., Lozano, J.-C. and Bouget, F.-Y. (2011b) A phosphate-regulated promoter for fine-tuned and reversible overexpression in Ostreococcus: application to circadian clock functional analysis. *PLoS ONE*, **6**, e28471.

Gan, L., Ladinsky, M. and Jensen, G. (2011) Organization of the smallest eukaryotic spindle. *Curr. Biol.* **21**, 1578–1583.

Grimsley, N., Péquin, B., Bachy, C., Moreau, H. and Piganeau, G. (2010) Cryptic sex in the smallest eukaryotic marine green alga. *Mol. Biol. Evol.* **27**, 47–54.

Haring, M., Offermann, S., Danker, T., Horst, I., Peterhansel, C. and Stam, M. (2007) Chromatin immunoprecipitation: optimization, quantitative analysis and data normalization. *Plant methods*, **3**, 11.

Heijde, M., Zabulon, G., Corellou, F. *et al.* (2010) Characterization of two members of the cryptochrome/photolyase family from Ostreococcus tauri provides insights into the origin and evolution of cryptochromes. *Plant, Cell Environ.* **33**, 1614–1626.

Heyer, W.-D., Ehmsen, K.T. and Liu, J. (2010) Regulation of homologous recombination in eukaryotes. *Annu. Rev. Genet.* **44**, 113–139.

Kamisugi, Y., Schlink, K., Rensing, S.A., Schween, G., Von Stackelberg, M., Cuming, A.C., Reski, R. and Cove, D.J. (2006) The mechanism of gene targeting in Physcomitrella patens: homologous recombination, concatenation and multiple integration. *Nucleic Acids Res.* **34**, 6205–6214.

Keeney, S., Giroux, C.N. and Kleckner, N. (1997) Meiosis-specific DNA double-strand breaks are catalyzed by Spo11, a member of a widely conserved protein family. *Cell*, **88**, 375–384.

Kilian, O., Benemann, C.S.E., Niyogi, K.K. and Vick, B. (2011) High-efficiency homologous recombination in the oil-producing alga Nannochloropsis sp. *Proc. Natl Acad. Sci. USA* **108**, 21265–21269.

Kotian, S., Liyanarachchi, S., Zelent, A. and Parvin, J.D. (2011) Histone deacetylases 9 and 10 are required for homologous recombination. *J. Biol. Chem.* **286**, 7722–7726.

Minoda, A., Sakagami, R., Yagisawa, F., Kuroiwa, T. and Tanaka, K. (2004) Improvement of culture conditions and evidence for nuclear transformation by homologous recombination in a red alga, Cyanidioschyzon merolae 10D. *Plant Cell Physiol.* **45**, 667–671.

Monnier, A., Liverani, S., Bouvet, R., Jesson, B., Smith, J.Q., Mosser, J., Corellou, F. and Bouget, F.-Y. (2010) Orchestrated transcription of biological processes in the marine picoeukaryote Ostreococcus exposed to light/dark cycles. *BMC genomics*, **11**, 192.

Moulager, M., Corellou, F., Vergé, V., Escande, M.-L. and Bouget, F.-Y. (2010) Integration of light signals by the retinoblastoma pathway in the control of S phase entry in the picophytoplanktonic cell Ostreococcus. *PLoS Genet.* **6**, e1000957.

Nelson, J. D., Denisenko, O., Sova, P. and Bomsztyk, K. (2006) Fast chromatin immunoprecipitation assay. *Nucleic Acids Res.* **34**, e2.

O'Neill, J.S., Van Ooijen, G., Dixon, L.E., Troein, C., Corellou, F., Bouget, F.-Y., Reddy, A.B. and Millar, A.J. (2011) Circadian rhythms persist without transcription in a eukaryote. *Nature*, **469**, 554–558.

Pfeuty, B., Thommen, Q., Corellou, F., Djouani-Tahri, E.B., Bouget, F.-Y. and Lefranc, M. (2012) Circadian clocks in changing weather and seasons: lessons from the picoalga Ostreococcus tauri. *BioEssays*, **34**, 781–790.

Schaefer, D.G. and Zryd, J.-P. (1997) Efficient Gene targeting in the moss Physcomirella patens. *Plant J.* **11**, 1195–1206.

Shinohara, A., Ogawa, H. and Ogawa, T. (1992) Rad51 protein involved in repair and recombination in S. cerevisiae is a RecA-like protein. *Cell*, **69**, 457–470.

Sodeinde, O.A. and Kindle, K.L. (1993) Homologous recombination in the nuclear genome of. *Proc. Natl Acad. Sci. USA* **90**, 9199–9203.

Stapleton, A. and Petes, T.D. (1991) The Tn3 beta-lactamase gene acts as a hotspot for meiotic recombination in yeast. *Genetics*, **2**, 39–51.

Zorin, B., Lu, Y., Sizova, I. and Hegemann, P. (2009) Nuclear gene targeting in Chlamydomonas as exemplified by disruption of the PHOT gene. *Gene*, **432**, 91–96.

