Bioresource Technology 184 (2015) 23–32

Contents lists available at ScienceDirect

Bioresource Technology

journal homepage: www.elsevier.com/locate/biortech

Review

Gene silencing in microalgae: Mechanisms and biological roles

Eun-Jeong Kim, Xinrong Ma, Heriberto Cerutti *

School of Biological Sciences and Center for Plant Science Innovation, University of Nebraska-Lincoln, Lincoln, NE 68588, USA

HIGHLIGHTS

- Most Archaeplastida microalgae contain histone H3 lysine 27 methyltransferase genes.
- Pervasive DNA cytosine methylation seems to rely on diverged DNA methyltransferases.
- RNA interference mechanisms, when present, may suppress transposons and viruses.
- The role of epigenetic silencing mechanisms in gene regulation remains uncharacterized.
- RNA interference is a useful tool for functional analyses and genetic engineering.

ARTICLE INFO

Article history:
Received 1 September 2014
Received in revised form 22 October 2014
Accepted 24 October 2014
Available online 30 October 2014

Keywords:
Epigenetics
Histone post-translational modifications
DNA methylation
RNA interference
miRNAs

ABSTRACT

Microalgae exhibit enormous diversity and can potentially contribute to the production of biofuels and high value compounds. However, for most species, our knowledge of their physiology, metabolism, and gene regulation is fairly limited. In eukaryotes, gene silencing mechanisms play important roles in both the reversible repression of genes that are required only in certain contexts and the suppression of genome invaders such at transposons. The recent sequencing of several algal genomes is providing insights into the complexity of these mechanisms in microalgae. Collectively, glaucophyte, red, and green microalgae contain the machineries involved in repressive histone H3 lysine methylation, DNA cytosine methylation, and RNA interference. However, individual species often only have subsets of these gene silencing mechanisms. Moreover, current evidence suggests that algal silencing systems function in transposon and transgene repression but their role(s) in gene regulation or other cellular processes remains virtually unexplored, hindering rational genetic engineering efforts.

1. Introduction

Algae are a diverse group of eukaryotic organisms with important roles in marine, freshwater, and terrestrial ecosystems (Worden and Allen, 2010; Tirichine and Bowler, 2011). The great potential of algae as feedstocks for renewable biofuel and biomaterial production is also gaining recognition (Hu et al., 2008; Radakovits et al., 2010; Gimpel et al., 2013; Leite et al., 2013). Microalgae are microscopic organisms capable of harnessing sunlight and CO_2 to synthesize useful chemical compounds, such as lipids and carbohydrates, which can be converted into fuels and other bioproducts. However, production of algae-based fuels is technically, but not yet economically, feasible (Lee, 2011; Chisti, 2013). The major economic bottlenecks cited in the literature include microalgae biological productivity, culture systems, crop protection, and harvesting/extraction processes (Hu et al., 2008; Chisti, 2013; Gimpel et al., 2013; Leite et al., 2013).

For large-scale fuel production reliant on algal photosynthesis key objectives will be achieving high productivity per unit of area, environmental (biotic and abiotic) stress tolerance, ease of harvesting and extraction, and a biomass profile optimized for biofuel conversion (Griffiths and Harrison, 2009; Radakovits et al., 2010). However, identifying in nature microalgal strains simultaneously endowed with all these traits has proven difficult (Hu et al., 2008; Griffiths and Harrison, 2009). Additionally, there has been limited success in increasing biomass productivity or oil content in algae by the genetic engineering of individual genes (Radakovits et al., 2010; La Russa et al., 2012; Gimpel et al., 2013), and this limitation emphasizes the importance of comprehending on a genome scale the metabolic and regulatory networks involved in these processes. Indeed, a significant barrier to advancement is that our knowledge of gene function and regulation is still fairly incomplete in most microalgae (Radakovits

* Corresponding author at: School of Biological Sciences and Center for Plant Science Innovation, University of Nebraska-Lincoln, E211 Beadle Center, Lincoln, NE 68588-0666, USA. Tel.: +1 402 472 0247.
E-mail address: hcerutti1@unl.edu (H. Cerutti).

http://dx.doi.org/10.1016/j.biortech.2014.10.119
0960-8524/

et al., 2010; Worden and Allen, 2010; Tirichine and Bowler, 2011). In this context, the study of algal gene silencing mechanisms may provide insights into the control of gene expression as well as facilitate the development of tools for rational genetic engineering.

The regulation of gene expression in eukaryotes involves complex mechanisms, operating at the transcriptional and post-transcriptional levels. Chromatin organization modulates the access of regulatory proteins to DNA and influences multiple aspects of transcription and other DNA-related processes (Bannister and Kouzarides, 2011; Ohsawa et al., 2013). Eukaryotic genomes are commonly organized into several types of chromatin, with euchromatin consisting of transcriptionally permissive or active domains and heterochromatin being characterized by densely packed silent regions (Casas-Mollano et al., 2007; Krauss, 2008; Bannister and Kouzarides, 2011). These functionally and structurally different chromatin states are marked by distinct covalent modifications on the DNA and on specific amino acid residues of the nucleosomal histones (Casas-Mollano et al., 2007; Bannister and Kouzarides, 2011; Saze and Kakutani, 2011). For instance, di- or trimethylation of histone H3 lysine 9 (H3K9) or of histone H3 lysine 27 (H3K27) is often associated with silenced chromatin (Krauss, 2008; Shaver et al., 2010; Bannister and Kouzarides, 2011; Saze and Kakutani, 2011; Derkacheva and Hennig, 2014). DNA cytosine methylation also plays a role in repression and in some organisms there appears to be a complex interplay between histone tail modifications and DNA methylation in establishing a silent chromatin structure (Law and Jacobsen, 2010; Saze and Kakutani, 2011; Du et al., 2012; Zhong et al., 2014).

RNA-directed mechanisms have also been co-opted by evolution to generate a broad spectrum of gene regulatory pathways. RNA-mediated silencing is a conserved process in eukaryotes by which small RNAs (~20–30 nucleotides in length) induce the inactivation of cognate sequences through a variety of mechanisms, including translation inhibition, RNA degradation, and/or transcriptional repression (Cerutti and Casas-Mollano, 2006; Carthew and Sontheimer, 2009; Meister, 2013). The function of long double stranded RNAs, as precursors of small RNAs, in triggering gene silencing was initially characterized in *Caenorhabditis elegans* and termed RNA interference (RNAi) (Fire et al., 1998). Yet, in slightly over a decade, RNAi has evolved into a fascinating biological phenomenon intersecting with multiple cellular pathways. Indeed, histone post-translational modifications, DNA cytosine methylation, and RNA-mediated mechanisms impinge on many cellular processes including, besides regulation of gene expression, DNA repair and recombination, chromatin structure, chromosome condensation/stability, as well as the suppression of viruses and transposable elements (Cerutti and Casas-Mollano, 2006; Carthew and Sontheimer, 2009; Cerutti et al., 2011; Ohsawa et al., 2013; Oliver et al., 2014). Moreover, gene silencing mechanisms seem to be important for the integration of environmental and intrinsic stimuli in the control of gene expression and their disruption leads to physiological and developmental abnormalities (Bannister and Kouzarides, 2011; Ohsawa et al., 2013).

In most algal species, both chromatin-associated and RNA-mediated silencing pathways remain largely uncharacterized, even at the level of identifying crucial gene factors in the sequenced genomes. This review will examine the existence of key histone lysine methyltransferases, DNA cytosine methyltransferases, and core components of the RNA-mediated silencing machinery in microalgae. However, algae are very diverse phylogenetically (Worden and Allen, 2010; Tirichine and Bowler, 2011) and, to simplify the identification of commonalities, the analysis will be restricted to microalgae in the Archaeplastida eukaryotic supergroup, which includes glaucophytes (Glaucophyta), red algae (Rhodophyta), green algae (Chlorophyta), as well as land plants (Streptophyta) (Table 1). We will also discuss briefly the known or inferred biological role(s) of gene silencing mechanisms in these aquatic organisms. It is anticipated that advances in the basic understanding of gene regulatory mechanisms in microalgae will enable optimization of metabolic pathways of interest through hypothesis-driven genetic engineering strategies.

2. H3K9 and H3K27 methyltransferases in microalgae

2.1. Phylogenetic analysis and domain organization of histone methyltransferases

The methylation of lysine residues in histones, with the exception of H3K79 methylation, is carried out by enzymes that contain an evolutionary conserved SET domain, named after three Drosophila genes (*Su(var)3-9*, *Enhancer of zeste*, and *Trithorax*) (Casas-Mollano et al., 2007; Bannister and Kouzarides, 2011; Huang et al., 2011; Derkacheva and Hennig, 2014). The SET domain constitutes the catalytic site of these lysine methyltransferases (KMTs), but flanking sequences, more distant protein domains, and possibly some cofactors are also important for enzyme activity and specificity (Huang et al., 2011; Krishnan et al., 2011; Derkacheva and Hennig, 2014). To begin characterizing the occurrence and the role(s) of H3K9 and/or H3K27 methyltransferases in microalgae, we surveyed 14 complete or near-complete algal genomes in the Archaeplastida supergroup for the presence of SET-domain polypeptides (Table 1).

Proteins with conserved SET domains were identified by either BLAST or PSI-BLAST searches of protein and/or translated genomic DNA databases, using as queries known *Arabidopsis thaliana* or *Homo sapiens* polypeptides containing SET motifs. Since several of the examined genomes are in draft stage, an important caveat in our analyses is that some proteins may be missing from the databases whereas others may have errors in the predicted gene structure. However, with few exceptions, we considered as potential homologs only proteins that exhibited enough sequence similarity to be aligned and used for phylogenetic tree construction. Interestingly, phylogenetic analysis of the extracted SET-domain proteins revealed that they could be grouped into several distinct classes (see also Huang et al., 2011) but we only examined in detail KMT1 and KMT6 homologs (Figs. 1 and 2), following the nomenclature proposed for yeast and metazoan lysine methyltransferases (Allis et al., 2007).

Members of the algal KMT1 class, like animal and plant KMT1 proteins, are likely responsible for H3K9 methylation, an epigenetic mark involved in gene silencing and heterochromatin formation (Casas-Mollano et al., 2007; Krauss, 2008; Bannister and Kouzarides, 2011; Huang et al., 2011). In the examined microalgae, KMT1 homologs appear to be limited to species of the Chlorophyta clade, including organisms in the Trebouxiophyceae (*Chlorella sorokiniana*, *Chlorella variabilis* NC64A, and *Coccomyxa subellipsoidea*) and Chlorophyceae (*Chlamydomonas reinhardtii* and *Volvox carteri*) classes (Table 1 and Fig. 1). *Micromonas pusilla* CCMP1545 also seems to code for a KMT1 related protein (Table 1). Yet, the corresponding gene is located in an island of the genome with no detectable homology to the closely related *Micromonas* sp. RCC299 (data not shown) and the functional significance of the encoded protein is currently unclear. Most algal KMT1 proteins show high sequence similarity to land plant KMT1 polypeptides, both within the SET domain and in the surrounding regions known as the Pre-SET and Post-SET motifs (Fig. 1). Additionally, all algal sequences contain an SRA (SET and RING associated) domain (Fig. 1), which recognizes the methylation status of CG and CHH DNA sequences (where H = A, T, or C) (Rajakumara et al., 2011). Land plant KMT1 proteins have been reported to fall into several distinct subgroups, indicative of functional diversification (Casas-Mollano et al., 2007;