Review

The Florida red tide dinoflagellate *Karenia brevis*: New insights into cellular and molecular processes underlying bloom dynamics

Frances M. Van Dolah\textsuperscript{a,}\textsuperscript{*}, Kristy B. Lidie\textsuperscript{a}, Emily A. Monroe\textsuperscript{a}, Debashish Bhattacharya\textsuperscript{b}, Lisa Campbell\textsuperscript{c}, Gregory J. Doucette\textsuperscript{a}, Daniel Kamkowski\textsuperscript{d}

\textsuperscript{a}Marine Biotoxins Program, NOAA Center for Coastal Environmental Health and Biomolecular Research, Charleston, SC, United States
\textsuperscript{b}Department of Biological Sciences and Roy J. Carver Center for Comparative Genomics, University of Iowa, Iowa City, IA, United States
\textsuperscript{c}Department of Oceanography, Texas A\&M University, College Station, TX, United States
\textsuperscript{d}Department of Marine, Earth and Atmospheric Sciences, North Carolina State University, Raleigh, NC, United States

\textsuperscript{*}Corresponding author at: NOAA CCEHBR, 219 Fort Johnson Road, Charleston, SC 29412, United States. Tel.: +1 843 762 8529; fax: +1 843 762 8700.

E-mail address: fran.vandolah@noaa.gov (F.M. Van Dolah)

The dinoflagellate *Karenia brevis* is responsible for nearly annual red tides in the Gulf of Mexico that cause extensive marine mortalities and human illness due to the production of brevetoxins. Although the mechanisms regulating its bloom dynamics and toxicity have received considerable attention, investigation into these processes at the cellular and molecular level has only begun in earnest during the past decade. This review provides an overview of the recent advances in our understanding of the cellular and molecular biology on *K. brevis*. Several molecular resources developed for *K. brevis*, including cDNA and genomic DNA libraries, DNA microarrays, metagenomic libraries, and probes for population genetics, have revolutionized our ability to investigate fundamental questions about *K. brevis* biology. Two cellular processes have received particular attention, the vegetative cell cycle and vertical migration behavior, which are of key importance due to their roles in the development of both surface populations that constitute blooms and subsurface cell aggregations that may serve to initiate them. High throughput sequencing of cDNA libraries has provided the first glimpse of the gene repertoire in *K. brevis*, with approximately 12,000 unique genes identified to date. Phylogenomic analysis of these genes has revealed a high rate of horizontal gene transfer in *K. brevis*, which has resulted in a chimeric chloroplast through the selective retention of genes of red, green, and haptophyte origin, whose adaptive significance is not yet clear. Gene expression studies using DNA microarrays have demonstrated a prevalence of post-transcriptional gene regulation in *K. brevis* and led to the discovery of an unusual spliced leader trans-splicing mechanism. Among the trans-spliced gene transcripts are type I polyketide synthases (PKSs), implicated in brevetoxin biosynthesis, which are unique among type I PKSs in that each transcript encodes an individual catalytic domain, suggesting a novel gene structure in this dinoflagellate. Clone libraries of 16S ribosomal DNA sequences developed from bloom waters have unveiled the temporal and spatial complexity of the microbial soup that coexists with *K. brevis* and its active involvement in both bloom growth and termination processes. Finally, the development and application of population genetic markers has revealed a surprisingly high genetic diversity in *K. brevis* blooms, long assumed to consist of essentially clonal populations. With these foundations in place, our understanding of *K. brevis* bloom dynamics is likely to grow exponentially in the next few years.
1. Introduction

Red tides have been documented in the Gulf of Mexico since 1530 when the Spanish explorer Cabeza de Vaca reported fish kills and discolored water in what is now Tampa Bay (Steidinger et al., 1998). In 1948 the phenomenon of water discoloration and its attendant fish kills were attributed to an undescribed dinoflagellate, which was named Gymnodinium breve (Davis, 1948). Because of its extensive adverse impacts on coastal environments and communities, the growth physiology and toxicity of this dinoflagellate (renamed Ptychodiscus brevis and most recently Karenia brevis) have now been under study for almost 60 years. However, research on the underlying molecular biology has only been initiated in the last decade, enabled by the availability of rapidly evolving molecular technologies from the biomedical field. The processes contributing to the development of K. brevis blooms span spatial and temporal scales from the molecular mechanisms within the dinoflagellate cell that respond to environmental cues to the oceanographic and even atmospheric processes that drive the local environmental conditions in which K. brevis lives. Significant progress has been made towards understanding the large-scale oceanographic processes that support the growth of K. brevis blooms and dictate their movement on the west Florida shelf, as reviewed by Walsh et al. (2006). However, further refinement of these models requires the integration of cellular responses with these physical forcings. Bloom initiation and bloom termination processes remain poorly understood, but are critical for the prediction and management of bloom impacts. Lack of information on these processes may in large part be because they require a more thorough understanding of the cellular level decisions through which K. brevis interacts with its environment. This review provides an overview of the advances made over the past decade in our understanding of the cellular and molecular processes in K. brevis that provide the necessary underpinning to further refine our understanding of bloom dynamics.

2. Cellular processes central to bloom growth

2.1. Vegetative cell cycle

Bloom formation of K. brevis proceed primarily through vegetative cell division. Therefore, understanding the mechanisms that regulate the vegetative cell cycle and the environmental cues to which it responds is important in order to understand bloom progression. In many phytoplankton, the cell cycle is entrained to the photoperiod, resulting in synchronous (1 division [div] d$^{-1}$) or phased cell division (<1 div day$^{-1}$). Cell division rates observed in both lab and field populations of K. brevis average 0.3 day$^{-1}$ with a maximum of approximately 0.6 day$^{-1}$ (Van Dolah et al., 2007), although synchronous division has been achieved in laboratory columns by repeated removal of the bottom third of the culture (Kamykowski et al., 1998a). Phased cell division was shown to occur only during the dark by Heil (1986) and the relative timing of the remaining cell cycle phases was subsequently determined in lab cultures (Van Dolah and Leighfield, 1999) and field populations (Van Dolah et al., 2007). In both lab and field populations, essentially all cells in a population are in G1 at the onset of the light phase or day. S-phase begins a minimum of 6 h into the light and continues for 12–14 h, mitosis occurs between 18 and 22 h after dawn, and the cell cycle is generally completed by the start of the next day. This pattern was shown to be under circadian control, as it persists in continuous light (Brunelle et al., 2007). Circadian entrainment appears to be mediated by the dark/light transition or “dawn” cue, as delaying the onset of light results in a delay in S-phase entry of an equivalent length of time (Van Dolah and Leighfield, 1999).

Given that S-phase entry is the key event controlling cell cycle progression, the mechanisms that mediate this process play a significant role in regulating bloom growth. Using antibodies and inhibitor studies, two conserved eukaryotic cell cycle regulators, cyclin and cyclin-dependent kinase, were shown to regulate both S- and M-phase entry in K. brevis (Van Dolah and Leighfield, 1999; Barbier et al., 2003). The presence of these conserved cell cycle proteins provided the basis from which to begin investigating the signaling pathways that mediate the “dawn” cue to the cell cycle and to identify cell cycle regulatory mechanisms that may lend themselves as biomarkers of bloom status or targets for control strategies. In Euglena, cAMP signaling appears to mediate diel cell cycle entrainment (Carrie and Edmunds, 1993). Blocking cAMP protein kinase activity inhibits cell cycle progression in the dinoflagellate Amphidinium operculatum (the Leighfield and Van Dolah, 2001). However, it does not appear to be involved in the immediate recognition of the dark/light cue, but rather may work downstream of that event to promote the G1/S phase transition. In K. brevis, a cryptochrome blue light receptor was identified that is a candidate for recognition of the dawn cue, given the differential effects of blue and red light on S-phase entry (Brunelle et al., 2007). However, the signaling pathways associated with this receptor are not currently known, and a role for cryptochrome in the observed blue light effects awaits confirmation.

2.2. Vertical migration behavior

The second cellular process that appears to play an equally critical role in the growth, movement, and possibly initiation of K. brevis blooms is its diel vertical migration behavior. K. brevis cells are not simply passive particles carried by turbulence or ocean currents, but rather are able to use vertical migration to maximize carbon fixation from photosynthesis near the surface during the day and dissolved nutrients at depth during the night. K. brevis can attain swimming speeds of up to 1 m h$^{-1}$ (McKay et al., 2006). Maximal swimming speed is attained during the day (McKay et al., 2006), and appears to be driven by phototaxis and geotaxis (Kamykowski et al., 1998b).

Under nutrient replete conditions in the laboratory, K. brevis vertical migration appears to be under circadian control, as surface accumulation continues during the subjective day when cells are shifted to constant dark (Heil, 1986). Downward swimming does not appear to be as directed: at night cells display a swimming
pattern that contributes to a diffusion of cells throughout the upper water column (Kamykowski et al., 1998b). More recent work has demonstrated that there is also an adaptive component to swimming behavior that is responsive to the internal cellular conditions, perhaps set by threshold values for certain biochemical constituents or the C/N ratio (Yamazaki and Kamykowski, 2000). Kamykowski et al. (1998a) found that approximately 50% of a K. brevis population migrated to the surface following a synchronous cell division. Differences in major biochemical constituents were found in mid-water populations of K. brevis versus those that had migrated to the surface during the day following division, with carbohydrate, lipid, protein and chlorophyll a concentrations generally higher in the cells found at mid-water than in surface cells (Kamykowski et al., 1998a, 1999). This suggested that daughter cells receive unequal shares of the parental resources and that this differential allocation influences their vertical migration behavior, where the nutrient poor daughters are more positively phototactic, thus facilitating their more rapid acquisition of biochemical stores through increased photosynthesis rates. The diel patterns found in biochemical constituents do not reflect simple acquisition rates, however. For example, Evans et al. (2001) showed that the diel pattern of neutral lipid content is responsive to light intensity exposure, but Schaeffer (2006) found that light intensity alone was insufficient to fully explain the diel pattern among different lipid components. The internal cues that trigger differential swimming behavior likely involve the integration of various metabolic and signaling pathways in K. brevis, that have only recently begun to be identified, and their further characterization will be facilitated by the molecular biological approaches described below.

This adaptive swimming behavior – the “choice” to ascend or not depending upon both external conditions and internal biochemical stores – may explain the vertical distributions of cells sometimes observed within K. brevis bloom patches studied over a diel cycle (Fig. 1). In the field, K. brevis blooms often occur in water columns under a broad range of light intensities to which they can readily adapt (Evans et al., 2001; Schaeffer et al., 2007) and in which the nutrient source is not obvious. During these times, aggregated cells can occur near fronts in daylight where nutrients from coastal water egress or fish kills may maintain populations in near surface waters (Walsh et al., 2004). Aggregated cells can also occur near the sediment interface at night where near-bottom nutrients either associated with upwelling plumes that reside below the euphotic zone or sediment pore waters may be utilized (Janowitz and Kamykowski, 2006). Sinclair et al. (2006a) demonstrated that K. brevis cells are capable of enhanced dark nitrate uptake after 12 h of nitrate deprivation that simulated an ascent into a nitrate depleted water column, while Sinclair et al. (2006b) demonstrated that K. brevis cells aggregate in a near bottom nitrate-rich layer if the overlying water column is nitrate depleted. More recently, Sinclair and Kamykowski (2007) observed that K. brevis cells will selectively swim through a thin sediment barrier set on a 100 mm mesh that split a plastic column into a nitrate depleted upper layer and a nitrate enriched lower layer. These laboratory studies indicate that K. brevis can actively seek nutrients and support the hypothesis that offshore, near-bottom aggregates of K. brevis may provide cell populations for bloom initiation, as upwelling favorable winds transport these populations toward the coast and as nearshore fronts interact with cell migrations to serve as an aggregation point for near surface K. brevis populations (Janowitz and Kamykowski, 2006).

3. The K. brevis genome

In the past decade significant investment has been made in developing the genomic resources with which to study the genetic structure and gene regulation critical to K. brevis bloom dynamics. Current molecular resources available for K. brevis are enumerated in Table 1 and their utility for furthering our understanding of K. brevis bloom dynamics is described below.

Like many dinoflagellates, K. brevis possesses a large haploid genome of 100 pg DNA cell−1, or approximately 1 × 1011 bp (Kim and Martin, 1974; Rizzo et al., 1982; Sigee, 1984; Kamykowski et al., 1998a), approximately 30 times the human genome. Unique among eukaryotes, dinoflagellates have permanently condensed chromatin which lacks nucleosomes typically involved in regulating chromosome condensation and eukaryotic gene expression. The condensed chromosomes of K. brevis have a characteristic banding pattern present in other dinoflagellates (Rizzo et al., 1982). This striated structure is proposed to result from formation of a liquid cholesterol DNA crystal in which parallel bundles of DNA filaments form stacked disks that make a continuous left-handed twist along the chromosome’s longitudinal axis (Bouligand and Norris, 2001). Basic DNA-binding proteins are probably involved in stabilizing this structure by neutralizing local electronegative charges that would result from tightly compacted DNA filaments (Hackett et al., 2005). At the periphery of these disks are loops of DNA that are less tightly compacted and comprise the actively transcribed DNA (Sigee, 1984; Bhaud et al., 2000). This unique chromosome structure suggests that dinoflagellates may also have evolved novel mechanisms for regulating gene expression and indeed some unusual genome characteristics and regulatory mechanisms in K. brevis described below may relate to this structure.

3.1. The expressed genome

The large size of the K. brevis genome makes it unfeasible to obtain a complete genome sequence with current technologies. However, the sequencing of cDNA libraries is an efficient approach to discovering expressed genes in organisms for which genomic data are unavailable, and this approach has been undertaken by two research groups. Lidie et al. (2005) analyzed a collection of 7000 expressed sequence tags (ESTs) from log phase K. brevis grown in nutrient replete conditions which provided the first
global insights into K. brevis genomic content. Approximately 29% of the ESTs had similarity to known sequences in the GenBank database, which identified the presence of conserved genes involved in cell cycle control, intracellular signaling, stress responses, intermediary metabolism, transcription/translation machinery, and the photosystem. K. brevis ESTs have an average GC content of 51%, lower than that found in the dinoflagellates Alexandrium tamarense (60%, Hackett et al., 2005) or Lingulodinium polyedrum (59%), but similar to Amphidinium (50%) and Crypteodinium cohnii (50%) (Codon Usage Database, Genbank Release 140.0). The 7000 K. brevis ESTs analyzed formed 5280 unigene clusters. Many gene clusters displayed single nucleotide polymorphisms (SNPs), most frequently in the third codon position such that the substitution does not alter protein sequence. Since K. brevis is haploid, the presence of SNPs suggests the presence of genes in multiple copies resulting from recent duplications. SNPs are similarly observed in L. polyedrum and A. carterae (Bachvaroff et al., 2004) and in A. tamarense (Moustafa and Bhattacharya, unpublished), suggesting that these gene duplications may be prevalent in all dinoflagellates. Few genes in dinoflagellates have been studied in detail, but those which have been occur in multiple copies, and often in tandem repeats, including peridinin chlorophyll binding (5000 copies in L. polyedrum; Le et al., 1997; Reichman et al., 2003), cyclic AMP dependent protein kinase (30 copies in L. polyedrum; Salois and Morse, 1997), and proliferating cell nuclear antigen (41 copies in Pfiesteria piscicida; Zhang et al., 2006).

To this EST collection (5' reads), an additional 5000 ESTs have been obtained from the initial K. brevis library and 10,000 ESTs were added from a “combined stress” cDNA library produced from cells grown under N- or P-limitation, cells exposed to heat, oxidative, or metal stresses, and cells in stationary phase (5' and 3' reads). All sequences have been deposited in the Genbank dbEST database. When combined with the original EST collection, the current collection of 5' ESTs clusters into 11,937 unigenes. A high rate of novel gene discovery in these libraries after 25,000 sequence reads suggests that this does not approach the entire expressed genome. An additional K. brevis EST collection developed by a second research group holds approximately 18,000 ESTs representing approximately 6000 unique genes (3' reads; D. Crawford, unpublished). When the latter ESTs are publicly released, the combined sequence data from these three datasets will comprise the largest resource of expressed sequences from any dinoflagellate.

Although ESTs provide insight into expressed genes, a genomic library is an essential tool needed to gain insight into the structure and organization of specific genes or gene clusters. Bacterial artificial chromosome (BAC) libraries are useful for this purpose because they can accommodate large insert sizes of >100 kb. In K. brevis, however, we found that BAC insert sizes >50 kb yielded few viable transformants (Van Dolah et al., unpublished). This is probably due to nucleotide modifications in K. brevis DNA, such as hydroxymethyluracil substitutions for thymidine, which can stall replication forks as the host bacterium seeks to repair this perceived error in the DNA strand (Otterlei et al., 1999). Consistent with this notion, end sequencing of the low frequency viable transformants with ~100 kb inserts identified a high proportion of plastid genes. The successful transformation of plastid sequences likely reflects the fact that the plastid does not utilize the modified bases that are present in the dinoflagellate nuclear genome. This was used to advantage to identify chloroplast-containing BAC clones for analysis of the K. brevis chloroplast genome, currently underway.

### 3.2. K. brevis genome evolution

Horizontal gene transfer (HGT) has been recognized as a major force in prokaryotic evolution, contributing up to 24% of some bacterial genomes (Nelson et al., 1999). The recent availability of genomic data for eukaryotic protists has revealed a surprising degree of HGT also in the evolution of eukaryotic protists (e.g., Andersson et al., 2007). In most cases the donor organism is the victim of engulfment by the recipient, sometimes leading to an endosymbiotic relationship. Plastids of photosynthetic eukaryotes were acquired by the endosymbiosis of a cyanobacterium (Fig. 2), whose genome was subsequently reduced by selected gene transfer to the nucleus of the host or by outright gene loss (e.g., Reyes-Prieto et al., 2006). The plastids most frequently found in dinoflagellates, containing chlorophyll c1 and peridinin as the major carotenoid, were subsequently acquired through the secondary endosymbiosis of a red alga (i.e., the chromalveolate hypothesis, see Fig. 2; Bhattacharya et al., 2004). The plastid genome of peridinin dinoflagellates is unlike that of any other eukaryote, in that it is unusually highly reduced such that it is comprised of a small number (14-16) of single-gene minicircles (Zhang et al., 1999). In contrast, the K. brevis plastid contains chlorophylls c1 + c2 and fucoxanthin, which are typical of haptophyte algae, indicating that its plastid is a replacement of the peridinin plastid via a tertiary endosymbiosis (Tengs et al., 2000; Ishida and Greene, 2002; Yoon et al., 2005). This finding raises the question of what happened to the nuclear-encoded plastid genes present in the ancestral peridinin dinoflagellate genome? A comparison of the EST libraries from K. brevis with one from the peridinin dinoflagellate Alexandrium tamarense revealed the absence of the nuclear-encoded plastid genes that had been transferred to the nucleus in the peridinin dinoflagellate (Hackett et al., 2006).
Fig. 2. The history of plastid endosymbioses in *K. brevis*. (A) The primordial algal plastid resulted from an ancient primary endosymbiosis that occurred ca. 1.5 billion years ago (Ga; Yoon et al., 2004). The cyanobacterial capture occurred in the common ancestor of the Plantae. Thereafter the common ancestor of chromalveolates (i.e., chromist + alveolate protists) putatively captured a red algal plastid through secondary endosymbiosis ca. 1.3 Ga. This was the progenitor of the broadly distributed peridinin plastid in dinoflagellates. Some dinoflagellates like *K. brevis* replaced the peridinin plastid with one of tertiary endosymbiotic origin (i.e., capture of a plastid from an alga containing a “secondary” plastid; see Bhattacharya et al., 2004). There are at least four different kinds of tertiary plastids in dinoflagellates with *K. brevis* containing one of haptophyte provenance. (B) The putative phylogeny of plastids described in (A) is supported by analysis of nuclear-encoded plastid-targeted proteins in Plantae and chromalveolates. This maximum likelihood tree, based on the analysis of a 5-protein concatenated data set (see Nosenko et al., 2006) shows the origin of these genes in the Plantae (RA, red algae + GR, green algae and land plants [the glaucophytes are missing]) ancestor from a cyanobacterial source (filled black circle, primary endosymbiosis), their subsequent origin in chromalveolates (HAP [haptophytes] + STR [Stramenopiles] + PD [peridinin dinoflagellates]) from a red algal source (filled blue circle, secondary endosymbiosis), and in *K. brevis* (FD, fucoxanthin dinoflagellate) from a haptophyte source (filled green circle, tertiary endosymbiosis). All branches have either (or both) a significant Bayesian posterior probability or a bootstrap support value >90%.

et al., 2005). These results imply that the haptophyte endosymbiosis resulted in a genome transformation in *K. brevis*, in which it lost (or silenced) the nuclear-encoded plastid genes that had been transferred from the ancestral red algal secondary endosymbiont (Yoon et al., 2005). Phylogenomic and phylogenetic analyses show that the resulting proteome of the *K. brevis* plastid is a chimeric mix of genes from red, haptophyte and green algal origin, acquired from multiple sequential endosymbioses resulting in large-scale intracellular HGT (also known as endosymbiotic gene transfer; see Li et al., 2006; Reyes-Prieto et al., 2006) from the endosymbiont to host nuclear genome (Nosenko et al., 2006), or HGT from external sources (Nosenko and Bhattacharya, unpublished). This unprecedented genomic plasticity in dinoflagellates is likely facilitated by their myxotrophic behavior, where the ability to engulf other algae and bacteria provides the opportunity to amass foreign genes in the dinoflagellate nucleus. This trait may allow for rapid adaptation to changing environmental conditions, thereby increasing their evolutionary fitness (e.g., see Nosenko et al., 2006).

4. Insights into global gene expression

4.1. Photoperiodic control of gene expression

The availability of the EST collections described above facilitated the development of a DNA microarray containing 4620 unique gene probes (Table 1; Lidie et al., 2005), which was more recently updated to include 10,263 unique gene probes from both control and stress cDNA libraries described above. Because of its central importance to cellular processes involved in *K. brevis* bloom formation, the microarray was first used to investigate global gene expression patterns under photoperiodic control (Van Dolah et al., 2007). Transcript profiles identified by microarray analysis are a composite of both changes in transcription and changes in RNA stability. This study found that 9.8% of genes were differentially expressed over the photoperiod, with most differentially expressed genes exhibiting either peak or minimum expression early in the dark. Both the overall pattern and the proportion of genes showing differential expression is consistent with a microarray study carried out in *Arabidopsis*, in which 11% of transcripts were found to be under photoperiodic control (Schaffer et al., 2001). In *K. brevis* the differentially expressed gene set had a surprisingly high representation (33%) of genes involved in post-transcriptional processing of RNA and protein turnover, as compared to *Arabidopsis* (4%). This is of interest because of the prevalence of post-transcriptional control reported for circadian processes in dinoflagellates (Morse et al., 1989; Mittag et al., 1998). Another predominant group of genes responding to light and dark were several photosystem genes and genes involved in energy acquisition. Most curiously, no change in the expression of cell cycle genes was observed under either the photoperiod (with 50% of cells cycling) or in constant light (70% of cells cycling). Given that the cell cycle is under circadian control, this raises the question of whether the cell cycle genes that are in most eukaryotes under transcriptional control are in fact post-transcriptionally regulated in *K. brevis*. Discerning these mechanisms is critically important for the development of biomarkers for bloom physiology. If gene expression is regulated post-transcriptionally, PCR-based measurements for
physiologically relevant markers will not be useful and antibody based markers may be a more productive approach.

4.2. Gene expression in response to acute stress

The ability of a *K. brevis* bloom to adapt to changing conditions in the coastal environment dictates its longevity and thus the severity of its negative impacts. Therefore, understanding the physiological basis for adaptation of *K. brevis* to changing environmental conditions is a prerequisite for understanding bloom longevity and bloom termination. Cellular responses to acute stress consist of a phylogenetically conserved gene network termed the “environmental stress response” (Gasch et al., 2000). Key components of this response include a temporary shutoff of general transcription, concurrent with the induction of transcription and translation of conserved stress proteins that serve to maintain other proteins in their properly folded configuration and to repair or recycle damaged proteins while the cell adapts to its altered conditions. Using Western blotting, Miller-Morey and Van Dolah (2004) identified several conserved stress proteins in *K. brevis* including heat shock proteins and superoxide dismutases, which were responsive to heat or oxidative stress. Their presence and their induction over a time course typical for eukaryotic cells suggested *K. brevis* possesses a typical environmental stress response network. Lidie and Van Dolah (2007a) used the *K. brevis* microarray to query global transcriptional responses to acute stresses including heat, peroxide, lead, and sodium nitrite. In all cases, a hallmark decrease was seen in gene transcripts involved in ribosome generation, translation, and energy acquisition, consistent with a general stress response that includes a transient shut-off of general mRNA transcription. However, transcription of stereotypical heat shock proteins and other stress related genes present on the array was not seen even under conditions previously shown to induce the expression at the protein level. This once again hints at post-transcriptional regulation in *K. brevis* of cellular processes typically under transcriptional control in eukaryotic organisms.

4.3. An RNA trans-splicing mechanism facilitates post-transcriptional RNA processing

The lack of transcriptional regulation of the cell cycle genes and acute stress responses suggests that *K. brevis* may utilize post-transcriptional mechanisms for regulating gene expression. Trypanosomes, kinetoplastid protists only distantly related to dinoflagellates, provide a model for how such post-transcriptional regulation may take place. Like dinoflagellates, trypanosomes lack identifiable transcriptional regulatory elements such as TATA boxes. Trypanosome RNA polymerase II transcription is constitutive and genes are transcribed in polycistronic units. To produce mature mRNAs from these pre-RNAs, a short non-coding RNA sequence termed the spliced leader (SL) is trans-spliced onto a message at a splice acceptor site located in intergenic regions of the polycistronic messages (Fig. 3). As a result, diverse mRNAs acquire a common 5’-sequence and, simultaneously, messages become polyadenylated via a separate mechanism. The presence of tandemly repeated genes in dinoflagellates and the demonstration that at least some of these are transcribed into polycistronic units (e.g., Rubisco; Zhang and Lin, 2003) suggested that a similar mechanism might exist in dinoflagellates. Indeed, Lidie and Van Dolah (2007a, 2007b) identified an identical 22 bp sequence on the 5’ end of ~100 diverse ESTs in the libraries described above. Concurrently, Zhang et al. (2007) identified the identical sequence in several species of dinoflagellates from all taxonomic orders.

![Fig. 3. The spliced leader trans-splicing mechanism in trypanosomes.](image-url)
primitive to derived, suggesting that this is a conserved feature of dinoflagellates that is absent from their nearest relatives the apicomplexans. The gene for this spliced leader was identified from K. brevis genomic DNA and found to be associated with the SS RNA gene, consistent with the location of SL genes in other organisms known to use this mechanism (Drouin and Monis de Sa, 1995). These findings suggest that K. brevis uses SL trans-splicing to process polycistrionic messages into mature mRNAs. The extent to which this mechanism is used by K. brevis or other dinoflagellates is not certain because most of the sequences in cDNA libraries do not represent full-length transcripts. However, using PCR with SL- and gene specific-primers, all K. brevis transcripts investigated to date, with the exception of plastid encoded genes, appear to include the 5’ SL sequence, including a number of the stress and cell cycle genes described above. Since its discovery a decade ago, spliced leader trans-splicing has been identified in a variety of eukaryotic protists and primitive chordates, which utilize this mechanism to varying degrees (Davis, 1996). The precise role of this mechanism in dinoflagellates has not yet been elucidated, but in trypanosomes the presence of the spliced leader appears to modulate mRNA stability or translation initiation (Maroney et al., 1995). Given the absence of nucleosomes and the compact structure of dinoflagellate chromosomes, regulation of gene expression at the post-transcriptional level may be more efficient than regulation of transcription.

5. Mechanisms regulating toxicity

K. brevis blooms are associated with extensive fish kills, marine mammal, bird, and turtle mortalities and human respiratory and food poisoning (Backer, this issue; Landsburg, this issue). These effects are attributed to the production of a suite of potent excitatory neurotoxins, the brevetoxins (PbTx, after Pychodiscus brevis). At least nine brevetoxin congeners have been characterized in K. brevis that are derived from two carbon backbones, termed A and B. PbTx congener composition changes during the course of a growth curve, with PbTx2 dominating during log phase and PbTx3 and PbTx1 increasing as cells move into stationary phase (Roszell et al., 1990). Furthermore, during log phase growth PbTx is largely intracellular, whereas in stationary phase cultures, extracellular toxin becomes more prevalent and the total toxin concentration in the culture increases (Roth et al., 2007). A similar trend is seen in natural blooms as they age (Pierce et al., 2001; Tester et al., 2008). Variation in cellular toxin quota of 10–70 pg cell⁻¹ in different clonal isolates grown under identical conditions suggests genetic variability in toxin biosynthetic capacity. Cellular toxin levels have been reported to respond to environmental conditions, with approximately a threefold increase in cultures under either N- or P-limitation relative to nutrient replete conditions (Greene et al., 2000). Similarly, light (Greene et al., 2000) and salinity (Maier Brown et al., 2006) are reported to alter cellular toxin concentrations.

Brevetoxins are ladder-like all-trans-fused polycyclic ether polypeptide compounds, a structure unique to dinoflagellates (Lin and Risk, 1981; Shimizu et al., 1986). Polypeptides are synthesized by complex enzymes called polypeptide synthases (PKSs) that, beginning with acetyl CoA, carry out successive Claisen condensations with additional acetalate groups, to produce a growing polypeptide chain. PKSs are typically classified as type I, II, or III. Early work in dinoflagellate PKSs (Snyder et al., 2003 see discussion below) suggests dinoflagellate polypeptides are synthesized by type I enzymes. Type I PKSs are large multifunctional proteins organized into “modules”, each containing multiple catalytic domains. The core catalytic domains within each module include the ketosynthase (KS), acyltransferase (AT), and acyl carrier protein (ACP) that are necessary for chain elongation, while “optional” domains include keto reductase (KR), dehydratase (DH), and enoyl reductase (ER) domains that carry out reduction of the β-keto group, and a thioesterase (TE) domain that terminates chain elongation. Each module catalyzes one Claisen condensation and various reductions or modifications depending upon what catalytic domains are present in that module. Modular type I PKSs build polypeptides in an assembly line manner where each module is used once and the growing polypeptide chain is passed to the subsequent module for the next chain elongation (Khosla et al., 1999).

Stable isotope incorporation studies found unusual truncations of acetyl groups in brevetoxin as well as methyl side chains, derived from methionine, that indicate brevetoxin biosynthesis may require atypical polypeptide synthases (Lee et al., 1986, 1989; Chou and Shimizu, 1987; Shimizu et al., 2001). The carbon truncations could be carried out on a PKS via an epoxidation followed by thioesterification, which would require the addition of epoxidase and thioesterase genes to the PKS gene or gene cluster (Shimizu et al., 2001).

The first studies to identify the genes for brevetoxin biosynthesis in dinoflagellates used degenerate PCR primers to type I and type II KS domains (Snyder et al., 2003). PCR products from multiple dinoflagellate species were homologous with known type I PKSs. Phylogenetic analysis placed three of four sequences obtained from K. brevis within a eukaryotic clade that included sequences of closely related apicomplexans (Snyder et al., 2003). The fourth sequence appeared to be of bacterial origin. Fluorescent in situ hybridization (FISH) was used to localize PKS genes following flow cytometry-sorting of cultures to separate K. brevis cells from bacterial cells. FISH localization suggested the bacterial origin of some PKS genes and dinoflagellate origin of other PKS genes previously identified through PCR strategies (Snyder et al., 2005).

Using a different approach, Monroe and Van Dolah (2008) identified eight additional PKS sequences in K. brevis through screening of the EST collections described above, including six KS domains, one KR domain and one sequence containing both ACP and KS domains. Phylogenetic analysis indicated that these sequences were clearly of eukaryotic, and not bacterial origin. Full length gene transcripts were obtained from the ESTs by 5’ and 3’ RACE (RNA amplification of cDNA ends), which demonstrated the presence of eukaryotic signatures (3’ untranslated region, poly[A] tail) and the spliced leader sequence, described above, at the 5’ terminus. These sequences are unusual in that each transcript encodes an individual domain, rather than a multi-domain protein seen in typical type I PKSs. The presence of the spliced leader suggests PKS expression may be under post-transcriptional control. This is consistent with microarray studies carried out on K. brevis under N- and P-limitation (Monroe et al., 2005) which showed no significant change in transcript abundance. The presence of separate transcripts for individual catalytic domains further suggests a novel gene structure in dinoflagellates (Fig. 4). Whereas type I PKS genes characterized to date encode multiple modules containing many domains, the gene structure suggested by the K. brevis PKS transcripts consists of separate PKS genes encoding for discrete domains. These domains therefore occur on separate polypeptide chains, more similar to the structure of type II PKSs but with sequence similarity closest to type I. Confirmation of the gene structure and organization of K. brevis PKSs now requires analysis of the genomic DNA libraries described above.
occurs has significant effects on their development and termination (Fukami et al., 1991; Doucette, 1995; Doucette et al., 1998; Skerratt et al., 2002). *K. brevis* exists in a sea of bacteria, viruses, fungi, and other microbes, all of which may exert an influence on its growth and toxicity either directly or indirectly. Of these interactions, the roles of bacteria in regulating bloom dynamics have received the most attention over the past decade. Bacterial communities associated with algal blooms are not simply random assemblages of marine bacterial species, but rather have a characteristic profile, commonly dominated by alpha-Proteobacteria, particularly those of the *Roseobacter* clade, some of which have been proposed to have growth promoting capability (Jasti et al., 2005). Another constituent, the gamma-Proteobacteria, has been alternatively attributed to both growth stimulation and bloom decline in different situations (Kodama et al., 2006). Members of the beta-Proteobacteria are encountered less commonly, but dominate the intracellular flora of *Gymnodinium instriatum* (Alverca et al., 2002).

A fourth group, the *Cytophaga–Flavobacterium* (CFB) clade, is also frequently represented within this community. Many CFB taxa occur in association with particles in the marine environment and are active in the remineralization of high molecular weight dissolved organic matter (Kirchman, 2001), which may lead to algal growth promotion. Conversely, certain members of the CFB are also included among the approximately 60 algicidal bacteria identified to date (Fukuyo et al., 2002; Mayali and Azam, 2004).

Natural bacterial populations have also been documented to produce iron binding ligands known as siderophores, which are capable of making this frequently growth-limiting trace metal available to phytoplankton (Soria-Dengg et al., 2001).

Many marine bacterial species are unculturable; therefore, understanding the composition and associated functional ecology of these microbes is being facilitated by the development of 16S rDNA community libraries. To date, one such library has been developed to date for *K. brevis*, including 165 rDNA sequences from several depths in waters containing various concentrations of *K. brevis* cells, as well as areas where this dinoflagellate is undetectable (Mikulska et al., unpublished data). This library shows a clear dominance of alpha-Proteobacteria (primarily members of the Rhodobacterales, including *Roseobacter* spp.) and a frequent presence of the Bacteroidetes when *K. brevis* cells are present in the water column at bloom concentrations. Interestingly, preliminary analysis of the library also suggests a potential relationship of bacteria capable of DMSP utilization with *K. brevis* presence, the latter well-known as a DMSP producer. While providing the first molecular-based snapshot of a *K. brevis* associated bacterial consortium, it is clear that data on the spatio-temporal variability and the functional attributes of dominant taxa present in bloom associated microbial communities are needed to fully understand how they may influence bloom progression.

The strong representation of Bacteroidetes in the 16S rDNA library from *K. brevis* bloom water is of interest because many of the known algicidal bacteria are members of this group, primarily *Cytophaga* (nomenclature under revision) and Flavobacterium. Representatives of both *Cytophaga* (strain 41-DBG2; Doucette et al., 1999) and the Flavobacteriaceae (strain S03; Roth et al., 2008a) isolated from Gulf of Mexico waters have been demonstrated to exhibit algicidal activity against *K. brevis*. Characterization of algicidal activity demonstrated that the latter (S03) required direct contact with *K. brevis* in order to elicit its algicidal effects while the former (41-DBG2), which produced a dissolved, heat labile algicide(s) (Twiner et al., 2004), did not. Regardless of the mode of algicidal attack, lysis of *K. brevis* cells by these two bacterial strains caused a release of dissolved brevetoxin (ca. 60% of the total toxin) into the surrounding water. The majority of this dissolved toxin comprised open A-ring PbTx-2 and ~3 derivatives that were ~100-fold less potent than their respective parent toxins (Roth et al., 2007). These findings suggest that attack of *K. brevis* by algicidal bacteria can lead to a reduction in toxicity and affect the fate and trophic transfer routes of brevetoxins during bloom events. Amongst the *K. brevis* clones tested against algicidal strains of bacteria, both susceptible and resistant clones have been identified. Resistance to attack was determined to be a function of the ambient bacterial flora present rather than an inherent trait of the dinoflagellate clone (Mayali and Doucette, 2002). Recently, a bacterium capable of preventing algicidal strains from attaining the threshold density needed to induce algal cell lysis (ca. 10^6 cells ml^-1) was isolated from a resistant *K. brevis* culture and identified as a member of the Flavobacteriaceae (Roth et al., 2008b). This work was the first to highlight the complex nature of microbial interactions and their potential consequences to *K. brevis* bloom dynamics.

Viruses have also been implicated in the lysis of dinoflagellates and termination of blooms (for review, Salomon and Imai, 2006). Viruses infecting microalgae are usually host specific, such that their abundance peaks during blooms, but falls to undetectable levels during non-bloom periods. Paul et al. (2002) isolated a heat labile “filterable lytic agent” from 0.2 μm filtered *K. brevis* bloom water that caused lysis of *K. brevis* cultures. Following cell lysis, high numbers of virus-like particles (4–7 x 10^9 ml^-1) were present in the culture supernatant. However, no identifiable virus particles were found in association with lysing cells; thus, the involvement of lytic viruses in *K. brevis* bloom termination remains an unsolved question.

### 7. Population genetics of *K. brevis*—what constitutes a bloom?

The genetic diversity within and among blooms of *K. brevis* is of significance in understanding both the origin of blooms that emerge in novel locations and the mechanisms driving the evolution of this species. The life cycle of *K. brevis* has not been entirely resolved (Steidinger et al., 1998); however, the presence of a sexual cycle is evident from the existence of planozygotes. Nonetheless, a resting cyst stage has not been documented, so the rate of genetic recombination remains entirely unexplored. In dinoflagellates, high numbers of virus-like particles (4–7 x 10^9 ml^-1) were present in the culture supernatant. However, no identifiable virus particles were found in association with lysing cells; thus, the involvement of lytic viruses in *K. brevis* bloom termination remains an unsolved question.

The majority of this dissolved toxin comprised open A-ring PbTx-2 and ~3 derivatives that were ~100-fold less potent than their respective parent toxins (Roth et al., 2007). These findings suggest that attack of *K. brevis* by algicidal bacteria can lead to a reduction in toxicity and affect the fate and trophic transfer routes of brevetoxins during bloom events. Amongst the *K. brevis* clones tested against algicidal strains of bacteria, both susceptible and resistant clones have been identified. Resistance to attack was determined to be a function of the ambient bacterial flora present rather than an inherent trait of the dinoflagellate clone (Mayali and Doucette, 2002). Recently, a bacterium capable of preventing algicidal strains from attaining the threshold density needed to induce algal cell lysis (ca. 10^6 cells ml^-1) was isolated from a resistant *K. brevis* culture and identified as a member of the Flavobacteriaceae (Roth et al., 2008b). This work was the first to highlight the complex nature of microbial interactions and their potential consequences to *K. brevis* bloom dynamics.

Viruses have also been implicated in the lysis of dinoflagellates and termination of blooms (for review, Salomon and Imai, 2006). Viruses infecting microalgae are usually host specific, such that their abundance peaks during blooms, but falls to undetectable levels during non-bloom periods. Paul et al. (2002) isolated a heat labile “filterable lytic agent” from 0.2 μm filtered *K. brevis* bloom water that caused lysis of *K. brevis* cultures. Following cell lysis, high numbers of virus-like particles (4–7 x 10^9 ml^-1) were present in the culture supernatant. However, no identifiable virus particles were found in association with lysing cells; thus, the involvement of lytic viruses in *K. brevis* bloom termination remains an unsolved question.

The genetic diversity within and among blooms of *K. brevis* is of significance in understanding both the origin of blooms that emerge in novel locations and the mechanisms driving the evolution of this species. The life cycle of *K. brevis* has not been entirely resolved (Steidinger et al., 1998); however, the presence of a sexual cycle is evident from the existence of planozygotes. Nonetheless, a resting cyst stage has not been documented, so the rate of genetic recombination remains entirely unexplored. In dinoflagellates, high numbers of virus-like particles (4–7 x 10^9 ml^-1) were present in the culture supernatant. However, no identifiable virus particles were found in association with lysing cells; thus, the involvement of lytic viruses in *K. brevis* bloom termination remains an unsolved question.

Please cite this article in press as: Van Dolah, F.M., et al., The Florida red tide dinoflagellate *Karenia brevis*: New insights into cellular and molecular processes underlying bloom dynamics. Harmful Algae (2009), doi:10.1016/j.hal.2008.11.004
A number of population genetic markers have been assessed in *K. brevis* in the past decade. Ribosomal large subunit (LSU) probes from the D1, D2, and 28S hypervariable regions were found to be identical between geographically distinct populations of *K. brevis*, making them unsuitable for assessing population level variability. However, they are unique to *K. brevis*, making them useful for identification at the species level, which is particularly useful in mixed bloom assemblages (Mikulski et al., 2005; Goodwin et al., 2005). Similarly the ribosomal small subunit (SSU) 18S coding regions and ITS sequences were found to be identical among isolates from Texas and Florida displaying phenotypic differences (Loret et al., 2002). In contrast, microsatellite markers, short stretches of DNA containing tandemly di- and tri-nucleotide repeat motif of variable length, have been successfully used to uncover genetic diversity among phenotypically distinct clonal isolates. Nine polymorphic loci have been characterized and two to six alleles found at each locus. For a sample of 13 *K. brevis* cultures (five from the same 1999 bloom sample), gene diversity ranged from 0.15 to 0.75 (Renshaw et al., 2006). This variability may suggest that bloom populations are genetically highly diverse, or this genetic pool may provide different subpopulations that bloom preferentially in response to different environmental conditions. To address this question, Campbell et al. (2004) have applied microsatellite markers to characterize the variability among clonal cultures isolated from a bloom off the Texas coast. More recently, the advent of single cell genotyping permits the evaluation of genetic variability within a bloom without the labor-intensive establishment of clonal isolates (Henrichs et al., 2007). This approach is currently being used to examine the temporal patterns of genetic diversity in bloom populations of *K. brevis* collected weekly during the 2005 and 2006 blooms off the Texas coast.

### 8. Future directions

The past decade has produced remarkable progress in the cellular and molecular biology of *K. brevis* and other dinoflagellate species. Yet for each topic discussed above, our insights remain rudimentary. Nonetheless, the application of the genomic resources developed over the past decade will undoubtedly engender rapid discoveries over the next few years that will contribute exponentially to our understanding of bloom dynamics. For example, our current insight into the nuclear genome of *K. brevis* is based largely on inference from the expressed sequences. We anticipate that the analysis of genomic libraries will yield substantial new information in rapid order. Improved technologies for high throughput DNA sequencing (e.g., 454, pyrosequencing) make genomic sequencing feasible that only a few years ago was unimaginable. The prevalence of post-transcriptional gene regulation in *K. brevis* suggests that novel mechanisms may regulate key processes involved in the growth of dinoflagellate blooms. The identification of these regulatory mechanisms in *K. brevis* will not only aid in our fundamental knowledge of *K. brevis* biology, but will also provide biomarkers that may be deployed on remote moorings associated with coastal ocean observing systems to better monitor the physiological status of blooms, and thereby improve upon existing bloom forecasting capabilities, including sensing conditions predictive of termination. The recent insights gained into the interactions of *K. brevis* with the microbial community suggest that the way we think of blooms should be reevaluated. In fact, the bloom of one species reflects a community response to a change in environmental conditions. Metagenomic characterization of the microbial community (e.g., bacteria, viruses, picoplankton, protists) and metabolomic analysis of bloom and pre-bloom waters may provide a powerful approach to address the questions heretofore intractable to traditional oceanographic approaches that have been applied. Can a metagenomic or metabolomic signature identify conditions conducive to bloom initiation or bloom termination? Finally, the application of the recently developed population genetics tools will shed light on genetic variation within and between blooms, elucidate the contribution of *K. brevis* and other recently identified *Karenia* species to blooms in the Gulf of Mexico, aid in defining the origin of blooms in different geographic regions, and potentially the assist in characterizing the life cycle of *K. brevis*.

### References


Please cite this article in press as: Van Dolah, F.M., et al., The Florida red tide dinoflagellate *Karenia brevis*: New insights into cellular and molecular processes underlying bloom dynamics. Harmful Algae (2009), doi:10.1016/j.hal.2008.11.004

F.M. Van Dolah et al. / Harmful Algae xxx (2009) xxx–xxx 9

G Model

HARALG-488; No of Pages 11


Kamykowski, D., Steidinger. MS Thesis. University of South Florida, St Petersburg, FL.


