A sensitive genetic-based detection capability for *Didymosphenia geminata* (Lyngbye) M. Schmidt: Phases Two and Three

CBER Contract Report 62

Client report prepared for MAF Biosecurity New Zealand

by

S. Craig Cary[1,2]

Brendan J. Hicks[1]

Kathryn J. Coyne[2]

Andreas Rueckert[1]

Chrissen E. C. Gemmill[1]

Catherine M. E. Barnett[1]

Independently reviewed by:

Susanna A. Wood[3]

[1] Centre for Biodiversity and Ecology Research
Department of Biological Sciences
School of Science and Engineering
The University of Waikato, Private Bag 3105
Hamilton 3240, New Zealand

[2] University of Delaware
College of Marine and Earth Studies
Lewes, Delaware 19958, USA

[3] Cawthron Institute
Private Bag 2, Nelson 7042,
New Zealand

October 2007

The University of Waikato
Te Whare Wānanga o Waikato

CBER University of Waikato
Centre for Biodiversity and Ecology Research

Table of contents

Tables

Figures

Appendices

Page

Reviewed by: *Approved for release by*:



Susanna A. Wood and Christina M.C. Vieglais Frances G. Velvin

Cite as:
Cary, S.C., Hicks, B.J., Coyne, K.J., Rueckert, A., Gemmill, C.E.C., Barnett, C.M.E. 2007. A sensitive genetic-based detection capability for *Didymosphenia geminata* (Lyngbye) M. Schmidt: Phases Two and Three. *CBER Contract Report 62*. Client report prepared for MAF Biosecurity New Zealand. Centre for Biodiversity and Ecology Research, Department of Biological Sciences, School of Science and Engineering, The University of Waikato, Hamilton, New Zealand. 87 p.

Executive summary

Ongoing surveillance to determine the distribution of *Didymosphenia geminata* (Lyngbe) M. Schmidt in the South Island and possible incursion into the North Island is a key component of MAF Biosecurity New Zealand's response to this non-indigenous, invasive freshwater alga. The present method for identifying *D. geminata* in water bodies involves the collection of algal scrapings and drift net-filtered water samples, their transport to a laboratory and the identification of *D. geminata* frustules under a microscope. However, microscopy lacks sensitivity at low cell concentrations, especially in samples containing extensive detrital material. It is also a time consuming, labour intensive and subjective method that restricts our ability to sample a wide range of rivers and sites within each river. The development of a highly specific and sensitive technique could dramatically increase our ability to analyse large numbers of samples in a short period of time and to detect the alga at low levels, thus increasing the efficiency of our surveillance programme without greatly increasing costs. In addition, a low-level detection capability would enable better targeting of public awareness messages and increase the probability of successful control, should a control tool be developed. This study sought to develop a sensitive DNA-based protocol capable of detecting *D. geminata* in environmental samples with extreme sensitivity. Our specific requirements for this new methodology were as follows:

- Robust field capabilities from collection to quantification
- Species or strain level specificity that has been environmentally validated
- Extreme sensitivity for low-level detection
- A broad dynamic range of detection (> 5 orders of magnitude)
- The highest degree of reproducibility
- An efficient, cost-effective, rapid, high-throughput laboratory capability.

In addition to developing a DNA-based surveillance method, a secondary objective was to leverage the molecular information acquired while developing the DNA surveillance method towards a phylogeographic study to shed light on the possible origins of the *D. geminata* population in New Zealand. As proposed, the entire DNA detection research programme was divided into three discrete phases of six months each. Phase 1, as reported earlier by Cary et al. (2006), was challenged with 1-1) designing and testing field collection and DNA extraction methods, and 1-2) identifying a gene target to discriminate *D. geminata* from other endemic taxa and develop a highly specific gene amplification protocol based on the polymerase chain reaction (PCR). Phase 2 (reported here) was challenged with 2-1) taking this PCR amplification method to a more sensitive and specific quantitative and fully validated level (i.e., QPCR), and 2-2) beginning the first comprehensive analysis of the phylogeography of *D. geminata*. During Phase 3 (also reported here), we 3-1) adapted the new detection protocols to achieve a field-based high-throughput capacity, 3-2) developed a rigorous quality control and quality assurance protocol, and 3-3) developed a manual for training our end-users in the use of the protocols and the application of the results.

Two different QPCR methods were developed and tested for *D. geminata* using the PCR primers developed and validated in Phase 1. The SYBR Green I QPCR assay directly monitors the incorporation of a DNA intercalating dye as the gene amplification process progresses. The Taqman QPCR assay also relies on a fluorescent signal being produced progressively each cycle of the PCR; however, the fluorescence is generated by cleavage of a fluorescent reporter dye from a target-specific (*D. geminata)* DNA probe which lies within the priming site of the forward and reverse primers. Each of the methods were rigourously tested and validated with control DNA and numerous environmental samples.

The SYBR Green I QPCR proved to be very sensitive over a wide range of *D. geminata* plasmid template concentrations. Moderate amounts of exogenous non-template DNA did not noticeably affect the sensitivity of the QPCR. The demonstrated lower limit of detection of the SYBR Green 1 assay corresponds to approximately 40 target sequences or ~1.5 *D. geminata* cells. The limitation of the SYBR Green QPCR was the generation of false positive amplification signals for negative control reactions (no added template) and for reactions containing template DNA below the equivalent of 1.5 *D. geminata* cells.

The robustness and specificity of the SYBR Green I QPCR assay was improved in the Taqman assay using a dual-labelled fluorescent probe specific to *D. geminata*. The primer and Taqman probe combination designed for *D. geminata* proved to be both sensitive and robust, with detection of approximately 30 copies (~0.1-1 cell) of the 18S ribosomal RNA gene from *D. geminata,* and rarely produced false positive amplification signals. The method was tested with environmental samples and found to be extremely sensitive, in some cases detecting *D. geminata* in samples where microscopy had reported a negative. From these analyses we have set the lower end of detection (Sensitivity Threshold) to be 1 pg of our calibrator sample which is equivalent to ~1.0-0.1 *D. geminata* cells in the QPCR Taqman reaction.

Although the Taqman QPCR method appears to be extremely sensitive and specific to *D. geminata*, the specificity of the method to detect only *D. geminata* in New Zealand waters required extensive validation. The primers and probe therefore underwent a robust 3-tiered validation procedure consisting of 149 positive QPCR amplification products being validated by gel electrophoresis (PCR amplicon length), high resolution melt (HRM) analysis (heat denaturation characteristics), and direct DNA sequencing of the QPCR product. To date, we have examined 174 samples from 76 rivers by QPCR, 54 from the South Island, 8 from the North Island, and 12 international (Norway 4, Canada 2, England 1, and USA 5; Table 3.1). 149 samples were positive for *D. geminata* (i.e., above threshold). All of these positive samples have been validated to be *D. geminata* using all three validation procedures. No North Island samples were positive for *D. geminata*. These results indicate that the Taqman QPCR method is specific for *D. geminata.*

The SYBR Green I QPCR is not the preferred method for surveillance efforts where low-level sensitivity and extreme specificity is needed. The use of *D. geminata*-specific primers and a probe in the Taqman QPCR assay provides greatly increased specificity and sensitivity and is better suited as a surveillance tool to determine the presence or absence of *D. geminata* from areas where it was not previously known. The Taqman assay also allows the ability to enumerate the number of cells present in a given sample, which could be useful for measuring the effectiveness of mitigation efforts.

The goal of the phylogeography study was to identify genetic markers with appropriate levels of variation to reconstruct phylogeographic patterns among samples of *D. geminata*. These markers would enable the identification of the origin(s) of this invasive organism and help in understanding the pattern(s) of genetic variation within and among samples in New Zealand. They would also indicate if multiple incursions of *D. geminata* have occurred in New Zealand. Only preliminary phylogeographic work was expected to be completed within the short time-frame available and the competing higher priority of developing a DNA-based detection capability.

We examined two potential markers: the 18S ribosomal RNA gene (18S rDNA) and its associated internal transcribed spacer regions (ITS). Generally the 18S can discriminate at the generic level and the ITS at the species or strain level. Preliminary results indicate that the partial 18S+ITS region provides a sufficient level of resolution to reveal the phylogeographic history and origin(s) of *D. geminata* in New Zealand. From the limited number of samples that have been analysed, it appears more likely that the New Zealand *D. geminata* originated

from an incursion from North America rather than Europe. This study can continue with the addition of many more samples now that the genetic marker has been identified.

The sensitivity of the DNA detection method mandates that strict protocols be adhered to from sample collection in the field to when they are analysed in the laboratory. We have been directly involved in the design and development of the drift nets and accompanying deployment protocols currently in use by MAF Biosecurity New Zealand for their delimiting surveys. The net specifically developed to concentrate *D. geminata* has the potential to be a primary source of site-to-site DNA contamination. Specific protocols for net DNA disinfection using household bleach are presented that are shown to remove all traces of *D. geminata* DNA that might be detected by the DNA method.

The purpose of an effective incursion detection surveillance strategy is to identify *D. geminata* at an early stage rather than when high biomass has accumulated at a site. This would allow mitigation efforts the best chance of reducing the impacts of *D. geminata*. In any given river, sample coverage and frequency will obviously affect the likelihood of detection. We have tested 6 different commercially available DNA extraction methods that have the ability to be "scaled up" to a high throughput level and potentially able to handle 50-100 samples in 48 hours. Each of these methods went through rigorous testing with direct comparison to our own highly optimized laboratory method.

One of the primary goals of this project was to transfer technologies to existing MAF Biosecurity New Zealand surveillance programmes and other potential end users to improve or enhance their surveillance efforts. As part of this objective, we have prepared a detailed manual of protocols and techniques developed during the project for the collection, extraction and QPCR analysis of *D. geminata*. We are prepared to offer *D. geminata* DNA method workshops at the request of interested parties.

After 18 months of research and development, we believe the DNA detection method has been adequately tested and validated and is now ready for full implementation. The method, while extremely robust, will continue to be tested and assessed to maintain the highest level of quality control and quality assurance possible. Because many high-risk waterways in New Zealand have yet to be tested with the DNA method, we plan to continue to fully validate (via electrophoresis gel, HRM, sequence) any QPCR positive sample we obtain. As new sequences from more closely related diatoms are obtained by us or through international databases, we plan to reassess our current method, and adapt it if necessary.

We recommend that the Taqman QPCR DNA detection method become the primary surveillance tool for the North Island of New Zealand, where *D. geminata* has yet to be detected. The power of the Taqman QPCR method lies in its negative predictive value. The sensitivity and specificity of the method means that any negative result indicates that the water body is highly likely to be free of *D. geminata*. Also, any positive results with the DNA detection method can be validated with other molecular methods (electrophoresis gel, HRM, sequence) or by the microscopic method, to rule out contamination and provide high confidence in the extreme low level detection capability.

Detecting *D. geminata* as early as possible, should it spread to the North Island, will increase the chance of a successful elimination attempt, should a control tool be developed. Early detection will also enable containment measures to achieve their maximum effectiveness by facilitating early communication of the positive status of a river. Having the knowledge that a river is affected before the organism is visible should enable MAF Biosecurity New Zealand and their *D. geminata* long-term management partners to target their social marketing campaigns towards freshwater users who are especially at risk of spreading the organism, and whose actions to slow the spread will reduce impacts to other river systems.

1 Introduction

We undertook an 18-month (3-phase) programme to develop, validate and deploy a genetic-based method for high-throughput detection and enumeration of *Didymosphenia geminata* (Lyngbye) M. Schmidt. This report is a direct extension from our previous interim report (Cary *et al.,* 2006) and addresses the subsequent progress in the development and validation of two DNA detection methods for *D. geminata*. This report is divided into 8 primary sections. Section 1 is an introduction linking the previous interim report to the final report and introduces the milestones and objectives to be reported in each of the subsequent sections. Sections 2-7 address the work conducted and results for each of our specific objectives in phase 2 and 3 including development and validation of the DNA method, advances in *D. geminata* phylogeography, development of specific sampling and high throughput methods for the DNA method, and planned methods for implementing technology transfer. Section 8 concludes the report addressing specific aspects of risks, quality control and quality assurance, and future recommendations for implementation and continuous improvement of the new detection method.

1.1 Overview of project and linkage to Phase 1

Our objectives for the first phase of the project were to develop and test sample collection strategies and extraction methods, and to design, test, and implement a low resolution genetic-based detection capability for *D. geminata*. During this period we thoroughly investigated and compared methods of stabilisation and extraction of high quality DNA from riverine samples. We also sequenced the 18S subunit of the ribosomal RNA gene (18S rDNA) of *D. geminata* and designed and tested a number of polymerase chain reaction (PCR) primer sets for their specificity to the *D. geminata* 18S rDNA. This process was essential in order to achieve the level of sensitivity and reproducibility required for this project. In Phase 2, we improved and validated the capabilities of the detection method using two PCR-based gene amplification methodologies to attain the highest sensitivity and specificity possible. During this phase we also began a global phylogeographic study to establish population level genetic markers that we hope will eventually allow us to determine the origin of *D. geminata* in New Zealand and the likelihood of whether there have been multiple incursions. In Phase 3, we advanced the field collection protocols to meet the needs of the quantitative PCR (QPCR) protocols and adapted the QPCR technologies to a high throughput process that allows hundreds of samples to be processed per week. We also describe a *D. geminata* collection and analysis manual and workshop agenda to transfer the technology to MAF Biosecurity New Zealand and our end user groups.

1.2 Overview of phase 2

Phase 2 addressed all aspects of the design and validation of a laboratory-based, high throughput QPCR detection and enumeration assay for *D. geminata*. It also included a preliminary survey of potential genetic markers in *D. geminata* to ultimately provide for a phylogeographic study to be conducted.

1.2.1 Quantitative PCR (QPCR) protocol development

Two different QPCR methods, SYBR Green I and Taqman, were developed for *D. geminata* using the primers developed and validated in Phase 1, and compared for their relative sensitivity for *D. geminata* detection.

1.2.1.1 The SYBR Green I QPCR assay

This PCR-based assay uses the *D. geminata* specific PCR primers developed in Phase 1 in concert with an intercalating fluorescent reporter dye. During the PCR reaction, as the *D. geminata* 18S rRNA gene is amplified, the SYBR Green I dye intercalates the new DNA and becomes photo excitable. When used in a QPCR thermal cycler, a powerful light emits at the right wavelength to excite the intercalated dye at a specific time during each cycle. Thus, as new DNA is produced by PCR, more dye is incorporated and the fluorescent signal increases proportionally. The emitted fluorescence is detected by a sensor and reported. Once the assay is complete, the amplified DNA can then be melted to separate the two DNA strands. The way the DNA melts will depend upon the order and composition of the bases in the PCR product, thus, different sequences will have a characteristic melt "curve". The number of peaks observed and at what temperature those peaks occur can be used diagnostically and thus further confirm the presence of *D. geminata*. We developed, tested and validated a SYBR Green I assay for the detection of *D. geminata* in environmental samples. Compared to the Taqman assay, SYBR Green 1 was not as suitable for surveillance of *D. geminata* due to its lower sensitivity, but the SYBR Green I tool could be useful for certain applications such as monitoring low levels of existing populations of *D. geminata* where the interest was in determining subtle and immediate growth responses to environmental conditions that might promote or limit *D. geminata* growth in the early stages of establishment.

1.2.1.2 The Taqman QPCR assay

The Taqman method is an extremely powerful and more specific alternative to the SYBR Green I to measure PCR product formation. Similar to the SYBR Green I assay, the Taqman assay relies on a fluorescent signal being produced progressively each cycle of the PCR. However, in the case of the Taqman assay, the fluorescence is generated by cleavage of a fluorescent reporter dye from a target-specific (*D. geminata)* DNA probe which lies within the priming site of the forward and reverse primers. The use of the two *D. geminata* specific primers (Phase 1) in conjunction with a new *D. geminata*-specific probe is designed to greatly increase the specificity and sensitivity of the detection reaction. The threshold cycle number (C_T) for each reaction is the cycle number at which the fluorescence emission changes significantly over baseline levels. By plotting the C_T values vs. known starting concentrations of standards, the concentration of *D. geminata* DNA and ultimately number of individual cells in a given sample can be determined by linear regression analysis. The QPCR method allows cell concentrations to be obtained on samples to a much finer resolution than microscopy when the samples have low cell numbers or contain large amounts of sediment or organic matter. This should aid in surveillance of new incursions, and will assist in monitoring the results of any mitigation efforts.

1.2.2 Validation of DNA Method specificity

The original *D. geminata*-specific QPCR primers and probe described above were designed based on a limited number of known diatom DNA sequences. It is therefore conceivable, although highly unlikely, that an unknown *D. geminata* "look-alike" might exist in nature and cross-react with the primers and probe resulting in a false positive. The primers and probe therefore underwent a robust 3-tiered validation procedure consisting of over 100 positive QPCR amplification products being validated by gel electrophoresis (PCR amplicon length), high resolution melt analysis (heat denaturation characteristics), and direct sequencing of the QPCR product. Explanation of each of these validation techniques is reported along with a summary of the validation results.

1.2.3 Preliminary phylogeographic study

At the current time we do not know how or from where *D. geminata* first entered New Zealand. Early on in the 2004 incursion response, it was decided that the most likely pathway for introduction was a human carrying contaminated aquatic recreational gear (Kilroy 2004). Later studies on the survivability of *D. geminata* cells removed from a river environment and left untreated showed that cells could remain viable for several months under optimum conditions of cool, moist, low-light (Kilroy et. al. 2006), indicating that a range of vectors, countries of origin and transport scenarios could have resulted in live didymo cells being accidentally carried into New Zealand. No further work has been done on the matter of determining the precise pathway or origin of introduction because it would not have aided incursion management decision-making at the time. However, now that *D. geminata* has been found in many more rivers, the issue of whether there have been multiple introductions, or only one, which is now spreading, has become a central question.

The rationale for wanting to know more about *D. geminata* genetic variation within the New Zealand populations is two-fold. It would help determine the geographical origin of the population which in turn would help determine the pathway by which the organism was likely introduced so that we can analyse strengthening border controls to reduce further introductions. It may also help with the investigation of potential sources of effective biocontrol agents.

There are two fundamental questions being addressed in the current study:

1) Is the genetic variability in the *D. geminata* populations WITHIN New Zealand rivers similar to the variability in the populations BETWEEN New Zealand rivers? This will help us determine if New Zealand has had multiple introductions or a single introduction which is spreading.

2) Is the genetic variability between *D. geminata* populations from the different geographic regions around the world sufficient to be able to develop a molecular marker to distinguish between various populations, and if so, from which geographical population(s) are the New Zealand population(s) most likely to have originated from?

1.3 Overview of Phase 3

Our primary objectives for this phase of the project were to (i) design a DNA-compatible long-term surveillance capability; (ii) develop, test and implement a rapid field compatible DNA-capability, and (iii) transfer the technology to MAF Biosecurity New Zealand personnel and other potential end users through a series of workshops and a field manual. In addition, we continued to rigorously test and refine the detection and enumeration protocols in both the laboratory and field as described above.

1.3.1 Field-based QPCR processing protocol

The sensitivity of the DNA detection methods mandates that strict protocols be adhered to from sample collection in the field to when they are analysed in the laboratory. We have been actively involved in the design and development of the drift nets and accompanying deployment protocols currently in use by MAF Biosecurity New Zealand for their delimiting surveys. The net was specifically developed to concentrate *D. geminata*, yet

has the potential to be a primary source of site-to-site DNA contamination. Specific protocols for net DNA disinfection and subsequent sampling are reported.

1.3.2 High throughput laboratory protocol

The purpose of an effective incursion detection surveillance strategy is to identify *D. geminata* at an early stage rather than when high biomass has accumulated at a site. This would allow mitigation efforts the best chance of reducing the impacts of *D. geminata*. In any given river, sample coverage and frequency will obviously affect the likelihood of detection. The DNA method that has now been developed has the ability to detect *D. geminata* at extremely low concentrations and is directly adaptable to a high throughput format allowing up to a hundred samples to be processed in a given day. For this milestone we have scaled up the methodology by coupling new high throughput sample processing strategies. We have worked closely with several vendors to determine the most cost effective processing pipeline for *D. geminata.*

1.3.3 Transfer technology to MAF Biosecurity and other end users

One of the primary objectives of this project was to transfer technologies to existing MAF Biosecurity New Zealand surveillance programmes and other potential end users to improve or enhance their detection efforts. As part of this objective, we have prepared a detailed manual of protocols and techniques developed during the project for the collection, extraction and QPCR analysis of *D. geminata*. We are prepared to offer *D. geminata* DNA detection method workshops at the request of interested parties. There is a particular need for hands-on training with all aspects of the sampling methodology, with the aim to minimize the incidence of contamination and hence reduce false positives which are easily registered by the highly sensitive DNA method..

1.4 Conclusions and recommendations

The report ends with a summary of the conclusions from each section highlighting specific areas of accomplishment and areas needing continued research. We also provide recommendations on the immediate need for implementation of the DNA method and possible areas requiring intensified sampling effort. We end the document with a risk analysis identifying areas in sample collection and analysis where contamination could occur and methods we have employed or recommend to reduce these risks to maintain high quality control and quality assurance

2 QPCR protocol development

2.1 General introduction

2.1.1 Tool criteria

The development and use of sensitive molecular genetic tools to detect and enumerate different harmful algal species directly from environmental samples has recently been reported (Coyne et al., 2005; Handy et al., 2006; Popels et al., 2003; Saito et al., 2002). In most cases, these methods have been adopted by monitoring or surveillance programs to enable the detection of the target organism at levels unattainable using standard microscopic tools. In all cases, these methods involve the use of gene-specific amplification technologies, i.e. polymerase chain reaction (PCR), that detect diagnostic genes in the organism of interest. Amplification reactions can now be quantified by quantitative PCR (QPCR) methods, either directly using a DNA intercalating fluorescent (SYBR Green I) dye that reports the production of the amplicon during the PCR process, or indirectly by monitoring the destruction of a fluorochrome-labelled species-specific DNA Taqman probe during the PCR. In both cases the eventual concentration of the amplicon and therefore the concentration of target gene (or cell number) within a sample can be calculated. In spite of their promise, however, the use of molecular methods can be costly and technically challenging to develop and implement. Each new assay developed must be backed up with considerable research and knowledge of the species and environment in which it is to be used.

Our approach in the development of these methods for *D. geminata* was to follow the successful design strategies developed by the Cary/Coyne laboratory at the University of Delaware for QPCR detection of twelve other harmful algal bloom species. Our specific requirements for these methods are that they:

1) have species or strain level specificity and have been environmentally validated,
2) are extremely sensitive for low level detection (<1 cell /mL),
3) have a broad dynamic range for enumeration (> 5 orders of magnitude),
4) have the highest degree of reproducibility,
5) are efficient, cost-effective and rapid, with a high throughput laboratory capability
6) can be adapted easily to standard field operations.

In this section, we describe the current QPCR processing pipeline for environmental samples, give a more detailed overview of the two QPCR methods we compared, and then report on the design, development and preliminary testing of each method on plasmid control and environmental samples. In addition, modifications to the extraction method reported in the interim report (Cary et al., 2006) are also included in this section.

2.1.2 Analysis strategy

The flow chart below diagrams the processing pipeline that environmental samples for assessment by QPCR follow (Figure 2.1). Our robust sampling protocol includes 3-fold redundancy in archiving and all of the needed controls to maintain the high quality control and quality assurance required for the assay. In this section of the report, we address the development of the detection and enumeration QPCR protocols aspects of the pipeline. For each of the techniques we report on the optimization of the assay, testing of assay specificity and sensitivity, and application with selected environmental samples. We also define the lower range of detection (threshold) of the DNA method to be used for future surveillance purposes.

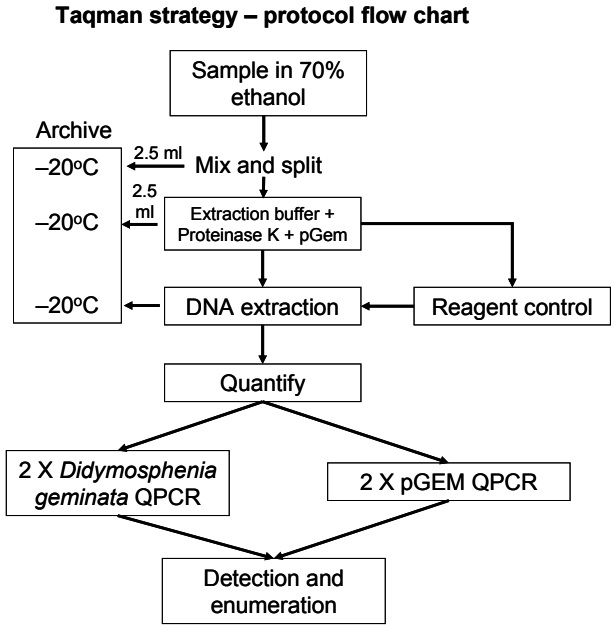


Figure 2.1 Processing pipeline that environmental samples would follow for Taqman analysis.

2.2 A QPCR primer – theory and practice

2.2.1 Two QPCR techniques - SYBR Green I and Taqman Probe

QPCR is an extremely powerful and efficient method to measure PCR product formation in real-time and has the ability to quantitatively identify the amount of starting nucleic acid target in the reaction. QPCR has the additional benefits of improved sensitivity and specificity when compared to normal end-point analysis in conventional PCR. QPCR uses real-time detection to measure the change in product concentration at several points during each PCR cycle (Heid *et al*., 1996). The increase in product concentration is measured as an increase in fluorescence emission (ΔR_n). Fluorescence is generated either by cleavage of a fluorescent reporter dye from a template-specific oligonucleotide probe or by a DNA intercalating dye. The threshold cycle number (C_T) for each reaction is the cycle number at which the fluorescence emission changes significantly over baseline levels. By plotting the C_T values versus known starting concentrations of standards, the concentrations of target DNA for each unknown sample can be determined by comparison to a standard or "calibrator" sample.

In the simplest reaction setup, the reporter dye is the nucleic acid specific SYBR Green I. SYBR Green I binds only to double-stranded DNA and upon excitation, emits fluorescence. During PCR, DNA is double-stranded only in the elongation step, when the new DNA is being made. The principle of the SYBR Green I reporter chemistry is illustrated in Figure 2.2.

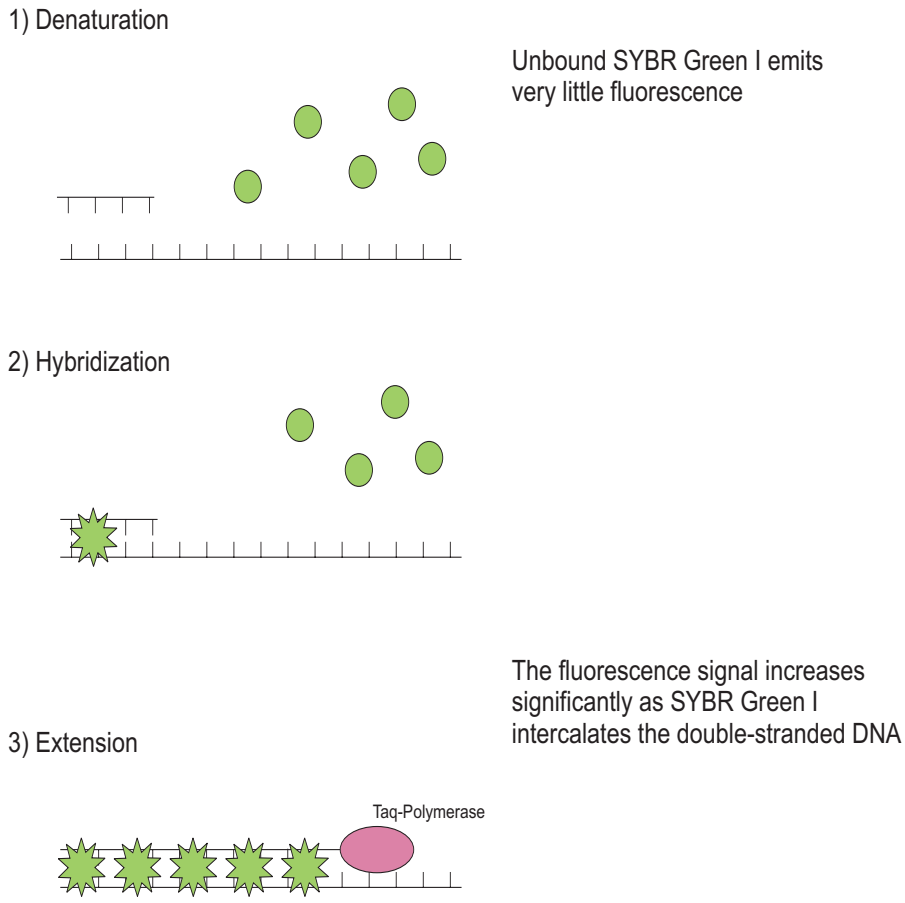


Figure 2.2. SYBR Green I chemistry. The dye only binds to double-stranded DNA and upon excitation emits fluorescence. The intensity of the fluorescence is proportional to the amount of DNA produced. Source: Cepheid brochure "SYBR Green I assays on the Smart Cycler system".

The Taqman probe technique increases both sensitivity and specificity of the reaction. QPCR using Taqman technology employs a target-specific oligonucleotide probe (Taqman probe), which contains a fluorescent dye at its 5´-end and a quenching dye at its 3´-end. The close proximity of the reporter and quencher prevents emission of fluorescence while the probe is uncleaved. Taqman probes are designed to hybridise to a specific DNA sequence which lies within the priming site of forward and reverse primers. During PCR elongation, the 5´-nuclease activity of the DNA polymerase hydrolyses the Taqman probe, separating the reporter and quencher so that fluorescence quenching no longer occurs. Fluorescence increases in each cycle proportional to the rate of probe cleavage. The principle of the Taqman QPCR is illustrated in Figure 2.3.

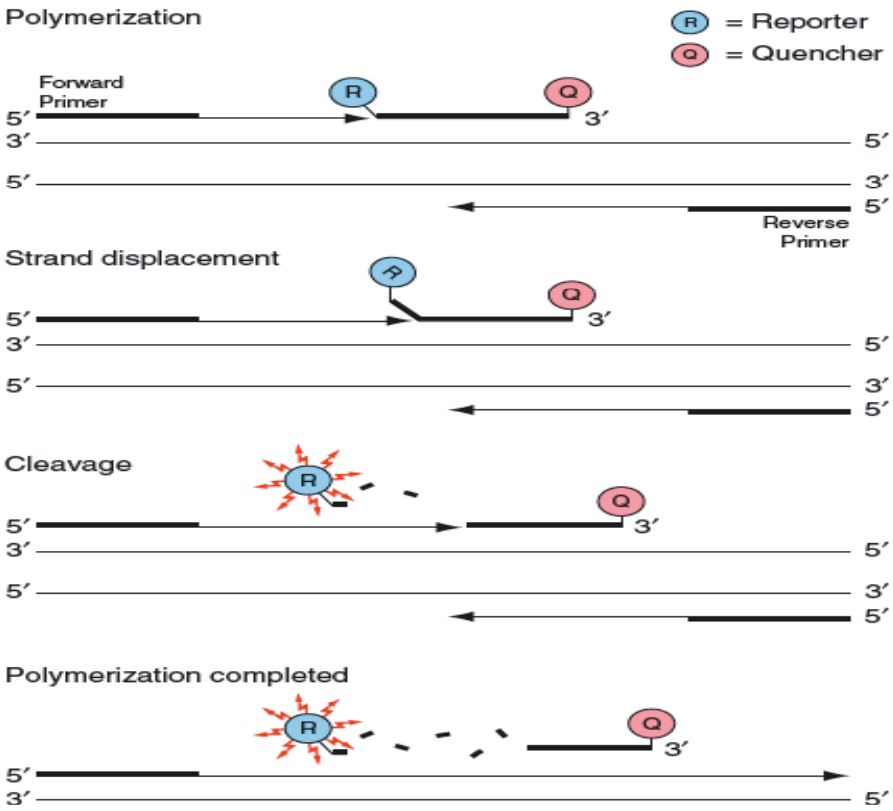


Figure 2.3. Principle of the Taqman quantitative PCR method. Source: Applied Biosystems, Inc.

The SYBR Green I method is convenient, cost effective, and can be applied directly to any PCR assay requiring only one primer pair without the need for a Taqman probe. However, SYBR Green binds non-specifically to all double-stranded DNA and does not distinguished between target and non-target DNA sequences such as primer-dimers and non-specific products. This binding of non-specific PCR products results in a fluorescence signal that is no longer proportional to the amount of amplicon, making the assay more qualitative than Taqman QPCR. The specificity of QPCR product formation can be confirmed using high resolution melting point analysis (HRM), which discriminates PCR products according to their melting temperature (T_m). The melting temperature of a PCR product is sequence-dependent and defined as the point at which 50% of the product is dissociated. The melting point analysis of a PCR amplicon usually results in a single specific melting point, which is distinctly different from the melting points of the primer-dimers and non-specific products. If multiple melting signals are detected the accuracy of the quantification results cannot be guaranteed. If the primers used in a SYBR Green I QPCR are well designed, this chemistry can work extremely well with non-specific background signals only showing up in very late amplification cycles. These non-specific amplicons can be distinguished from specific amplicons by melt-point analysis performed at the end of every QPCR run with SYBR Green I.

The Taqman assay provides the highest level of specificity and sensitivity of the QPCR methods. The added specificity provided by the probe itself renders the assay highly quantitative. In general, positive Taqman results do not require confirmation and validation by HRM due to the extreme specificity of the process.

2.2.2 Application of the QPCR method in the environment and the use of an internal reference standard.

QPCR has been used extensively in biomedical research for the detection and quantification of pathogens in patient tissues or medical supplies (Klein, 2002; Lovatt, 2002). We and others have adopted this methodology and demonstrated the potential for the quantitative analysis of harmful algal species in the environment (Bowers *et al*., 2000, Saito *et al*., 2002, Popels *et al.*, 2003; Coyne *et al.*, 2005, Handy *et al.*, 2005, Handy *et al.*, 2006, Demir *et al.*, in press). For this specific application, DNA is extracted from samples with known concentrations of algal cells and amplified using primers specific to the algal DNA. A standard curve is generated by making serial dilutions of the DNA extracted from the sample with known cell concentration. This is then used to determine cell concentrations in environmental samples. An example is given in Figures 2.4a and 2.4b for a toxic marine dinoflagellate *Pfiesteria piscicida*. Alternatively, the use of a single calibrator sample can be employed after validation (described in detail below). While this process is relatively straightforward, DNA extracted from environmental samples, in particular, can vary significantly in quantity and quality, affecting the outcome of QPCR by several fold. Co-precipitation of inhibitory compounds for PCR also confounds quantitative molecular analyses of environmental samples (Tebbe and Vahjen, 1993; Wilson, 1997) and frequently produces false negative results.

We recently developed a method in which we introduce a known concentration of exogenous plasmid DNA (pGEM) into the DNA extraction buffer as an internal standard (Coyne *et al.*, 2005). Since the target DNA is extracted in the presence of the internal standard, inherent variability in extraction efficiencies and the presence of inhibitors equally affect both the target and standard. The concentration of target DNA may then be measured by QPCR using target-specific primers and either SYBR green I or a Taqman probe and then corrected by QPCR determination of the internal standard copy number (refer to Section 2.4.2.4 for further information). The introduction of an exogenous internal standard during the extraction phase also reduces variability due to human error such as personal differences in laboratory technique, day-to-day proficiency, and pipetting or dilution errors. The introduction of the internal control makes this methodology truly quantitative and more amenable to the rigors of routine surveillance efforts.

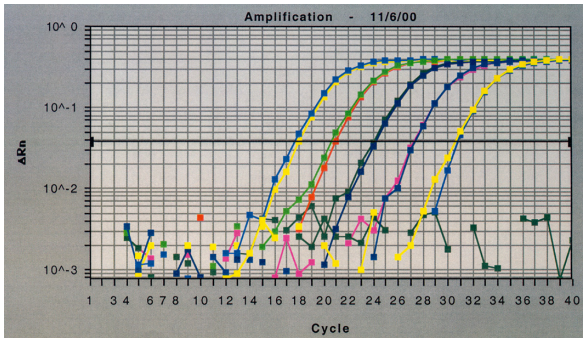


Figure 2.4a QPCR amplification plot of 10-fold dilutions of *Pfiesteria piscicida*.

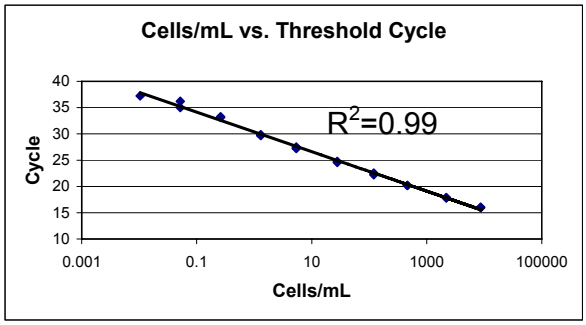


Figure 2.4b Log plot of C_T vs. *Pfiesteria piscicida* concentration in cells/mL of QPCR analysis of environmental samples spiked with known concentrations of *P. piscicida*.

Figure 2.4. QPCR amplification plot of 10-fold dilutions of *Pfiesteria piscicida* and the relationship of threshold cycle to cell concentration.

2.3 QPCR detection of Didymosphenia geminata using SYBR Green I

The objective of this part of the study was to determine the sensitivity range of the *D. geminata*-specific primers (D602F and D753R) developed in Phase 1 of the project (Cary *et al.*, 2006) when used with SYBR Green I. For this purpose, tenfold serial dilutions of plasmid DNA containing the 18S rDNA of *D. geminata* were subjected to QPCR and the sensitivity range determined by linear regression. The influence of exogenous DNA on the sensitivity of the SYBR Green I assay was also investigated.

2.3.1 Determination of the sensitivity range of the SYBR Green I based QPCR for *Didymosphenia geminata* using plasmid DNA

The QPCR was performed with a Rotor-Gene 6000 (Corbett, Life Sciences, Australia) in 0.1 ml PCR tubes using a 72-well rotor. The amplification reactions employed 12.5 µl containing 6 mM $MgCl_2$, 0.2 mM dNTPs, 1× Taq PCR reaction buffer, 0.75 units of AmpliTaq PCR-polymerase (Roche), 0.125× SYBR Green I (Invitrogen, Inc.) and 200 nM of each primer (D602F and D753R) (Invitrogen, New Zealand). The amount of plasmid template DNA ranged over ten orders of magnitude from 2.5 ng to 2.5 ag. An amplification negative control with no added template was included in each experiment. PCR cycling used the following conditions: 95°C for 2 min followed by 40 cycles at 95°C for 10 sec, 55°C for 15 sec and 68°C for 15 sec. The last temperature step included the fluorescence acquisition for SYBR Green I. After completion of PCR cycling, melting point analysis of all reactions was performed by gradually increasing the temperature by 1°C from 60°C to 99°C.

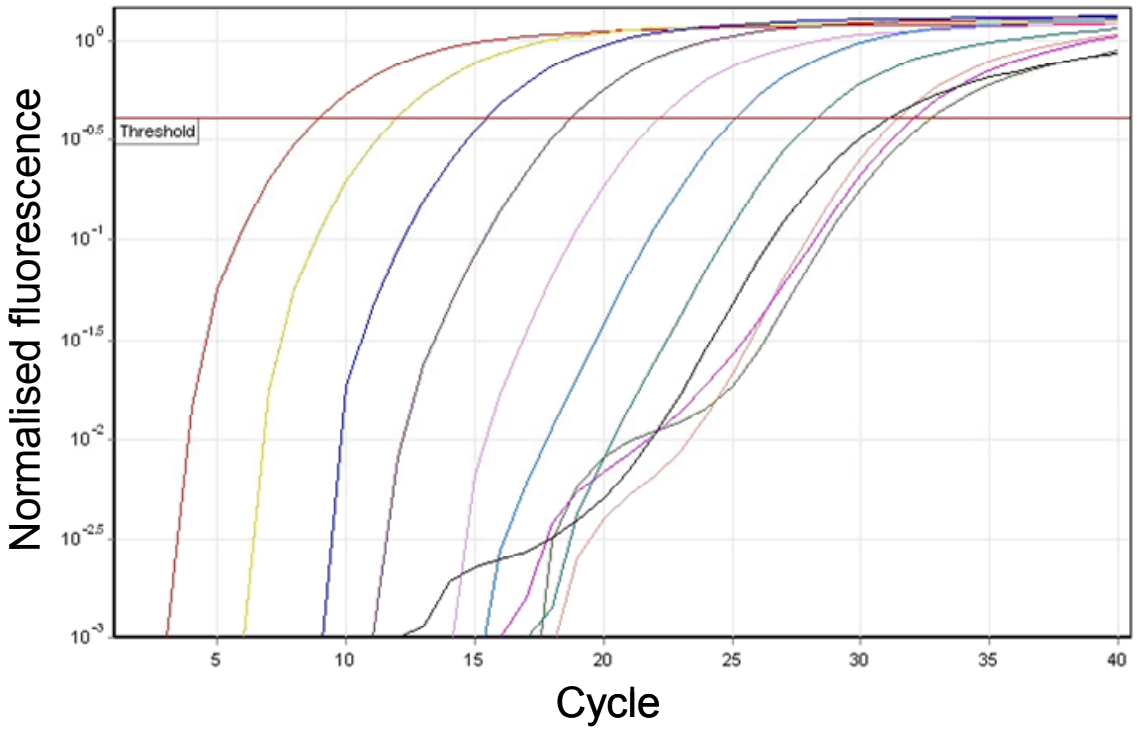


Figure 2.5. Sensitivity range of the SYBR Green QPCR for *Didymosphenia geminata* (quantitative PCR (QPCR) amplification plot). The QPCR used tenfold serial dilutions of *D. geminata* 18S rDNA plasmid DNA ranging from 2.5 ng to 2.5 ag.

The sensitivity of the QPCR was determined using known amounts of plasmid DNA with inserts containing the 18S rDNA of *D. geminata*. The QPCR had a linear ($R^2 = 0.999$) range of detection over eight orders of magnitude with a lower limit of detection of 250 ag ($2.5x10^{-16}$ g). This equates to approximately 40 target copies or equivalent to ~1.5 *D. geminata* cells. The PCR negative control (NTC) and samples containing template concentrations below 250 ag produced positive PCR signals ($C_T > 32.7$) which were not in the demonstrated linear range of detection. The upper range of detection was approximately 2.5 ng of plasmid DNA (Figures 2.5 and 2.6).

The accuracy of amplification products was verified by using melting point analysis subsequent to QPCR. The melting temperature (T_m) for each reaction was derived from the negative first derivative of the dissociation curve (data not shown) plotted over the temperature (-dF/dT versus T).

All reactions containing template produced a melting point signal for the 18S rDNA amplicon. The average melting temperature (T_m) was 84.0°C (Table 2.1; Figure 2.7). A second melting peak (peak 1; range: 76.5°C – 77.2°C) was detected for reactions containing no or less than 2.5 pg of template. This melting temperature was characteristic for primer-dimers.

Table 2.1. Melting point analysis *Didymosphenia geminata* plasmid DNA QPCR products.

No.	Name	Sample type	Template per reaction	Primer-dimer peak (°C)	*D. geminata* peak (°C)
1	Plasmid dilution 1	Standard	2.5 ng	N.D.	84.3
2	Plasmid dilution 2	Standard	250 pg	N.D.	84.3
3	Plasmid dilution 3	Standard	25 pg	N.D.	84.3
4	Plasmid dilution 4	Standard	2.5 pg	N.D.	84.2
5	Plasmid dilution 5	Standard	250 fg	76.5	84.0
6	Plasmid dilution 6	Standard	25 fg	77.2	84.2
7	Plasmid dilution 7	Standard	2.5 fg	77.0	84.0
8	Plasmid dilution 8	Standard	250 ag	76.7	83.8
9	Plasmid dilution 9	Standard	25 ag	77.0	83.5
10	Plasmid dilution 10	Standard	2.5 ag	77.0	83.5
11	PCR negative control	NTC	-	76.7	N.D.

N.D.: No detection

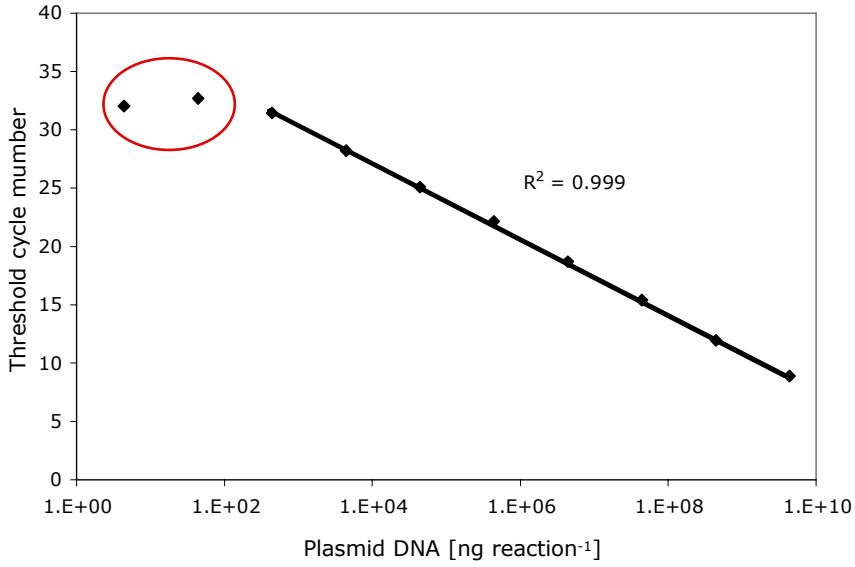


Figure 2.6. Sensitivity range of the SYBR Green I quantitative PCR (QPCR) for *Didymosphenia geminata* (linear regression line). The QPCR used tenfold serial dilutions of *D. geminata* 18S rDNA plasmid DNA ranging from 2.5 ng to 2.5 ag. The corresponding C_T values were plotted over log quantities of template attaining an $R^2 = 0.999$ at the higher target concentrations. The points in the circle shows the drop off in sensitivity at the lower target concentrations.

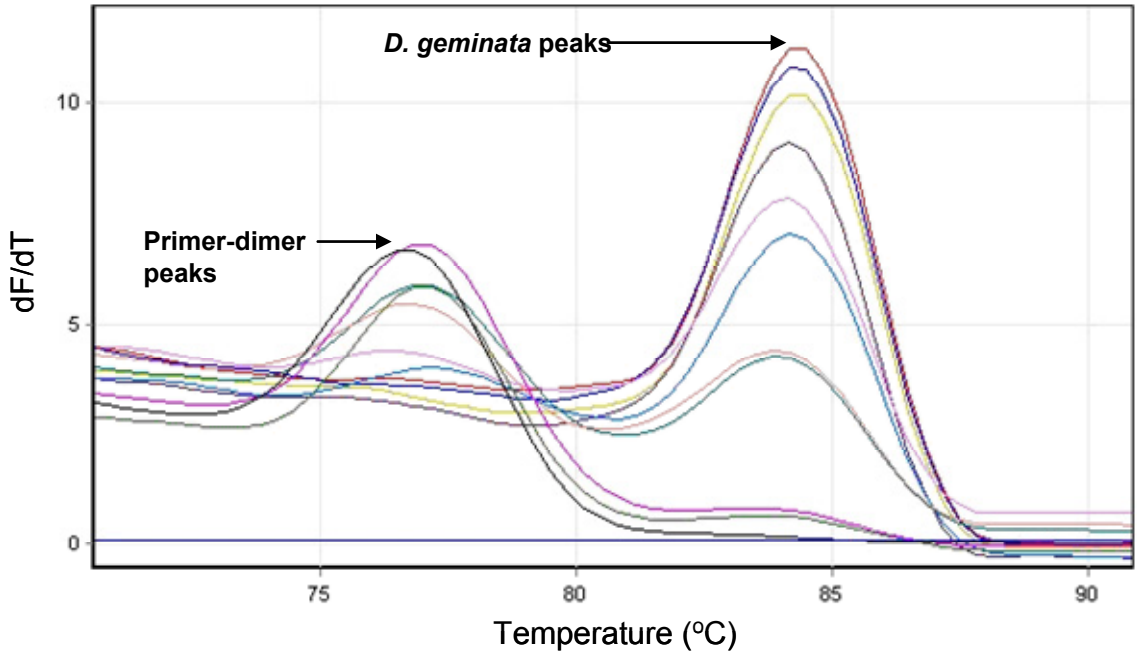


Figure 2.7. High-resolution melting curves of quantitative PCR (QPCR) reaction amplifying a small region of the 18S rDNA of *Didymosphenia geminata*. The QPCR used tenfold serial dilutions of *D. geminata* 18S rDNA plasmid DNA ranging from 2.5 ng to 2.5 ag.

2.3.2 Detection of *Didymosphenia geminata* in the presence of exogenous DNA

To investigate the influence of exogenous DNA (i.e., co-purified non-template DNA) on the sensitivity of the SYBR Green I QPCR, 15 ng of exogenous DNA (*Anoxybacillus flavithermus*) was added to each PCR. All reaction and cycling conditions were identical to the conditions described above.

The sensitivity of the QPCR was determined in the presence of exogenous DNA using known amounts of template plasmid DNA. The SYBR Green QPCR had a linear ($R^2 = 0.999$) range of detection over eight orders of magnitude with a lower limit of detection of 250 ag ($2.5x10^{-16}$ g), which equates to approximately 40 target copies or 1.5 *D. geminata* cells. The upper range of detection was approximately 2.5 ng of plasmid DNA (Figures 2.8 and 2.9).

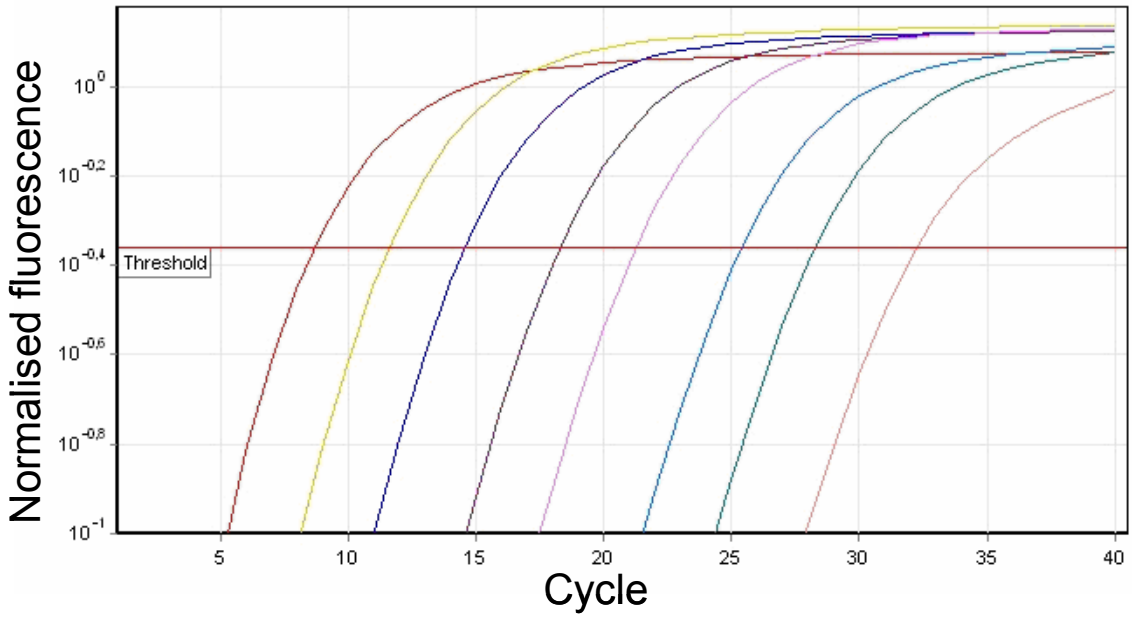


Figure 2.8. Sensitivity range of the SYBR Green I quantitative PCR (QPCR) for *Didymosphenia geminata* in the presence of exogenous DNA (QPCR amplification plot). The QPCR used tenfold serial dilutions of *D. geminata* 18S rDNA plasmid DNA ranging from 2.5 ng to 2.5 ag.

2.3.3 Field application of the SYBR Green I assay

To investigate the relative sensitivity of the SYBR Green I QPCR with environmental samples, 21 samples from Fiordland (1), Marlborough (3), Otago (1), Southland (12), Taranaki (1), West Coast (2), and USA (1), were amplified (Figure 2.10). All reaction and cycling conditions were identical to the conditions described above. The C_T values for these samples were relatively diverse, ranging from ~5 to ~35 cycles. However, the subsequent HRM (Figure 2.11) indicates that much of the fluorescence seen in samples with higher C_T values, and thus lower starting template copies, may be due to spurious primer-dimers. HRM curves for these samples with a relatively high C_T value, e.g., 30, demonstrate a significant secondary peak between 75 - 80°C. In particular, a sample from Stony River, Taranaki, North Island, New Zealand, produced an amplification curve in the SYBR Green I QPCR, which would be suspect because *D. geminata* has not been identified in this region. When viewed with HRM analysis, a clear primer-dimer peak was seen. Thus, a sample negative for *D. geminata* may still produce an amplification curve in SYBR Green I QPCR, namely a false positive.

By comparison, this secondary primer-dimer peak is virtually non-existent in samples which had a lower C_T value, e.g., ~15, and thus a higher starting template number. Overall, SYBR Green I QPCR is sufficiently sensitive to quantitatively detect *D. geminata* at relatively high starting concentrations, but at low concentrations, the interference of the primer-dimers present is confounding and could lead to false positives.

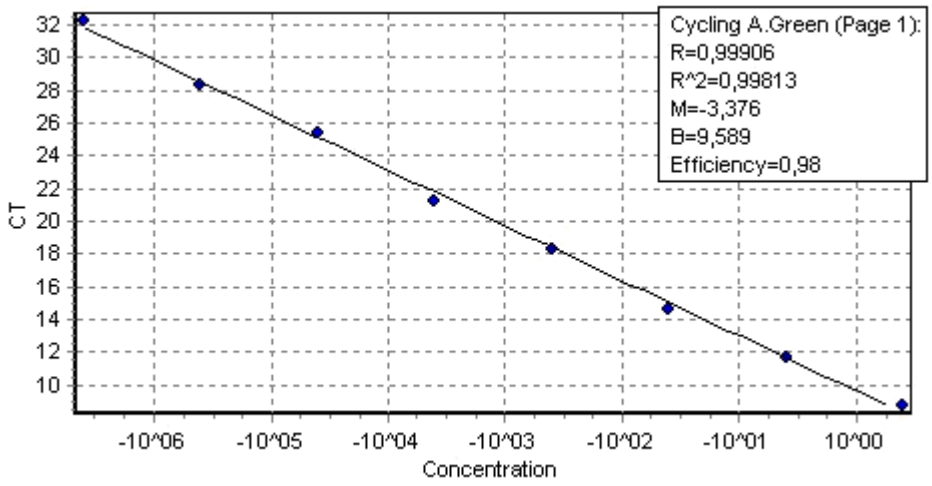


Figure 2.9. Sensitivity range of the SYBR Green I QPCR for *Didymosphenia geminata* in the presence of exogenous DNA (linear regression line). The quantitative PCR used tenfold serial dilutions of *D. geminata* 18S rDNA plasmid DNA ranging from 2.5 ng to 2.5 ag. The corresponding C_T values were plotted over log quantities of template.

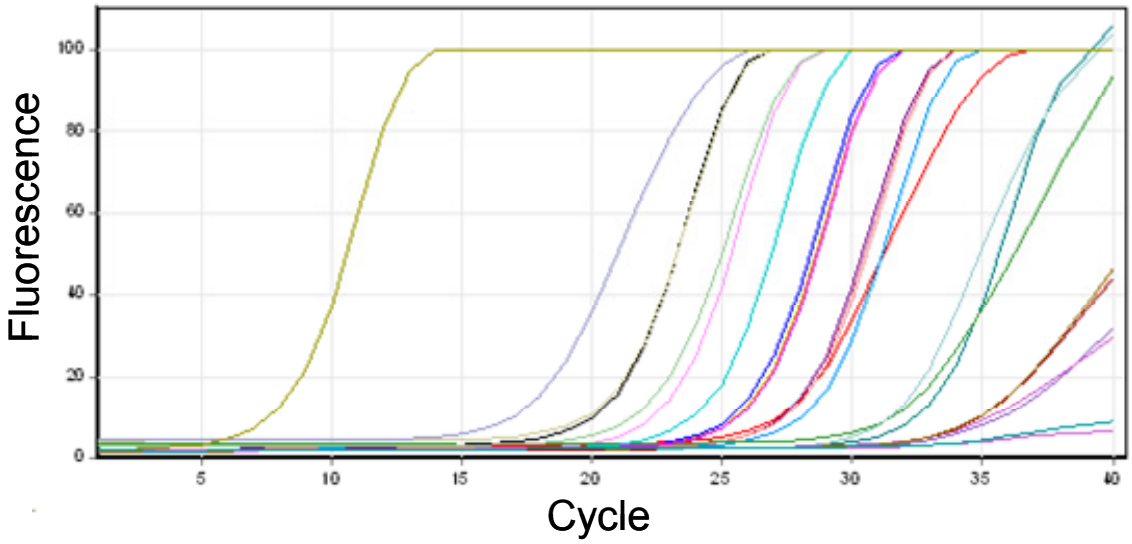


Figure 2.10. Quantitative PCR using SYBR Green I QPCR amplification plot of DNA extracted from samples collected from 21 rivers: Fiordland (1), Marlborough (3), Otago (1), Southland (12), Taranaki (1), West Coast (2), and USA (1), The assay also included a positive control (*Didymosphenia geminata* 18S rDNA plasmid DNA) and a no-template control (NTC).

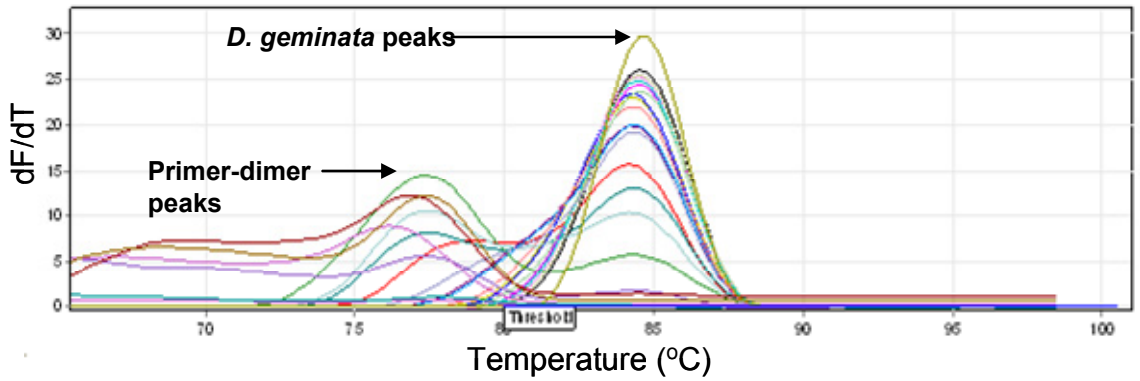


Figure 2.11. High-resolution melting curves for the 21 environmental samples amplified by SYBR quantitative PCR. Samples which had low C_T values correspond with single narrow peaks at ~84°C. Conversely, samples with high C_T values correspond with double peaks at ~77°C (primer-dimers) and 84°C.

2.3.4 Summary

The SYBR Green I QPCR for *D. geminata* proved to be very sensitive over a wide range of plasmid template concentrations. Moderate amounts of exogenous non-template DNA did not noticeably affect the sensitivity of the QPCR. The demonstrated lower limit of detection corresponds to approximately 40 target sequences or 1.5 *D. geminata* cells.

The limitation of the SYBR Green based QPCR was the generation of false positive amplification signals for negative control reactions (no added template) and for reactions containing template DNA below the equivalent of 1.5 *D. geminata* cells. Despite this constraint, SYBR Green I QPCR is a straightforward and cost-effective method most applicable in routine monitoring efforts where *D. geminata* is known to exist. However, SYBR Green I QPCR is not the preferred the method for surveillance efforts where low-level sensitivity is needed to detect presence/absence of founding populations.

2.4 QPCR detection of Didymosphenia geminata using a Taqman probe – a quantitative assay

2.4.1 Objectives

For this project, we used the *D. geminata*-specific primers (D602F and D753R) that were described in our previous report (Cary *et al.,* 2006) for the Taqman assay. Our strategy for developing and optimising the Taqman reaction was to:

(i) design a Taqman probe specific for *D. geminata*,
(ii) optimise the Taqman protocol for both *D. geminata* and the pGEM internal standard,
(iii) determine the efficiency and sensitivity of the Taqman reaction,
(iv) evaluate the sensitivity of the Taqman reaction in the presence of exogenous DNA,
(v) validate the use of a calibrator in the Taqman QPCR for relative enumeration,
(vi) evaluate the potential for multiplexing *D. geminata* and pGEM primers and probes in the same reaction,
(vii) examine two modifications to the extraction and amplification protocols,
(viii) evaluate reproducibility for multiple sample collections from the same site, and
(ix) test the QPCR reaction protocol on samples collected along a transect of Buller River (Nelson).

2.4.2 Protocols

2.4.2.1 Design of the Didymosphenia geminata-specific Taqman probe

The purpose of this portion of the study was to design a *D. geminata*-specific Taqman probe. The normal requirements for the probe design (length, T_m, base composition etc.) were adhered to in focusing the probe around a distinctive variable region between D602F and D753R on the 18S rDNA. The location for the hybridisation probe (D-641) was visually identified using multiple sequence alignment (ClustalW; European Bioinformatics Institute; http://www.ebi.ac.uk/clustalw/) of eleven 18S rDNA diatom sequences including *Amphora montana*, *Anomoeoneis sphaerophora*, *Cymbella* sp., *Didymosphenia geminata*, *Didymosphenia minuta*, *Dickieia ulvacea*, *Encyonema triangulatum*, *Eolimna minima*, *Eolimna subminuscula*, *Fragilaria striatula* and *Gomphonema parvulum* (Figure 2.12). The hybridization site of the Taqman probe started at nucleotide position 641 and consisted of 23 nucleotides, i.e. 5′-TCAGAAACTGTCATCCGTGGGTG-3′. The Taqman probe was synthesized with a 6-FAM reporter dye at the 5′-end and a Black Hole Quencher at the 3′-end (Integrated DNA Technology, USA).

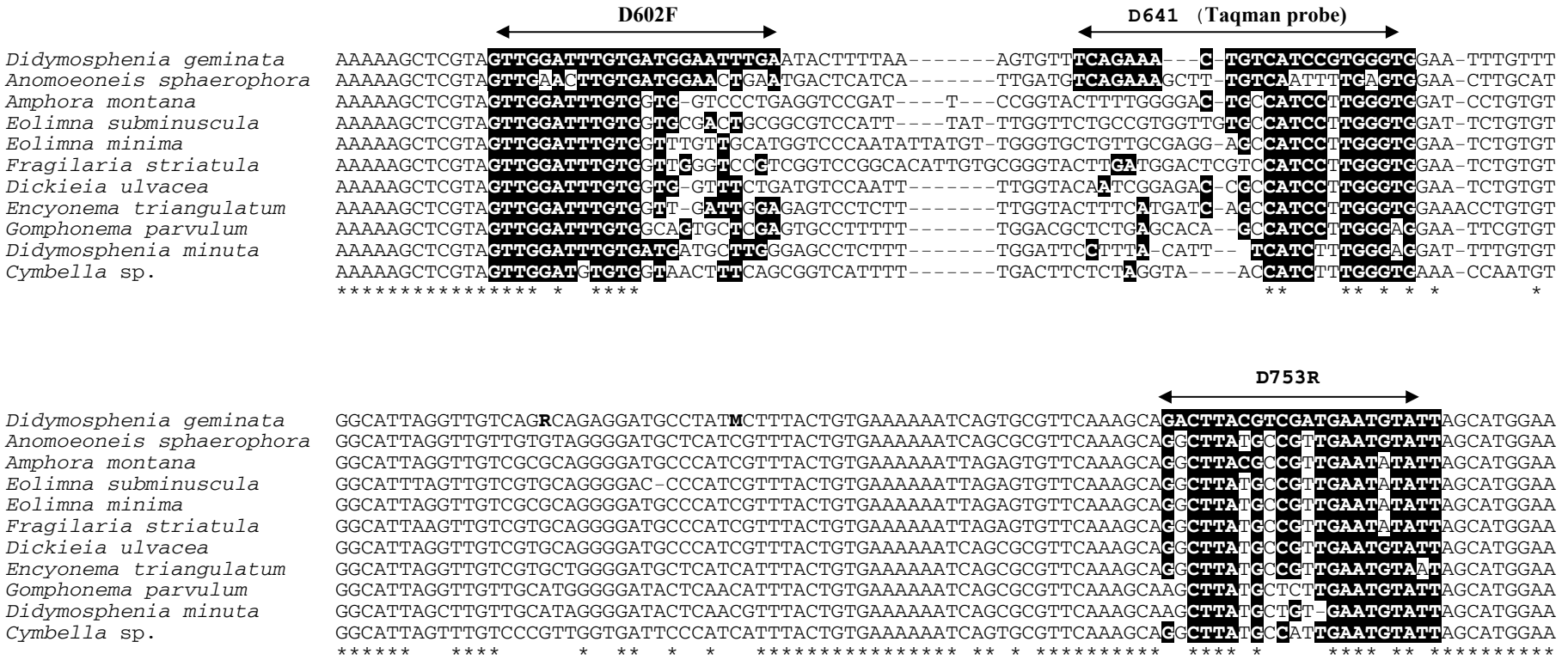


Figure 2.12. CLUSTAL W (1.82) multiple sequence alignment of partial 18S rDNA sequences for the selection of the Taqman probe and primers for *Didymosphenia geminata*. Consensus sequences within the primer and probe sites are highlighted in black.

2.4.2.2 Optimisation of Taqman reaction for D. geminata and pGEM templates

Primer and probe concentrations were optimized for QPCR on a Rotor-Gene 6000 (Corbett Life Science, AUS) using 100 ag of cloned plasmids containing the 18S rDNA sequence for *D. geminata* as template. Quadruplicate reactions were run with a matrix of 50 nM, 300 nM, and 900 nM primer concentrations. Primer concentrations for reactions with the lowest average C_T values were chosen as the optimum concentrations. For the *D. geminata*-specific primers, the optimum primer concentrations were 50 nM of D602F and 300 nM of D753R. For the pGEM primers, the optimum primer concentrations were 300 nM of M13F and 900 nM of pGEMR. The probe concentrations for each reaction were 200 nM.

The QPCR was performed in 0.1 mL PCR tubes using a 72-well rotor. The amplification reactions employed a volume of 12.5 µL containing 6 mM $MgCl_2$, 0.2 mM dNTPs (Roche), 1× PCR reaction buffer, 1 unit of Platinum *Taq* polymerase (Invitrogen, New Zealand), 50 nM of forward primer D602F (5′-GTTGGATTTGTGATGGAATTTGAA-3′), 300 nM of reverse primer D753R (5′-AATACATTCATCGACGTAAGTC-3′), 200 nM of probe (5′-FAM-CACCCACGGATGACAGTTTCTGA-BHQ-3′) (Integrated DNA Technology, USA), 0.8 µg of non-acetylated BSA (Sigma) and 1 µl of plasmid template (100 ag). The full protocol for setting up the PCR mastermix is given in Appendix 1. An amplification negative control with no added template was included in each experiment. PCR cycling used following conditions: 95°C for 2 min and 45 cycles of 95°C for 10 sec and 60°C for 45 sec. The last temperature step included the fluorescence acquisition.

The internal reference standard, pGEM was included in the CTAB extraction buffer at a concentration of 20 ng mL^{-1}. Primers and probe for amplification of pGEM are described in Coyne *et al.* (2005), but reaction conditions were re-optimized for use with the Rotor-Gene 6000 QPCR machine. Optimum conditions consisted of a 12.5 µL reaction containing 0.25 U *Taq* Polymerase, 1X reaction buffer, 6 mM $MgCl_2$, 0.2 mM dNTPs, 300 nM M13F primer, 900 nM pGEMR primer, 200 nM pGEM probe and 1.25 µL plasmid template (1 pg μL^{-1} final concentration). PCR cycling used is as per *D. geminata* protocol above. We also routinely will include 0.8µg of BSA in each reaction to help with possible environmental inhibitors.

2.4.2.3 Efficiency and sensitivity of the Taqman assay

Efficiency:

Once the optimum primer concentrations were determined, the efficiency of the reaction was evaluated by running a set of 10-fold dilutions of plasmid and plotting the C_T vs. log plasmid concentration. Efficiency is calculated from the slope of the curve:

$$E = 10^{(-1/slope)}$$

A reaction with an efficiency value of 2.0 is considered 100% efficient (Livak and Schmittgen, 2001).

The C_T vs. log concentration using the *D. geminata* primers yielded a linear relationship (R^2 = 0.987) with 10-fold dilutions starting with 4.032 X 10^{-2} ng plasmid down to 4.032 X 10^{-8} ng plasmid. The slope of the line is -3.3229, and the calculated efficiency of the reaction was 1.9996 (Figure 2.13).

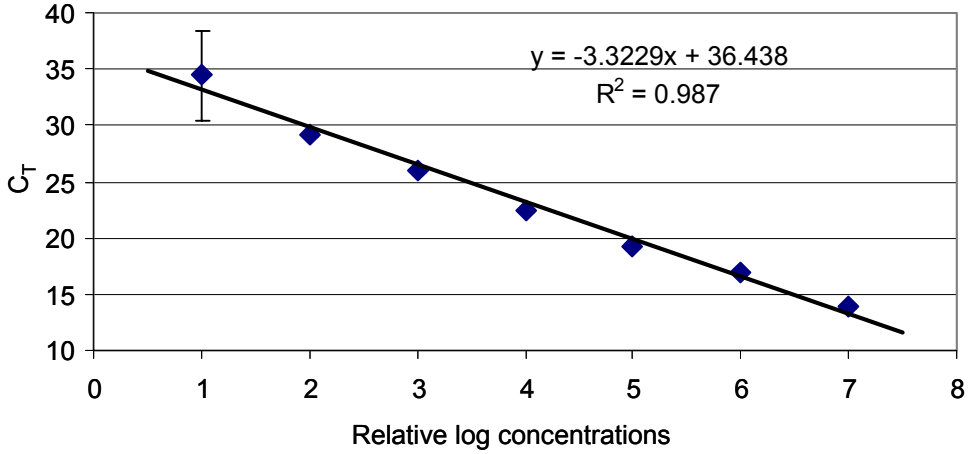


Figure 2.13. Graph of C_T vs. relative log concentrations of *Didymosphenia geminata* 18S rDNA plasmid dilutions.

Sensitivity:

The sensitivity of detection was calculated for copy number of template in the reaction from the mass of the plasmid + DNA insert for the *D. geminata* rDNA plasmid. The plasmid is 3,957 nucleotides in length and has a molecular weight of 1.22251515 X 10^6 g/mole. The *D. geminata* rDNA insert has 448 As, 445 Gs, 507 Ts, and 319 Cs. The molecular weight of an A is 313.12 g/mole, G is 329.21, T is 304.2, and C is 289.18. There were five ambiguous nucleotides in the sequence that have not been resolved. Each of these was given a molecular weight of 300 g/mole. The molecular weight of the double-stranded plasmid plus DNA insert was then calculated to be 3.5396758 X 10^6 g/mole.

To determine the sensitivity, eight 10-fold dilutions of a known mass of plasmid were prepared. These dilutions contained a range of plasmid from 0.016128 to 0.0000000016128 ng/µL. An aliquot of 2.5 µL of each diluted plasmid was used in duplicate QPCR reactions and the C_T values were determined (Table 2.2).

Table 2.2. C_T values for duplicate reactions of 10-fold dilutions of plasmid DNA containing the 18S rDNA for *Didymosphenia geminata*. N.D., no detection.

Dilution	DNA (ng)	C_T			
		Replicate 1	Replicate 2	Average	Standard deviation
1	0.04032	14.02	13.84	13.93	0.127
2	4.03E-03	16.85	16.86	16.86	0.007
3	4.03E-04	19.16	19.32	19.24	0.113
4	4.03E-05	22.46	22.23	22.35	0.162
5	4.03E-06	26.19	25.80	26.00	0.275
6	4.03E-07	29.36	29.03	29.20	0.233
7	4.03E-08	37.30	31.63	34.47	4.009
8	4.03E-09	N.D.	N.D.		

Using the molecular weight of the plasmid and Avogadro's number (6.022 X 10^{23} per mole), the lowest dilution that could be detected (number 7) was calculated to contain 68 copies of the plasmid with insert equivalent to ~ 2 *D. geminata* cells.

2.4.2.4 Validation of Taqman QPCR for relative enumeration using the delta delta C_T ($\Delta\Delta C_T$) method

There are two options for quantitation: (1) a standard curve can be run with every set of samples for both the target (*D. geminata*) template and the internal standard (pGEM); or (2) a calibrator sample can be run to determine cell concentrations of unknown samples relative to the calibrator using the $\Delta\Delta C_T$ method. For the $\Delta\Delta C_T$ method to be valid, the *D. geminata* and pGEM reaction efficiencies must be similar so that the difference in C_Ts (ΔC_T) for the ten-fold dilutions of the target (*D. geminata*) and the internal standard (pGEM) remain constant (Livak and Schmittgen, 2001) .

To validate the $\Delta\Delta C_T$ method for the *D. geminata* Taqman assay, template DNA was prepared using a dilution series of a mix of 0.01 ng/µL of the *D. geminata* and pGEM plasmid. The difference in C_T values for each dilution was calculated using the following equation:

$$C_T(\text{target}) - C_T(\text{pGEM}) = \Delta C_T$$

ΔC_T was then plotted vs. the relative log concentrations of the plasmids. For the $\Delta\Delta$Ct method to be valid, the slope of the line must be ± 0.1. The results yielded a slope of 0.0497, validating the use of the $\Delta\Delta$Ct method (Figure 2.14) for the *D. geminata* Taqman assay.

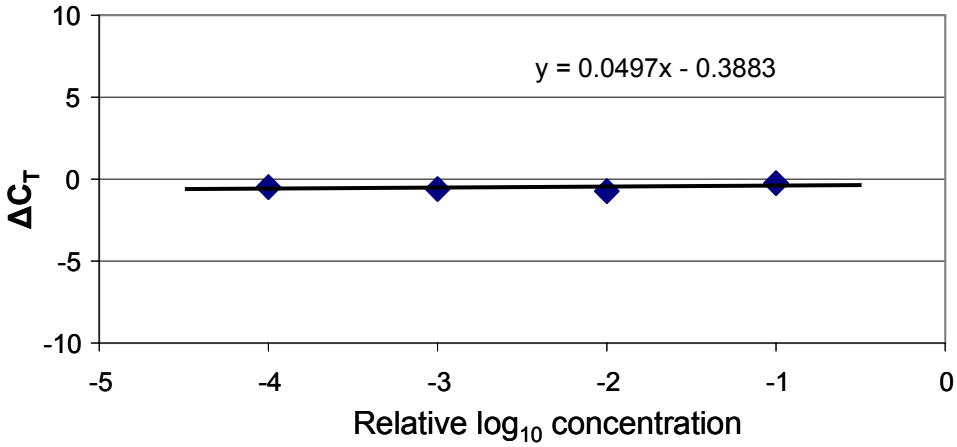


Figure 2.14. Graph of ΔC_T against log concentration of *Didymosphenia geminata* and pGEM plasmid dilutions.

2.4.2.5 Multiplexing the D. geminata and pGEM reactions

We evaluated the potential for multiplexing (combining two assays into one reaction) the *D. geminata* and pGEM Taqman assays using plasmids. The pGEM concentration is essentially constant for all samples, but *D. geminata* DNA concentration varies, thus a dilution series of *D. geminata* 18S rDNA plasmid by itself or with a constant concentration of pGEM was prepared. Standard conditions describe above were used. The C_T values for the multiplex reactions were compared to C_T values for simplex (single primer pair) reactions.

The C_T values for *D. geminata* reactions with and without pGEM primers were identical for the most concentrated samples (Figure 2.15). However, when the ΔC_T (*D. geminata* – pGEM) was greater than ~5 C_T units (in other words, *D. geminata* template was much less than pGEM template), the C_T values (and standard deviation) for the *D. geminata*

reactions with pGEM primers (i.e., multiplex) were significantly higher than the C_T values for *D. geminata* without pGEM primers (i.e., simplex) (Figure 2.15). This significant difference led us to proceed with the validation of only the simplex reaction.

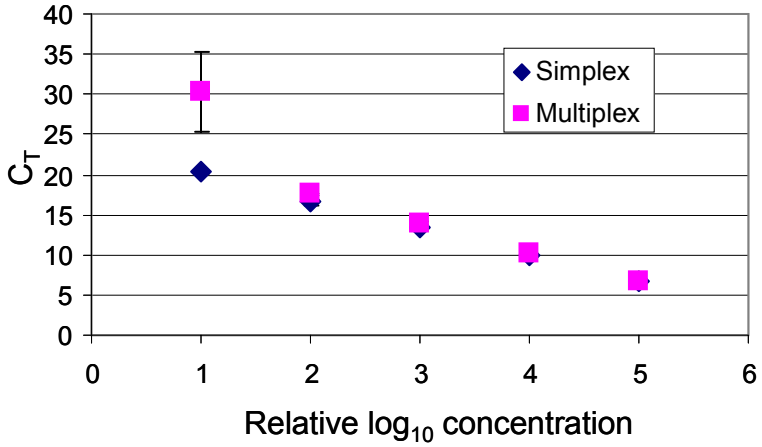


Figure 2.15. Graph of C_T values against relative log concentrations of *Didymosphenia geminata* plasmid used in multiplex vs. simplex QPCR reactions.

2.4.2.6 Determination of the sensitivity range of the Taqman assay for D. geminata in the presence of exogenous DNA

To investigate the influence of exogenous DNA (co-purified non-template DNA from other organisms) on the sensitivity of the Taqman QPCR, exogenous DNA (*Anoxybacillus flavithermus*) was added to tenfold serial dilutions prepared from the calibrator sample 40B. For this sample a cell concentration of 30,859 ± 3,349 cells/mL was determined by microscopic cell count (Section 2.4.3.1). A total of 2 mL of this sample was subjected to CTAB DNA extraction (Appendix 2, full CTAB protocol) resulting in DNA being extracted from 61,718 cells. The total amount of DNA recovery from this sample was 1040 ng resuspended in 10 µL of low TE. For QPCR analysis the calibrator sample was routinely diluted to 10 ng/µL, which correlated to approximately 1,200 *D. geminata* cells. A ten-fold serial dilution of the 10 ng/µL sample was carried out in a solution containing 15 ng/µL of *A. flavithermus* DNA. An aliquot of 1 µL from each dilution step was then used as template for QPCR analysis. Taqman QPCR conditions were identical to the conditions described in Section 2.4.2.3.

The QPCR had a linear ($R^2 = 0.997$) range of detection over five orders of magnitude with a lower limit of detection of 1 pg correlating to an equivalent of approximately 1-0.1 *D. geminata* cells (Figure 2.16). The PCR negative control containing no added template did not cross the threshold line within 45 amplification cycles (data not shown). This study shows that when compared to the same reaction without exogenous DNA, we have lost some sensitivity.

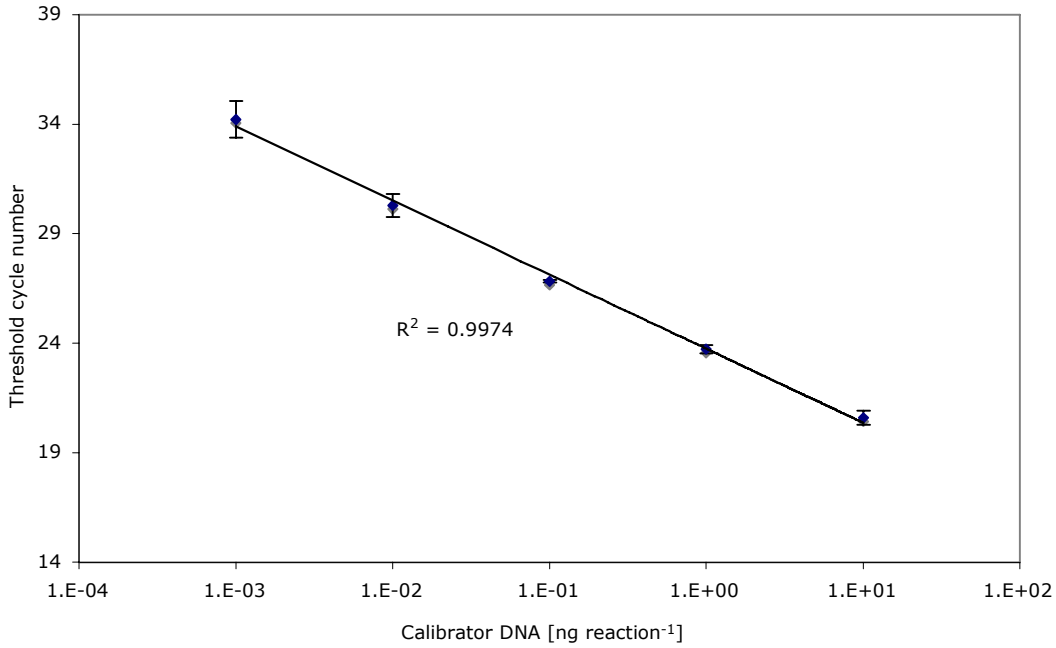


Figure 2.16. Sensitivity of QPCR for *Didymosphenia geminata* in the presence of exogenous DNA. Dilutions were prepared from the calibrator sample 3 (40B) (Table 2.4) and the mean values of C_T plotted over the log quantities of template. Error bars represent standard deviations for three replicates.

2.4.2.7 Modifications to the extraction protocol

Additional chloroform treatments and Proteinase K (PK) to improve DNA yield.

We wanted to determine if there would be any difference in quality or quantity of DNA extracted with only one chloroform (chl) extraction vs. two chloroform extraction steps. Environmental samples were collected from 8 sites over 2 days in the Buller, Gowan, and Owen Rivers (Table 2.3, Figure 2.17). Samples were collected by suspending a 40-µm mesh drift net that was 240 mm in diameter in flowing water for 2 min.

Table 2.3. Locations of samples collected and *Didymosphenia geminata* cell concentrations at each site on 30-31 October 2006 as determined by QPCR.

Site	Latitude and longitude	Elevation (m)	Location	Water velocity (m/s)	Volume filtered (m^3)	Cells in total volume	Cells/L
1	41º 47.119' S 172º 48.600' E	581	Buller R at St Arnaud's bridge	0.55	2.99	187	0.06
2(A)	41º 43.226' S 172º 41.131' E	486	Buller R at Howard R	0.69	3.75	77,143	21
2(B)	41º 43.226' S 172º 41.131' E	486	Buller R at Howard R	0.69	3.75	63,863	17
2(C)	41º 43.226' S 172º 41.131' E	486	Buller R at Howard R	0.69	3.75	80,276	21
3	41º 42.185' S 172º 39.458' E	376	Buller R near Howard Junction	1.25	6.79	1,203,501	177
4	41º 44.744' S 172º 35.313' E	376	Gowan R upstream of holiday camp	0.76	4.13	83,661	20
5	41º 43.165' S 172º 34.341' E	323	Gowan R near Buller R confluence	1.45	7.87	324,500	41
6	41º 40.059' S 172º 27.994' E	221	Owen R upstream site	0.15	0.81	0	0
7	41º 41.253' S 172º 26.999' E	198	Buller R at Owen R Recreation Reserve	0.21	1.14	467,098	410
8	41º 47.810' S 172º 19.509' E	113	Buller R at Murchison	0.16	0.87	132,748	153

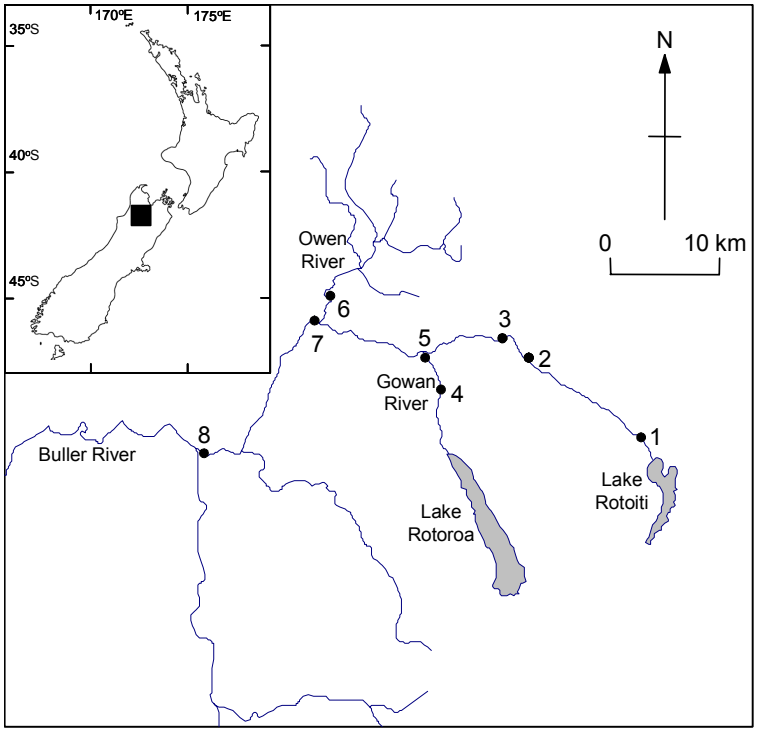


Figure 2.17. Location of sampling sites in the Buller River catchment.

Four samples were collected from site 2 on 30 October to determine reproducibility. DNA was extracted from 1-2 mL of net samples using the CTAB protocol (Cary *et al.*, 2006). Samples were resuspended in LoTE buffer and the concentrations and 260:280 ratios were determined using the NanoDrop. Samples with low 260:280 ratios (<1.8) were extracted a second time, with an second chl:isoamyl alcohol (IAA) extraction step added.

Our results demonstrate that a second chl:IAA extraction decreased the yield of DNA but did not significantly alter the 260:280 ratio. The yield varied from 67 ng/µL to over 1.8 µg/µL (Table 2.4).

In addition, we set out to determine whether the addition of a PK treatment to the DNA extraction protocol significantly improved the yield and intergrity of the extracted DNA. Two 1-mL aliquots of a net-sample from the Wild Natives River, Fiordland, were extracted using the original CTAB protocol (Cary *et al.*, 2006). An additional two 1 mL aliquots were subjected to PK treatment before the CTAB extraction by the addition of 30 µL of PK (10 mg/mL) and incubated at 55°C for 1 hour. Upon completion of both extractions, NanoDrop analysis was performed. The results demonstrated a 2-3 fold increase in the total nucleic acid yield in samples subjected to PK, and a significant improvement in the 260:280 ratios. Following this observation, we integrated the PK treatment for environmental samples into our CTAB protocol (see Appendix 2).

Table 2.4. Evaluation of chloroform extraction. Each sample with absorbance ratios at 260:280 nm wavelengths of less than 1.8 (260:280 (1)) was subjected to a second DNA extraction including an additional chloroform extraction (260:280 (2)).

Site	Sample	[DNA] (1) (ng/µL)	260:280 (1)	[DNA] (2) (ng/µL)	260:280 (2)
1	1	67.65	1.91		
2	2-1 (40-1)	86.00	1.64	63.18	1.76
2	2-2 (40-2)	96.73	1.62	68.49	1.73
2	2-3 (40-3)	111.12	1.21	105.76	1.86
2	2-4 (40-4)	80.99	1.69	32.3	1.71
3	3 (40B)	336.43	1.98		
4	4	1865.96	1.99		
5	5	1367.06	1.97		
6	6	315.68	1.58	255.86	1.63
7	7	435.35	1.91		
8	8	391.13	1.82		

Addition of BSA to improve Taqman reactions

Bovine serum albumin (BSA) is known to increase the efficiency and sensitivity of PCR reactions by effectively removing inhibitory contaminants from the reaction. We wanted to test the effect of BSA on QPCR reactions for serial dilutions (10 to 0.001 ng/µL) from an environmental sample. Our results indicated that some inhibition was present at higher DNA concentrations and that the BSA effectively removed this. There was little apparent inhibition for the low concentration samples, however the addition of BSA made no difference in C_T values for these samples (Table 2.5).

Table 2.5. Evaluation of bovine serum albumin (BSA) addition to quantitative PCR (QPCR) reactions of 10-fold dilutions of DNA extracted from an environmental sample. (–) indicates samples without BSA. (+) indicates samples where BSA was included in the QPCR reaction. N.D., not detected.

DNA concentration (ng/µL)	C_T				
	Replicate 1	Replicate 2	Replicate 3	Average	Std. Dev.
10 (–)	N.D.	N.D.	N.D.		
10 (+)	40.67	N.D.	41.79	41.23	0.791
1 (–)	23.81	24.38	23.94	24.04	0.298
1 (+)	24.39	24.13	24.03	24.18	0.185
0.1 (–)	27.58	27.47	27.85	27.63	0.195
0.1 (+)	27.58	27.59	27.66	27.61	0.043
0.01 (–)	32.55	31.09	32.35	31.99	0.791
0.01 (+)	31.26	32.00	33.04	32.10	0.894
0.001 (–)	N.D.	N.D.	36.58	36.58	
0.001 (+)	N.D.	N.D.	N.D.		

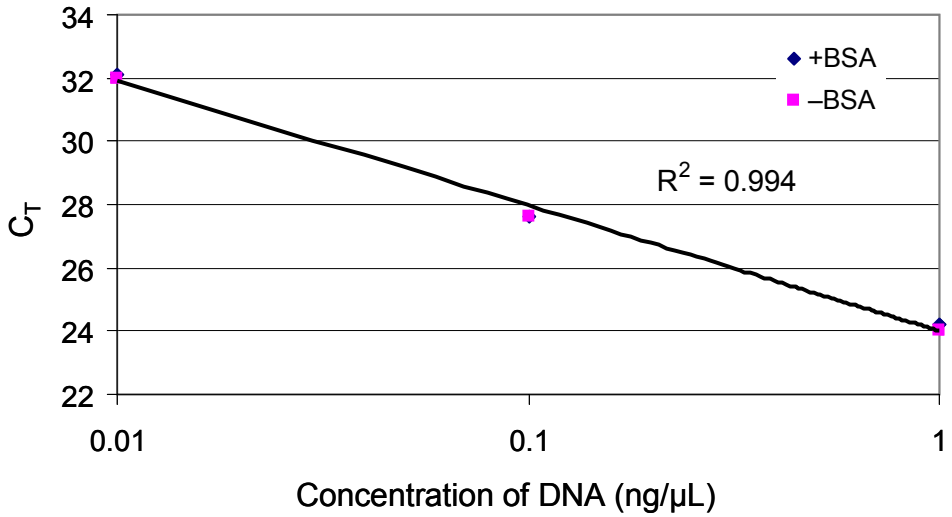


Figure 2.18. Graph of C_T value vs. concentration of DNA extracted from environmental sample with and without added bovine serum albumin (BSA).

Two samples with low 260:280 ratios were subsequently evaluated by QPCR. Samples 8(40-1) and 8(40-4) (Tables 2.4 and 2.6) were diluted to 1 and 0.1 ng/µL. Both of the diluted samples were amplified. Inhibition can be evaluated by the ΔC_T of 10-fold dilutions. As described above, the efficiency of the reaction is $E = 10^{(-1/slope)}$ and a value of 2.0 is considered 100% efficient. Using the C_T values of the 10-fold dilutions, we calculated the slope and efficiency of each of the reactions for these two samples:

ΔC_T sample 8(40-1) slope = 3.5 Efficiency = 1.93
ΔC_T sample 8(40-4) slope = 3.13 Efficiency = 2.08

These results demonstrate that even though the 260:280 ratios were low, the QPCR reaction was not affected (Figure 2.18).

2.4.2.8 Assay reproducibility for DNA extracted from environmental samples

Four samples were collected from one site in the Buller River (Section 2.4.3.1) to determine reproducibility in net sampling. We wanted to assess if field collections needed to be undertaken in duplicate or if a single net sample was representative of the population. Aliquots of 2 mL from each sample were extracted using the CTAB method (Appendix 2) and Taqman QPCR was run using *D. geminata* and pGEM primers/probes as described in Section 2.4.2.2. Average concentrations (relative to the calibrator sample 40-B) and standard deviation for the four samples were found to be 0.850 ± 0.231 (Table 2.6).

Table 2.6. Evaluation of replicate samples collected at site 2 on the Buller River, (Nelson), 30 October 2006. Concentrations of *Didymosphenia geminata* in each sample 40-2 through 40-4 are relative to concentrations to sample 40-1.

Sample number	ΔCt (*D. geminata* – pGEM)	*D. geminata* concentration relative to sample 40-1	Sample volume (mL)	Relative *D. geminata* per sample
40-1	4.965	1	45	1
40-2	5.380	0.750	32	0.5333
40-3	4.875	1.064	35	0.8278
40-4	4.545	1.337	35	1.0406

Replicate samples from this sample collection were found to vary by approximately 25% in cell concentrations of *D. geminata* (Table 2.3). However, if the "outlier" (Sample 40-2) is removed from the analysis, results of the remaining triplicate samples are more consistent, with an average relative concentration of 0.956 ± 0.113, or about 11% variability expected. For presence/absence determinations, this variation in cell concentration between samples will not affect the results, but if the cell enumeration capability of the QPCR method is required, for instance, in measuring outcomes of a control treatment, then at least two and preferably three drift net samples should be taken per site.

2.4.3 Application of the DNA method to environmental samples

2.4.3.1 QPCR results for Buller, Owen, Gowan river sampling using the Taqman assay

Didymosphenia geminata cells with chloroplasts intact were counted from sample 7 (40B) for use as calibrator. Cell enumeration was carried out on this sample in triplicate using an inverted microscope (Olympus CKX41) and an Utermöhl settling chamber (Utermöhl 1958). For this sample, there were 30,859 ± 3,350 cells/mL of 70% ethanol. We extracted 2 mL of the 70% ethanol sample, so that the total number of cells extracted was 61,718 cells. This sample was then used as a calibrator, with all reactions normalized to the pGEM signal. The number of *D. geminata* cells per ethanol fixed sample and per m^3 of water sampled was calculated (Table 2.3) using the equation described below. Note that for each sample in the QPCR reaction, it is assumed that 2 mL of the ethanol-fixed sample is extracted and that the volume of CTAB for each unknown and calibrator is the same. The water velocity at each site was recorded with a flow meter (Flo-mate 2000, Marsh-McBirney, USA). The cells/m^3 of water collected was calculated from the velocity of the river and the diameter of the net used for collection as follows:

$$\text{Cells/m}^3 = [\Delta\Delta C_T \text{ (unknown)}] * (\text{cells in 2 mL of ethanol-fixed calibrator}) * (\text{total volume of ethanol}/2) / [\text{water velocity (m/s)} * \text{area of collection net (m}^2) * \text{time for collection (s)}]$$

These results demonstrate the low levels possible of the Taqman QPCR method for detecting *D. geminata* in environmental samples. The method was able detect *D. geminata* at very low concentrations in the sample from Site 1 (Buller River at St Arnaud's bridge) where no *D. geminata* mass on the river bed was visible. With the Taqman QPCR method now operational, we began the process of specificity validation using samples collected from New Zealand and internationally.

2.5 Defining the lower end of detection

With any sensitive detection method, a lower end of sensitivity must be defined for routine surveillance purposes. The challenge is to assign a level of confidence around an identifiable point of detection. This becomes the cut off for what would be defined as a positive as opposed to a questionable result. After running hundreds of repeat and replicate Taqman assays using dilutions of our calibrator control sample 40B (see section 2.4.3.1), it is clear that this cut off should be set at 1 pg of template DNA, which is equivalent to ~0.1 cell in the reaction tube (Figure 2.19). We have established this as the "Sensitivity Threshold" of detection for the *D. geminata* Taqman assay. Any sample presenting a C_T above the Threshold will be deemed positive warranting further validation. Any sample producing a C_T that falls below the threshold (BT) will be considered questionable and require a second QPCR, re-extraction, or possibly re-sampling of the river including the use of benthic swabs (Cary *et al.,* 2006).

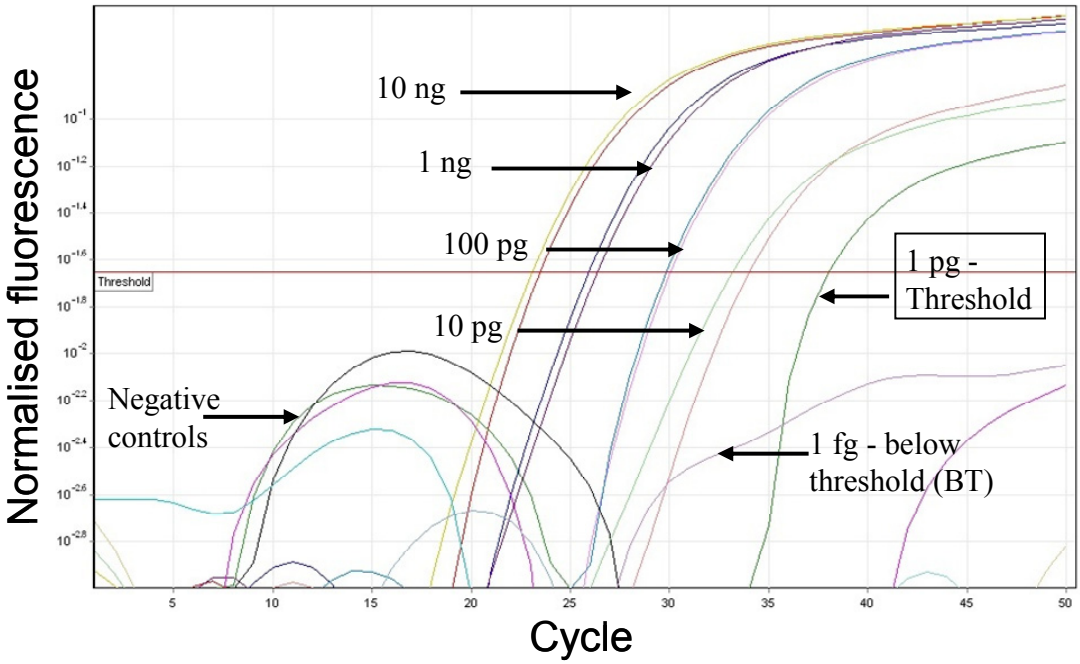


Figure 2.19. Taqman profiles of calibrator sample (40B) at 10 X dilutions identifying the 1 pg threshold of detection.

2.6 Conclusion and recommendations

The SYBR Green I QPCR for *D. geminata* proved to be extremely sensitive over a wide range of plasmid template concentrations. Moderate amounts of exogenous non-template DNA did not noticeably affect the sensitivity of the QPCR. The demonstrated lower limit of detection was 250 ag of plasmid DNA, which corresponds to approximately 1.5 *D. geminata* cells (assuming that one genome contains 30 ribosomal operons per cell).

The major limitation of the SYBR Green I assay was the generation of false positive amplification signals for negative control reactions (no added template) and for reactions containing low input copy numbers of template. This was due to the detection of non-specific amplification products. This hampers analysis of the presence or absence of *D. geminata* in environmental samples with very low cell abundance. Although melting point analysis can

assist in verifying the correctness of the amplicon, the results obtained in this study were sometimes ambiguous.

The robustness and specificity of the SYBR Green I QPCR assay could be improved using a dual-labelled fluorescent probe specific to *D. geminata*. In Taqman QPCR, the release of fluorophore is directly linked to the synthesis of the target sequence while primer-dimers and non-specific products remain undetected. The primer and Taqman probe combination designed for *D. geminata* proved to be both sensitive and robust, with detection of approximately 68 copies of the 18S rDNA from *D. geminata*. We are currently evaluating probes with longer length than the one described here. Longer probes have increased T_m and should yield increased sensitivity and accuracy of detection. We also validated the $\Delta\Delta C_T$ method for use with the *D. geminata* primers and probe. The analysis confirmed that ΔC_T for dilutions of the target template DNA is consistent over a wide range of concentrations, indicating that a single calibrator sample of known cell concentrations may be used to accurately determine the cell concentrations of unknown samples.

The method was tested with "real" samples collected from the Buller, Owen and Gowan Rivers. Several of the sites had visible *D. geminata* while three sites had no visible *D. geminata*. Two of the three sites with no visible *D. geminata* were positive by QPCR and all positive samples were confirmed to contain *D. geminata* DNA by denaturing gradient gel electrophoresis analysis (Cary *et al.,* 2006). Although the Taqman method appears to be extremely sensitive and specific to *D. geminata*, the specificity of the method to detect only *D. geminata* in New Zealand waters will continue to require validation. From these analyses we have set the lower end of detection (Sensitivity Threshold) to be 1 pg of our calibrator sample which is equivalent to ~1.0 *D. geminata* cells in the QPCR Taqman reaction.

3 Validating the DNA Method

3.1 Introduction

The genetic techniques described in Section 2 have been developed to be highly sensitive (single-cell level) in detecting a diagnostic gene of *D. geminata* from environmental samples. Because of the specificity and sensitivity of the DNA method, and the obvious political and environmental issues that it addresses, it was imperative that the method be stringently validated before it is used routinely for surveillance. Validation is necessary because the small pieces of DNA targeted by the method must be proven to be unique to *D. geminata* and not found in any other closely related organism. As described in Cary *et al.*, (2006), the design of the method hinges on locating a *D. geminata*-specific variable region of the target gene (18S ribosomal RNA). Uniqueness can only be determined by sequence comparison to other closely related organisms that currently exist in our database. Since we are working with environmental samples collected from a broad range of locations, there is the possibility, albeit remote, that an uncharacterised "look alike" diatom might exist that would produce a false positive signal. In order to validate the methodologies, we have been running exhaustive multi–level genetic tests on currently infected rivers to assure the specificity of the method. The benchmark that we set was that at least 100 *D. geminata* positive samples from different sites around New Zealand and the world will be further validated by at least two of three additional genetic techniques (gel electrophoresis, high resolution melt curve analysis, and direct sequencing of the amplicon). Each of these methods provides insight into the identification of the gene being detected in each sample at different levels of resolution.

3.2 Objectives

3.2.1 Gel electrophoresis

The first level of validation relies on the fact that the *D. geminata* QPCR method will produce an amplification product of a very specific size (151 base pairs). The assumption is that any positive amplification yielding a product not of this size would be a false positive. Figure 3.1 shows a gel image with typical *D. geminata* QPCR (D602F and D753R) amplification products from duplicate reactions of seven Buller River samples referenced to a DNA ladder (L). Each of the reaction from these samples yielded the correct size amplicon.

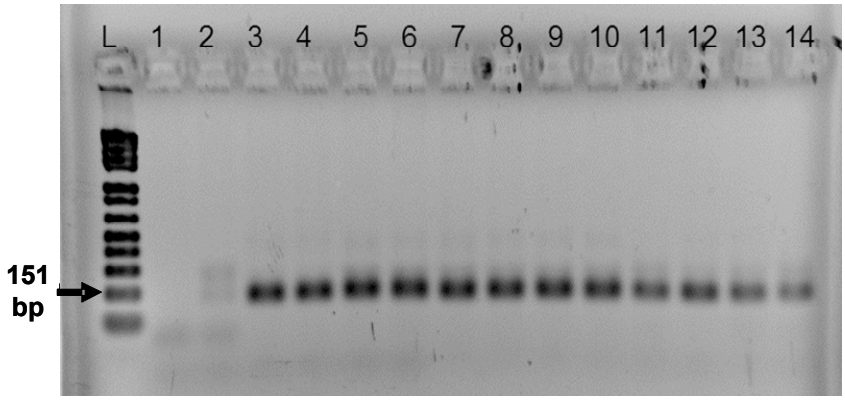


Figure 3.1. Gel image with typical *Didymosphenia geminata* quantitative PCR (D602F and D753R) amplification products from duplicate reactions of seven Buller River samples referenced to a DNA ladder (L).

3.2.2 High resolution melt (HRM) analysis

The next level of validation employs a method that allows us to carefully monitor the characteristics of a QPCR amplification product as it denatures with elevated temperature and compare these characteristics to known *D. geminata* QPCR standards. Double-stranded DNA will denature (melt) in a progressive, predictable, and reproducible way with increasing temperature. In effect, the double strands of DNA will "unzip" (denature or melt) as the temperature is increased and "re-zip" (anneal) together as the temperature is lowered. The timing and temperature at which this melting happens depends on the length and actual base pair sequence of the DNA, so much so that by visualising the characteristics of this process one can distinguish DNA populations that differ by only a single base pair (Ririe *et al.,* 1997). The process is visualised by adding a dye to the DNA that intercalates the double strand DNA rendering it fluorescent when excited with the appropriate wavelength of light. Dye not intercalated will not fluoresce. As denaturation takes place and the strands slowly unzip, the intercalated dye is released and assumes its non-fluorescent state. By monitoring the total fluorescence during the melting process we can calculate the point at which 50% of the DNA is denatured or where 50% of the dye has been released. This point is known as the T_m. *D. geminata* has a very specific T_m of 84.5 °C (Figure 3.2). A variation of +/– 2°C would indicate that the QPCR amplicon was not *D. geminata* and would warrant further investigation.

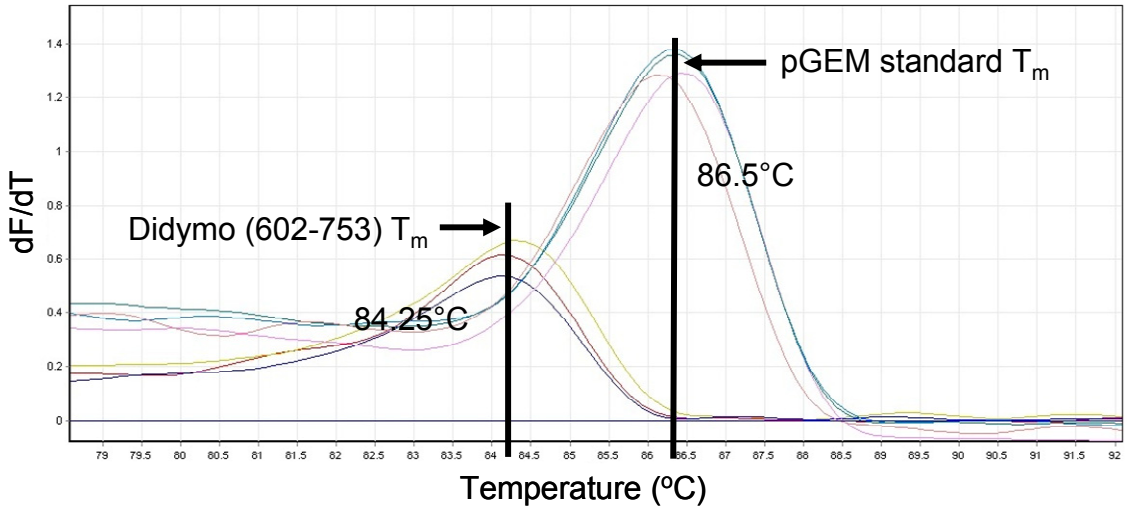


Figure 3.2. High resolution melt analysis illustrating the difference in T_m between the pGEM standard and the *Didymosphenia geminata* HRM results.

The Rotor-Gene 6000 QPCR instrument has the ability to conduct HRM analysis. By incorporating the intercalating dye into the master mix of the QPCR reaction or adding it after the reaction is complete, we can in a matter of minutes perform HRM analysis on all of the samples. This has now become our standard protocol for all *D. geminata* and pGEM QPCR reactions.

3.2.3 Sequence analysis

The third and most conclusive validation level comes from direct sequencing of PCR products generated from all positive QPCR reactions. Because the QPCR reactions use a

fluorecent probe that would interfere with sequencing, a second PCR reaction without the fluorescent probe must be run specifically to obtain sequencing template. We normally use the standard D602F and D753R primers as used in the QPCR reactions. However, recently we have been using D602F and the *D. geminata*-specific D1670R primer to obtain a longer segment to sequence. Each of the resulting sequences were aligned with the original *D. geminata* sequence and manually checked for similarity. Any inconsistencies in the sequence were checked and validated against the original sequencing chromatograms.

3.3 Protocols in brief

Sampling

Initially we set out to sample three different locations of each of the then 13 infected rivers in the South Island, one at the site of first discovery of *D. geminata* (the lower Waiau River, Southland) and then at the two flanking ends of distribution. This would provide adequate coverage of each river and supply the necessary positive control samples for the initial stage in the validation process. As newly infected sites were discovered by our end-users or during the scheduled delimiting surveys, we added these to our pool of positive samples. All international samples (28 total) collected for the phylogeography study were also included in the validation process. We also carried out "trial runs" on samples of unknown status provided by DOC Southland, by NIWA researchers (Kilroy and Dale, 2006) and from the North Island delimiting surveys (Duncan et al., 2007). All samples considered in the validation process were collected using the drift net protocol. Collectively we analysed 174 samples, 149 of which we validated as positive for *D. geminata*, originating from 62 rivers (48 in the South Island of New Zealand and 14 international rivers).

Gel electrophoresis

To check for the appropriate amplicon size, all positive QPCR reaction products were run out on a 1% agarose TAE gel at 100 volts for 20 min. The gels were then stained with ethidium bromide for 20 min, destained in water for 20 min, and then visualised on a UV light source as described in our interim report (Cary *et al.*, 2006).

High resolution melt analysis (HRM)

Our HRM analysis capability came on line in early December 2006. Since that time all QPCR experiments terminated in an HRM analysis. HRM analysis was carried out subsequently to PCR by adding 5 µM of SYTO 9 Green (Invitrogen, New Zealand) to the PCR reactions. The analysis was performed similar to the melting point analysis using the SYBR Green I dye (Section 2.3) with the difference that only the HRM channel was used for signal acquisition. The PCR products were melted using 0.5°C increments between 60°C to 99°C.

Sequence analysis

PCR products were generated from the DNA from each of the QPCR positive samples using the D602F-D753R or the D602F-D1670R primers sets. The PCR for subsequent sequencing was carried out in 0.2 mL thin walled tubes, containing a total volume of 50 µL and a final concentration of 2.5 mM $MgCl_2$ (Roche, Germany), 1X PCR buffer - $MgCl_2$ (Roche, Germany), 0.2 mM of each dNTPs (Roche, Germany), 2.0 units of Taq DNA polymerase (Roche, Germany), 0.2 µM *D. geminata*-specific primers (IDT, USA), 0.24 µg BSA (Sigma, New Zealand), and 40 ng of genomic DNA (Appendix 3). Experiments were performed in a MJ Research Peltier thermal cycler under the following parameters: an initial denaturation step at 94°C for 2 min and then 40 cycles of 94°C for 20 secs, 55°C for 1 min,

and 72°C for 2 mins. A final extension step of 72°C for 5 mins followed to complete the experiment. To confirm the success of the experiment, 5 μL of each reaction was run on a 1% agarose gel with sodium boric acid (SBA) buffer, stained with ethidium bromide, and visualised with UV transillumination. The remainder of each reaction was purified using a Genscript PCR purification kit and quantified by spectrophotometry. Purified PCR products were then submitted to the University of Waikato DNA Sequencing Unit for direct sequencing, using the corresponding *D. geminata*-specific primers at 5 μM. Upon receipt of sequencing results, sequences were imported into the alignment tool DNAstar and aligned to the known *D. geminata* sequence (see Appendix 3 for the full protocol).

3.4 Results

To date (6 June 2007), we have examined 174 samples from 76 rivers by QPCR, 54 from the South Island, 8 from the North Island, and 12 international (Norway 4, Canada 2, England 1, and USA 5; Table 3.1). 149 samples were positive for *D. geminata* (i.e., above threshold). All of these positive samples have been validated to be *D. geminata* using gel electrophoresis and direct sequence analysis. As would be expected, in all cases the amplification product produced during these QPCR runs, when subjected to gel electrophoresis, was of the correct size (151 bp) for *D. geminata.* In all cases, we saw <1% DNA sequence variation in any one sample when compared to the original *D. geminata* sequence. This level of variability is consistent with normal sequencing error. No North Island samples were positive for *D. geminata*.

Table 3.1. Summary of analyses using the DNA Method carried out on New Zealand and international samples to 30 June 2007. HRM, high resolution melt analysis.

Region	Number of rivers analysed	Number of samples			
		Analysed for *D. geminata*	Positive for *D. geminata*	Sequences of *D. geminata*	HRMs carried out
North Island, NZ	8	9	0	0	6
South Island, NZ	54	137	121	93	58
International	12	28	28	28	0
Total	74	174	149	121	64

Since the introduction of HRM analysis to the validation process in December 2006, we have analyzed over 60 samples. In all cases, the QPCR amplicon from each of the tested samples produced a melt curve or T_m in the region consistent with *D. geminata* (~83°C*)*. However, we have observed that the actual T_m might shift as much as ±1°C depending on the concentration of template in the original reaction. This shift does not occur every time but appears related to the template concentration and sample quality. We are continuing to investigate this shift but feel confident that with the redundancies built into the validation protocols that any shift can be reconciled.

3.5 Conclusions and recommendations

The primary reason for developing a *D. geminata* QPCR Taqman protocol was to attain the highest level of specificity possible for the assay. As mentioned above, with over 100

validated positive QPCR reactions, we have yet to detect a false positive. Using each of the validation approaches, the QPCR method appears highly specific to *D. geminata*.

HRM analysis provides a rapid and highly sensitive alternative to direct sequencing. Major differences in sequence and/or length of the QPCR product can be detected. However, under our use, we detected a noticeable variability in the T_m for *D. geminata.* While we have not been able to ascertain the reason for the variability, we are confident that when used in concert with either the gel method or direct sequencing, validation can be achieved with a high degree of assurance. We hope that with further testing we will be able to understand the HRM variability to the extent that this method will become definitive and replace the more expensive and time consuming direct sequencing method.

Our protocol, while very specific and highly optimized, still remains experimental as long as samples from new and untested locations are being analyzed. As such, the validation process using gel, HRM, and sequence analysis will remain part of the QC/QA protocol. Every positive QPCR sample will still be subjected at least to HRM and if necessary sequence analysis for the foreseeable future.

From the extensive analysis described above, we feel the DNA Method has been validated to the extent necessary to release the protocol to our end-users and begin full implementation in New Zealand.

4 Preliminary phylogeography study of *Didymosphenia geminata* (Lyngbye) M. Schmidt: using molecular markers to reveal geographic history

4.1 Introduction and objectives

4.1.1 Internal transcribed spacers (ITS)

The goal of this preliminary study was to identify a molecular marker with appropriate levels of genetic variation to reconstruct phylogeographic patterns among samples of *Didymosphenia geminata* (Lyngbye) M. Schmidt. These samples were collected both from New Zealand sites and globally. This will enable identification of the origin(s) of these samples of this invasive organism and help in understanding the pattern(s) of genetic variation within and among samples in New Zealand. Provan and co-workers (2005) successfully used a similar approach to recover the invasive history of another alga, *Codium fragile*.

D. geminata is a morphologically variable species with a problematic taxonomic history (Antoine and Benson-Evans, 1984; Metzeltin and Lange-Bertalot, 1995). Currently it is placed within the Class Bacillariophyceae, Order Cymbellales, in the Family Gomphonemataceae, which also includes the genera *Gomphoneis, Gomphonema, Gomphopleura* and *Reimeria* (The Taxonomicon).

Previous molecular phylogenetic studies using 18S ribosomal (e.g., Beszteri *et al.*, 2001; Rynearson and Armbrust, 2004) and mitochondrial (Ehara *et al.,* 2000) DNA sequence variation have been useful at higher, infra-familial levels. These studies indicate that these markers would not provide sufficient variation at the specific and intra-specific levels. As such, we turned our attention to the internal transcribed spacers (ITS) of nuclear ribosomal DNA (nrDNA; Figure 4.1). This marker has been used at lower taxonomic levels in other diatoms to resolve relationships among species and populations of algae (Kooistra *et al*., 1992) and more specifically among diatoms (Zechman *et al*., 1994). ITS reveals semi-cryptic species (Créach *et al.,* 2006), genetic variation in sexual compatibility types (Behnke *et al*., 2004), and, importantly, variation among geographically isolated samples (Godhe *et al*., 2006).

Our approach has been to develop novel methods of investigation to enable direct amplification of molecular markers from environmental samples and to bypass the lengthy and difficult stages of culturing. To avoid co-amplification of contaminants (e.g., Stevens *et al.,* 2007), we employed *D. geminata*-specific primers (Cary *et al.,* 2006; Table 4.2) in combination with universal ITS primers (Figure 4.1).

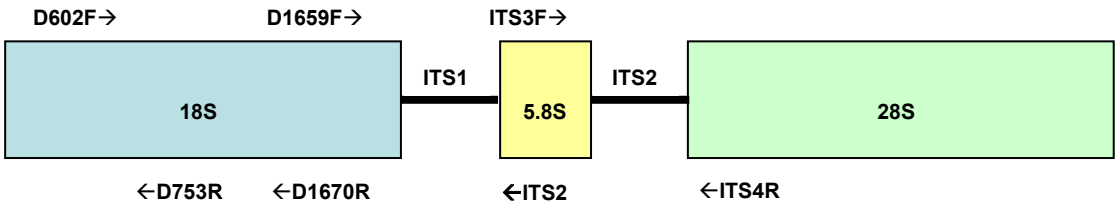


Figure 4.1. Schematic of nrDNA regions targeted for this study. The relative position of primers is indicated above (forward) and below (reverse) the conserved genes (18S, 5.8S, and 28S).

4.1.2 RAPD Analysis of *D. geminata* samples in conjunction with whole genome amplification

A second, longer term goal is to develop *D. geminata*-specific polymorphic markers, such as random amplified polymorphic DNA (RAPDs) and inter simple sequence repeats (ISSRs), to further assess levels of variation at a finer scale. To be able to do this, we are conducting whole genome amplification (WGA) of *D. geminata*. We will then use this contaminant-free DNA to develop *D. geminata*-specific sequence characterised amplified region (SCAR) markers from amplification products generated via RAPD and ISSR analyses (e.g., Bornet *et al.*, 2005). This approach will be less time-consuming and less costly than developing *D. geminata* specific microsatellite markers. However, if co-dominant markers are desired, then microsatellites could be developed from the DNA produced via the WGA.

A rapid population-level genetic analysis of *D. geminata* samples is desirable for tracing the domestic and/or international source of incursion. We have chosen the RAPD method as our prime candidate due to its general applicability and relatively neutral representation of community members. Compared with similar techniques such as automated ribosomal spacer analysis (ARISA) and denaturing gradient gel electrophoresis (DGGE), RAPD requires only standard molecular genetics lab equipment (e.g., PCR thermocycler, gel electrophoresis apparatus) and is not prone to PCR bias typical of gene or locus-specific analyses. The method enables a genomic level analysis of individuals or groups of individuals from the same population. These genome surveys of organisms collected from different geographic regions can then be compared allowing identification of subtle variations in population structure. The method relies on the use of a single eight-nucleotide random PCR primer and its ability to reproducibly amplify, in both the forward and reverse direction, sections of the genome. The resulting amplification product, when run out on a standard agarose gel, will produce an identifiable and reproducible fingerprint of the genome.

The RAPD method requires that the sample be extremely clean and representing only the organism in question. For *D. geminata*, this is problematic as it cannot be cultured and often occurs in a complex consortium in nature. We therefore elected to hand pick cells from each sample and use a new Whole Genome Amplification protocol to acquire enough *D. geminata* DNA to perform the RAPD analysis. The WGA technique is based on strand displacement amplification using Phi29 DNA polymerase (Lasken and Egholm 2003). Eventually, this will be performed on each of the international samples and several New Zealand samples, and the fingerprints compared.

4.2 Protocols in brief

4.2.1 Internal transcribed spacers (ITS)

Samples used in this preliminary study were previously collected and stored in ethanol or CTAB (Table 4.1; Cary *et al.*, 2006). For the purposes of this study we routinely use the DNeasy Plant kit (Qiagen).

Table 4.1. Collection locations and dates for samples *Didymosphenia geminata*.

Code	River	Region	Date collected	Longitude (°)		Latitude (°)	
LW1	Lower Waiau	Southland, NZ	29-May-06	167.678931	E	45.63587	S
BC-5	Boulder Creek	Montana, USA	15-May-06				N
C-1	River Coquet	UK	4-May-06	1.89833611	W	55.33753	N
N-2	Atna	Norway	14-Sep-06	10.8330889	E	61.7317	N
OR3-E1	Oreti River	Southland, NZ	11-May-06	168.138847	E	45.53235	S

Addition of Proteinase K in our extraction protocols improved the yields from samples stored in ethanol. All DNA extractions were evaluated for quantity and quality of DNA by electrophoresis on 1% agarose TBE gels.

We routinely made small modifications to the amplification protocol based on prior PCR results and template quantity. Standard PCR mixtures contained 1X PCR buffer ($MgCl_2$ + and $MgCl_2$–; Roche), 2.5 mM $MgCl_2$, 0.1 mM dNTPs (Roche), 0.2 μM of each primer (Table 4.2), <5% BSA (Sigma), and 0.3 U AmpliTaq PCR-polymerase (Applied Biosystems). The amount of DNA used in the PCRs varied depending on the quality and quantity and prior results. PCR products were electrophoresed on 1% agarose TBE gels and photographed. *D. geminata*-specific primers complimentary to the 18S region of nrDNA (Cary *et al*., 2006) were used as a starting point to amplify most of the 18S and the entire ITS regions. We tested each pair of primers (Figure 4.2) to make certain that 1) we had a single ITS product which is indicative of targeted amplification of *D. geminata* and that we did not amplify putative contaminants, and 2) that the PCR products were of the correct relative size. We have been able to PCR amplify and directly sequence the nrDNA target regions for a number of samples. To facilitate better sequencing reads and to again verify our PCR products, we have also cloned the entire D602F to ITS4R PCR fragment (ca. 2KB) region for three samples, BC5, LW1 and OR3. Cloning of the PCR product was performed with a TOPO TA cloning kit (Invitrogen, New Zealand) following the manufacturers instructions. PCR products were purified for sequencing using PEG or ExoSAP. Sequencing was carried out at the University of Waikato DNA Sequencing Unit.

Table 4.2. Primer sequences used to amplify partial 18S rDNA and ITS regions for *Didymosphenia geminata. D. geminata* primers were developed by Cary *et al*. (2006). Universal primers are from White *et al.* (1990).

Primer	Sequence (5' to 3')	Specificity
D602F	GTT GGA TTT GTG ATG GAA TTT GAA	*Didymosphenia geminata*
D753R	AAT ACA TTC ATC GAC GTA AGT C	*Didymosphenia geminata*
D1565F	CCT AGT AAA CGC AGA TCA TCA G	*Didymosphenia geminata*
D1659F	GCT GGG GAT TGC AGC TA	*Didymosphenia geminata*
D1670R	CAC CAG TAA AGG CAT TAG CTG	*Didymosphenia geminata*
ITS2R	GCT GCG TTC TTC ATC GAT GC	universal
ITS3F	GCA TCG ATG AAG AAC GCA GC	universal
ITS4R	TCC TCC GCT TAT TGA TAT GC	universal

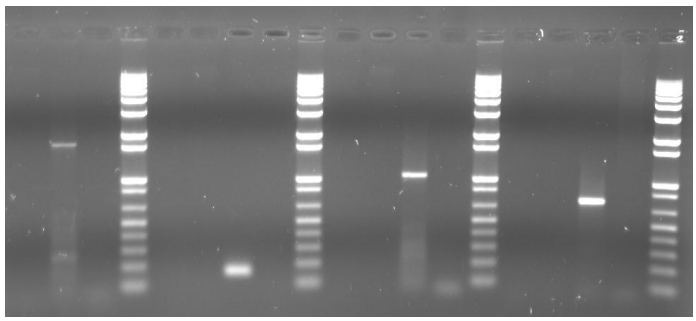


Figure 4.2. nrDNA PCR products for the various primer combinations for *Didymosphenia geminata*. All products were of the expected size. Lane 1, D602F and ITS4R, ca. 2KB; lane 2, ladder; lane 3, D602F and D753R, ca. 150bp; lane 4, ladder; lane 5, D602F and 1670R, ca. 1000bp; lane 6, ladder; lane 7, D1659F and ITS4R, ca. 850bp; lane 8, ladder.

To obtain high quality, reliable DNA sequences we obtained both forward and reverse sequencing reactions. We also worked towards complete overlaps of these reads to obtain the highest confidence in sequence variation that is possible. Sequences were aligned and edited by eye using Sequencher 4.1. All final sequences have been compared to those in GenBank via the BLASTn search algorithm. Final DNA sequences will be deposited at GenBank. Phylogenetic analyses using 18S sequences from our samples and related species from GenBank (Table 4.3) were conducted using Maximum Likelihood methods as implemented in PAUP* 4.1. We conducted two levels of analysis, one focusing on intraspecific relationships among samples of *D. geminata* comprised of partial 18S+ITS data, and to put this work into context, we also conducted higher-level analyses among different genera of naviculoid diatoms using partial 18S data. *Fragilaria striatula* was designated as the outgroup taxon. We performed 100 bootstrap replicates to indicate support for the various clades.

Table 4.3. Genbank accessions for samples used in the 18S analysis.

Species	GenBank accession number
Amphora montana	AJ243061
Anomoeoneis sphaerophora	AJ535153
Dickieia ulvacea	AY485462
Encyonema triangulatum	AJ535157
Eolimna minima	AJ243063
Eolimna subminuscula	AJ243064
Fragilaria striatula	AY485475
Gomphonema parvulum	AJ243062
Navicula cryptocephala	AJ297724

4.2.2 RAPD Analysis of *D. geminata* samples in conjunction with whole genome amplification

Didymosphenia geminata is a large diatom and can be easily hand-picked using micro-pipettes and an inverted microscope. In order to run a RAPID analysis on *D. geminata*, we needed to collect at least 100 individual cells from any given sample. Since RAPD works with any genomic DNA, individual *D. geminata* cells must be inspected and washed to minimize bacterial contamination. Aliquots of selected samples were placed into the back well of a 24-well plate in which each well was filled with 500 µL deionized water. Individual cells were hand-picked using a micro-pipette and transferred to a new well, this process was repeated several times. Washed cells were transferred to a 1.5 mL Eppendorf tube.

The washed cells were resuspended in 1 mL of deionized water and sonicated in an ultrasonication bath for 2 min. Samples were gently pelleted by centrifugation (2,500 g, 2 min), and 800 µL of supernatant was removed. This sonication process was repeated again before DNA was purified from the remaining solution using a QIAamp DNA Mini Kit (QIAGEN) following the manufacturer's directions. Negative controls were included at each step of the purification process. However, since the number of *D. geminata* cells that can be practically processed in this manner is limited (typically around 100 cells), the amount of DNA that can be extracted from the cells is likely to be insufficient for downstream genome-wide analyses. This problem is solved by using WGA based on Phi29 DNA polymerase,

since a typical WGA reaction yields 4-7 μg of DNA from 10 ng of sample and does so in a representative manner.

RAPD reactions are PCRs performed with an arbitrary primer (e.g., a random octamer), which generates a collection of DNA fragments specific for any given genomic DNA sample. This collection of DNA fragments forms a unique pattern of bands when analyzed with DNA agarose gel electrophoresis analysis. A selection of random primers will be tested with *D. geminata* samples to identify one or two primers that give patterns with clear and well-separated bands. The selected primer(s) will be used to amplify various environmental *D. geminata* samples, and the results will provide us with useful and representative information about the *D. geminata* communities.

4.3 Results

4.3.1 Internal transcribed spacers (ITS)

When we combined the *D. geminata* primers with universal primers used routinely to amplify the internal transcribed spacer (ITS) regions of nrDNA, we were able to selectively amplify *D. geminata.* This ca. 2KB product comprised most of the 18S gene together with the ITS regions (partial 18S+ITS) (Figure 4.1).

BLASTn results provided high confidence that we have successfully sequenced a unique diatom (Figure 4.3). Our sequence was nested within other sequences of diatoms and sister to two species of a genus, *Gomphonema*, thought to be closely related based on morphology. Based on these results, and visual, microscopic examination of our collections, we have confidence that we have sequenced *D. geminata.*

Pairwise genetic distances values within *D. geminata* were small (range 0.003-0.012), with both New Zealand samples being most similar to the sample from Boulder Creek, Montana, USA (Table 4.4; Appendix 4). Pairwise genetic distances across the naviculoid diatom partial 18S data ranged from 0.000 (between *D. geminata* samples from the UK and USA) to 0.1189 (between *F. striatula* and *Cymbella* sp.) (Table 4.5).

Table 4.4. Pairwise genetic distances of *Didymosphenia geminata* for partial 18S+ITS sequence. BC5, Boulder Creek, Montana, USA; LW1, Lower Waiau, Southland, New Zealand; N2, Nidelva, Norway; OR3, Oreti River, Southland, New Zealand; UKC1, River Coquet, UK.

	BC5	LW1	N2	OR3
BC5				
LW1	0.003			
N2	0.009	0.011		
OR3	0.005	0.007	0.012	
UKC1	0.004	0.006	0.009	0.006

The partial 18S+ITS sequence data resolved relationships among samples of *D. geminata* (Figure 4.4). The two New Zealand samples form a weakly supported clade (BS=56) along with the sample from Montana, USA. All *D. geminata* samples form a strongly supported clade (BS=100) in the partial 18S analysis (Figure 4.5) and are sister to *Anomoeoneis sphaerophora*. *Gomphoneis minutae v*ar. *cassiae* and *Gomphonema parvulum* form another well supported clade (BS=98) and are sister to *Cymbella* sp.

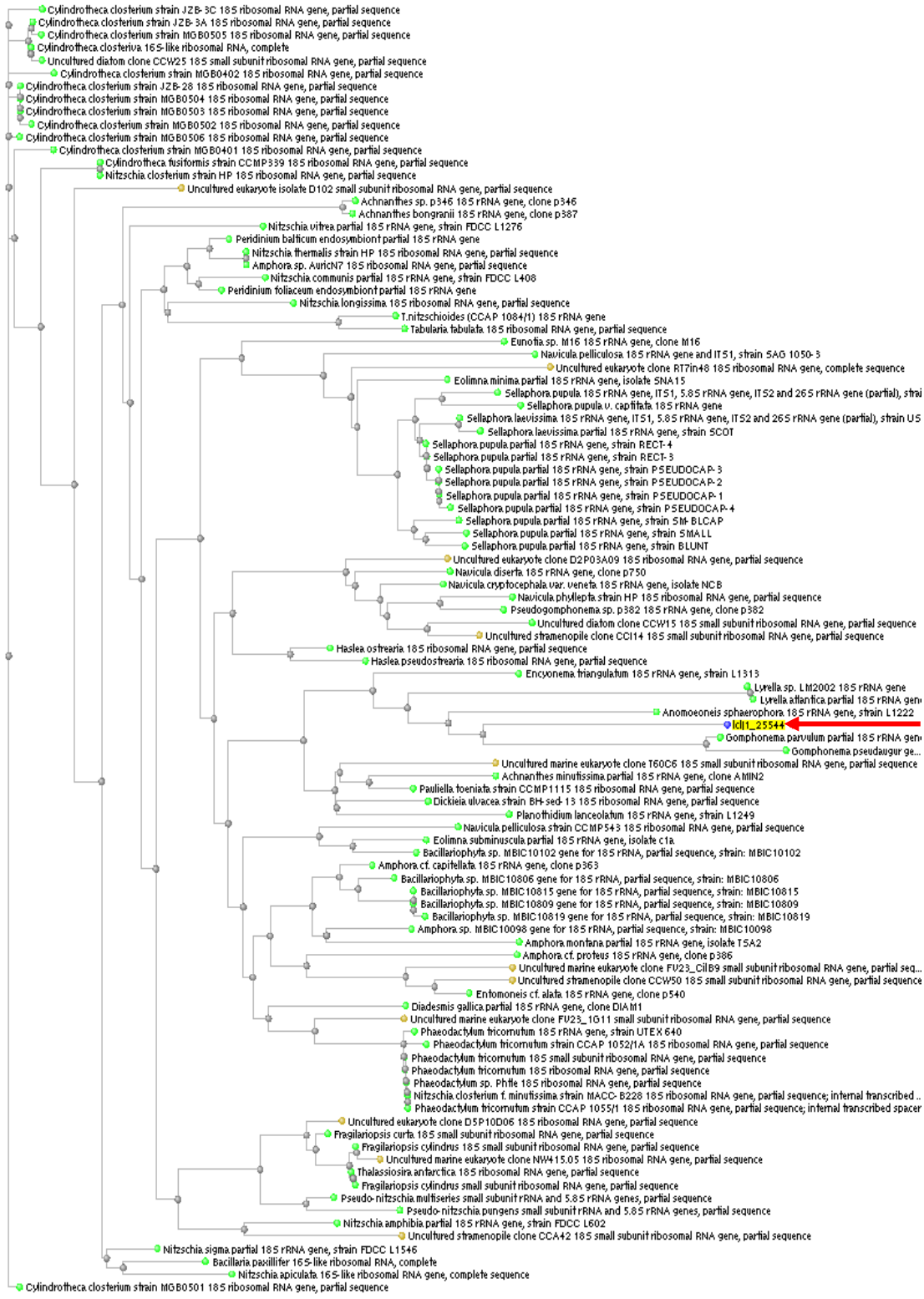


Figure 4.3. Neighbor Joining tree of the BLASTn search results. Note that our partial 18S + ITS sequence of *Didymosphenia geminata* (UKC1, denoted with arrow) is nested deeply within a clade exclusively composed of diatom species (denoted by green dots), and is sister to two *Gomphonema* spp.

Table 4.5. Pairwise genetic distances of naviculoid diatoms for partial 18S sequence. BC5, Boulder Creek, Montana, USA; LW1, Lower Waiau, Southland, New Zealand; N2, Nidelva, Norway; OR3, Oreti River, Southland, New Zealand; UKC1, River Coquet, United Kingdom.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1 *Didymosphenia geminata* LW1															
2 *Didymosphenia geminata* OR3	0.0032														
3 *Didymosphenia geminata* BC5	0.0011	0.0022													
4 *Didymosphenia geminata* N2	0.0086	0.0097	0.0076												
5 *Didymosphenia geminata* UKC1	0.0011	0.0022	0.0000	0.0076											
6 *Gomphoneis minutae* var *cassiae*	0.0553	0.0564	0.0543	0.0618	0.0542										
7 *Gomphonema parvulum* AJ243062	0.0467	0.0477	0.0457	0.0532	0.0456	0.0248									
8 *Cymbella* sp.	0.1042	0.1053	0.1033	0.1108	0.1031	0.0998	0.0965								
9 *Amphora montana* AJ243061	0.0640	0.0640	0.0630	0.0683	0.0629	0.0651	0.0618	0.1176							
10 *Anomoeoneis spaerophora* AJ535153	0.0380	0.0391	0.0370	0.0445	0.0369	0.0565	0.0457	0.1037	0.0608						
11 *Dickieia ulvacea* AY485462	0.0432	0.0443	0.0423	0.0497	0.0422	0.0445	0.0423	0.0968	0.0412	0.0358					
12 *Encyonema triangulatum* AJ535157	0.0464	0.0475	0.0455	0.0529	0.0454	0.0401	0.0402	0.0945	0.0466	0.0413	0.0260				
13 *Fragilaria striatula* Ay485474	0.0705	0.0694	0.0695	0.0737	0.0694	0.0673	0.0662	0.1189	0.0599	0.0664	0.0455	0.0478			
14 *Eolimna minima* AJ243063	0.0594	0.0583	0.0585	0.0648	0.0583	0.0563	0.0595	0.1172	0.0411	0.0531	0.0335	0.0389	0.0512		
15 *Eolimna subminuscula* AJ243064	0.0605	0.0595	0.0596	0.0660	0.0595	0.0587	0.0490	0.1113	0.0401	0.0466	0.0259	0.0379	0.0490	0.0357	
16 *Navicula cryptocephala* AJ297724	0.0578	0.0589	0.0568	0.0643	0.0567	0.0589	0.0535	0.1117	0.0448	0.0491	0.0360	0.0404	0.0457	0.0436	0.0359

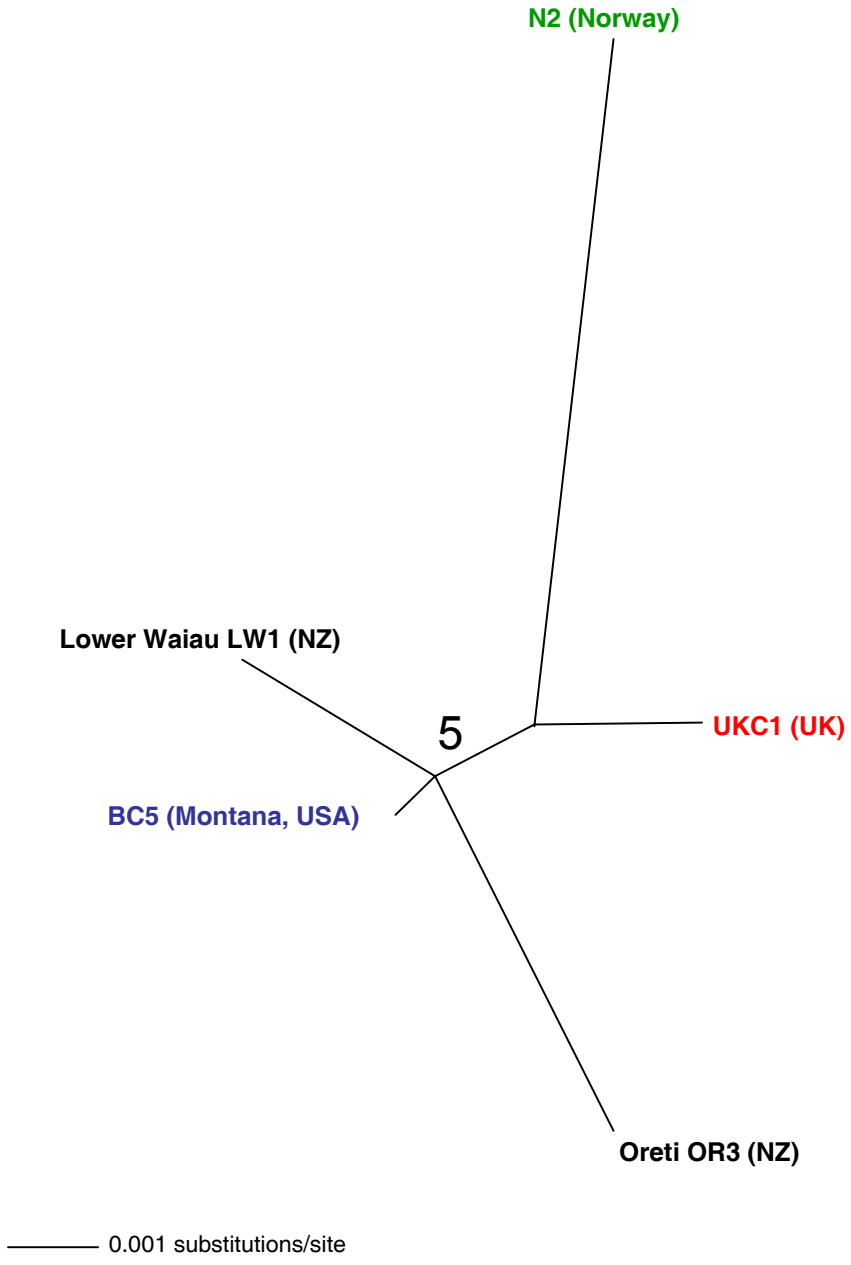


Figure 4.4. Unrooted maximum likelihood phylogenetic tree of the partial 18S+ITS sequence for *Didymosphenia geminata*. Bootstrap values are given above branch. BC5, Boulder Creek, Montana, USA; LW1, Lower Waiau, Southland, New Zealand; N2, Nidelva, Norway; OR3, Oreti River, Southland, New Zealand; UKC1, River Coquet, United Kingdom.

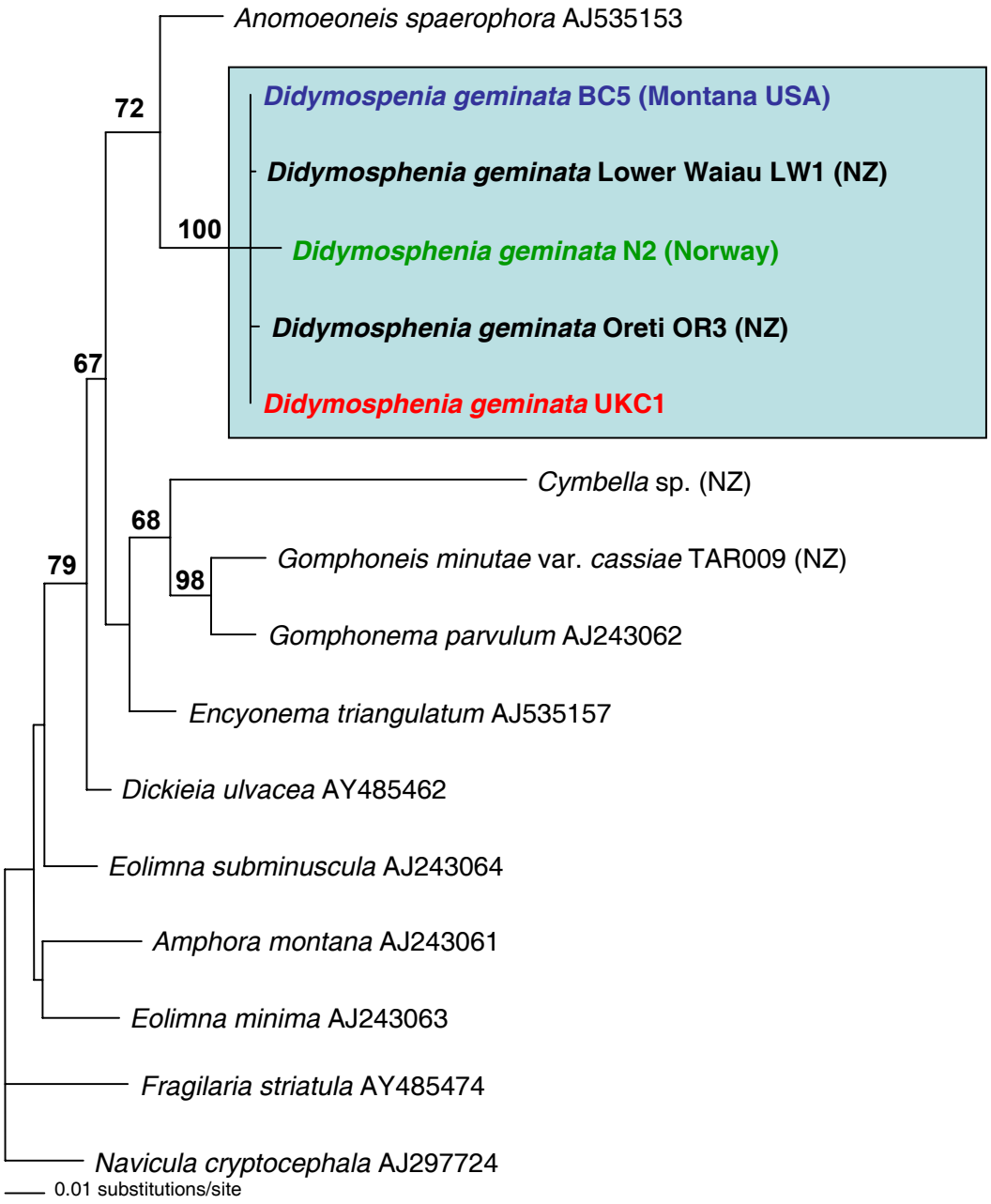


Figure 4.5. Rooted maximum likelihood phylogenetic tree of the partial 18S sequence data for five samples of *Didymosphenia geminata* and other diatom genera (accession numbers are given for those obtained from GenBank). Bootstrap values are given above branches.

4.3.2 RAPD Analysis of *D. geminata* samples in conjunction with WGA

We have successfully isolated *D. geminata* cells from the environment, cleaned the samples, and extracted DNA from the cells. The thoroughness of the cleaning process was verified by two QPCR assays. The first used universal bacterial 16S rDNA primers. No notable amplification was observed. The DNA sample was then verified to contain *D. geminata* using our *D. geminata* SYBR green I assay. Completion of the RAPD analysis will require further time and resources.

4.4 Conclusions and recommendation

4.4.1 Internal transcribed spacers (ITS)

The partial 18S+ITS region provides sufficient level of resolution to reveal the phylogeographic history and origin(s) of *D. geminata* in New Zealand (Figure 4.4). The partial 18S data allows for resolution at higher taxonomic levels (Figure 4.5). The partial 18S+ITS region should continue to be very informative in future, allowing more detailed phylogeographic studies of *D. geminata*. Other markers are available if the limit of resolution is reached with ITS (e.g., Provan *et al*., 2004; Provan *et al*., 2005; Zechman *et al.,* 1994).

Future detailed analyses, based on morphology, would determine if *D. geminata* is a morphologically variable species, or if it represents a polymorphic complex of cryptic species. Metzeltin and Lange-Bertalot (1995) suggest that *D. geminata* is composed of three morphotypes, *geminata s.s.*, *capitata* and *subcapitata*. Specimens from the Lower Waiau River have been identified as *D. geminata* morphotype *capitata* (Kilroy, 2004) which is the morphotype found across Northern Europe (D. Metzeltin pers. comm., as cited in Kilroy, 2004). However, in our analyses, the two *D. geminata* samples collected in New Zealand are more similar genetically to the sample from Montana, USA, than to either of the two samples from Europe. Morphological analyses are needed to accompany future phylogeographic genetic comparisons. Additional fine-scale sampling and a combined approach of morphological and molecular analyses would allow us to address this issue in a robust manner. Relationships and taxonomic clarification of the other conspecific species (*D. siberica*, *D. curvata*, *D. clavaherculis*, and *D. pumila*, sensu Metzeltin and Lange-Bertalot (1995)) is also in need of further investigation and revision.

Given the small dataset and the preliminary nature of the phylogenetic analyses, caution is required in their interpretation. However, it is clear that the samples from the UK and Norway are distinct from the other samples. To achieve greater confidence in these results and to be able to infer more meaningful relationships, our sequence database needs to be expanded with samples from *within* and *between* river systems, both in New Zealand and globally.

4.4.2 RAPD Analysis of *D. geminata* samples in conjunction with WGA

The use of RAPD in conjunction with WGA is likely to provide us the ability to examine *D. geminata* population structure and movement with great confidence and sensitivity. However, as agreed in our contract with MAF Biosecurity New Zealand, only preliminary phylogeographic work was expected to be completed in the short timeframe allowed and along with the competing, higher priority of developing a DNA detection method for surveillance. Future work will be required to complete the whole genome amplification and RAPD analysis. The groundwork for the completion of this study has been laid, and we are prepared to optimize the protocol and begin to apply the methods on various *D. geminata* samples within the next few months.

5 QPCR-compatible field sampling protocols

5.1 Introduction

In this section we address the sampling requirements for increasing sensitivity of the DNA based detection method, minimising the possibility of contamination, and improving ease of sampling. Though 40 µm is a standard size for phytoplankton sampling, and has been shown to effectively trap drifting cells of *D. geminata*, such fine mesh size has the complicating factor of trapping other drifting material (both organic and inorganic). This has the potential to reduce the success of identifying *D. geminata* genomically by interfering with the DNA extraction process through exogenous DNA and inhibitory substances, e.g., humic acid. Thus our aim was to focus sampling on the typical *D. geminata* cell sizes (35 µm wide x 100-130 µm long) and reject larger or smaller material. Our previous report detailed the addition of a 250 µm prefilter to the front of the sampling nets (Cary *et al.,* 2006). We have adopted a 250-µm prefilter as a standard procedure for all net sampling for *D. geminata* in rivers.

In early 2007, concerns were raised that procedures recommended by MAF Biosecurity New Zealand (immersion in 2% household bleach for 1 min) to kill *D. geminata* and thereby decontaminate sampling gear might still result in false positives when used with the DNA detection method. Although this concentration of bleach kills the *D. geminata* cells, it does not necessarily degrade DNA. Thus, DNA in dead *D. geminata* cells from one site might remain in a sampling net, contaminating the next sample from a new site. While dead cells will not invade the new site, the extreme sensitivity of our Taqman QPCR techniques means that we needed to assess whether DNA from dead cells is still extractable and able to be amplified. The purpose of these experiments was to determine the concentration and incubation time of household bleach required to completely degrade DNA of *D. geminata.* We aimed to develop a robust decontamination method for field sampling equipment for use with the molecular detection methods described in this report.

Further, decontamination procedures have so far relied on the use of Janola brand bleach, which is a premium product, with active ingredients in addition to the main ingredient sodium hypochlorite, and although widely available, is also one of the most expensive brands. We carried out a limited comparison of the amount of active ingredient in other brands of bleach available in New Zealand supermarkets to assess their relative effectiveness at degrading DNA.

The objectives were:

1. To test the effectiveness of 40, 75, and 100 µm mesh sizes for the net of the sampling sock.
2. To test the effectiveness of 2-5% bleach to degrade *D. geminata* cells under immersion times of 1-15 mins.
3. To compare the concentration of the active ingredients in commercially available household bleaches.

5.2 Protocols in brief

5.2.1 Net design and deployment

Triplicate drift samples were collected in 240-mm diameter plankton nets of mesh sizes 40, 75, and 100 µm fitted with a 250-µm mesh prefilter to exclude larger material. The nets were suspended for 2 mins at site BU-8 in the Buller River at Howard River (41º 43.226' S, 172º 41.131' E), South Island, New Zealand on 31 October 2006. The water velocity at the net mouth was 0.69 m/s. Each sample was preserved separately in 70% ethanol. Cell enumeration was carried out on

each sample using an inverted microscope (Olympus CKX41) and an Utermöhl settling chamber (Utermöhl 1958). Triplicate counts were carried out on each sample.

5.2.2 Net decontamination and DNA denaturation

Denaturation Experiment 1: In the first experiment, samples of viable *D. geminata* from Southland were incubated in the field for 15 mins with various concentrations of commercially available Janola regular premium bleach (4.2% sodium hypochlorite, i.e., w/v 42 g/L). The extent of DNA degradation was measured using standard PCR followed by gel electrophoresis and Taqman QPCR. One additional incubation was carried out using 2% Janola for 1 min to determine if the current MAF Biosecurity New Zealand protocol to kill *D. geminata* on angling equipment is also robust enough to denature *D. geminata* DNA, thus completely removing the DNA signal.

Janola-treated *D. geminata* samples were washed in laboratory water and preserved in 70% ethanol for shipping. In the laboratory, the *D. geminata* suspensions were centrifuged and the pellets transferred to 1.5 mL Eppendorf tubes. The remaining ethanol was allowed to evaporate for 1 hour under laminar flow. Subsequently, the samples were CTAB-extracted following the protocol outlined in Cary *et al.,* (2006; Appendix 2) and subjected to Taqman QPCR and standard PCR followed by gel electrophoresis. Three control samples were included in this study: A) a field-positive control sample with no Janola treatment but with CTAB extraction, B) a PCR-positive control containing 1 pg of a plasmid containing the *D. geminata* 18S rRNA gene and C) a PCR-negative control containing no added template DNA.

Denaturation Experiment 2: This experiment was designed to give clear guidance to the delimiting teams due to go out in early May 2007. Its aim was to see if delimiting survey teams could use a single bleach concentration that would meet the needs of both decontamination (i.e., killing *D. geminata* so it cannot be transferred between sites) and denaturation (i.e., complete destruction of *D. geminata* DNA so that amplification is impossible).

In this new experiment, we exposed *D. geminata* DNA to two different bleach types held at 2% concentration for 1, 5, 10, or 15 mins. We aimed to see if a bleach concentration less than 5% could denature *D. geminata* DNA when incubated for 15 mins or less.

The tests used Janola regular premium bleach (4.2% sodium hypochlorite, i.e., w/v 42 g/L) and Chlor-O-Gene bleach (3.15% sodium hypochlorite, i.e., w/v 31.5 g/L). A very fresh-looking submerged colony of *D. geminata* was collected on the 19 April 2007 from the Mararoa River, Southland. The same experiment was carried out with each of the two bleaches in the field.

A fingernail-sized piece of *D. geminata* was clipped with scissors from a colony and put into each of eighteen 50-mL sample tubes. Duplicate tubes were incubated for 1, 5, 10, and 15 mins, and to one tube (the control) no bleach was added. Water was decanted out of the tube through fine mesh to retain the *D. geminata*, and then 2% bleach was added to each tube. After incubation for the appropriate time, the bleach was poured off through a new piece of mesh and the sample was washed twice with bottled water. Finally, 70% ethanol was added to the tubes and they were then tightly capped. This procedure was repeated for the second bleach, and all samples were sent to University of Waikato for analysis.

The DNA was extracted on 23 April 2007, followed by Taqman QPCR using our established protocols described in Section 2.4.

5.3 Results

5.3.1 Net design

The 40 μm net caught six times more cells on average than either the 75 μm or 100 μm mesh sizes (Table 5.1; ANOVA $P < 0.001$). However, cell counts in all samples were made more difficult by large amounts of sediment, and there was markedly more sediment in the 40 μm net sample than in the other mesh sizes.

Table 5.1. Number of *Didymosphenia geminata* cells caught in drift nets set for 2 mins in the Buller River, South Island, New Zealand.

Replicate number	Number of cells/mL of sample
100μm net	
1	1,840
2	1,120
3	1,700
Mean	1,553
75 μm net	
1	1,020
2	1,600
3	1,660
Mean	1,427
40 μm net	
1	9,000
2	10,600
3	7,400
Mean	9,000

5.3.2 Denaturation of DNA from viable *D. geminata* with household bleach

Denaturation Experiment 1: When subjected to Janola bleach for 15 mins, all concentrations from 1.56% to undiluted bleach denatured DNA of *D. geminata*, to the point where it could no longer by amplified by our techniques (Table 5.2). Incubation in 2% bleach for 1 min, however, failed to denature the DNA and we were able to see positive amplifications. The untreated *D. geminata* sample and the PCR positive control sample both produced conclusive PCR signals. The QPCR traces of the real-time amplifications are shown in Figure 5.1.

In order to verify these results, the QPCR products from all the reactions were run out on an agarose gel (Figure 5.2). The gel results concur with Figure 5.1 with only the 2% (1 min) Janola treated sample (lane 8), the untreated control sample (lane 9) and the PCR positive control (lane 10) producing amplicons. In addition, the original DNA extractions from treatments 1-9 were also run out on an agarose gel (Figure 5.3) to determine if any DNA of any size could be visualised. The only extractions that visually yielded any DNA were the 2% Janola 1 min. treatment (lane 8) and the positive control (lane 9).

Table 5.2. Results of QPCR analysis of *Didymosphenia geminata* field samples after incubation with varying concentrations of household Janola bleach (original concentration of sodium hypochlorite 42 g/L at the date of manufacture) for 15 mins.

Lane number	Treatment type	Percent bleach	Sodium hypochlorite (g/L)	OD_{260} DNA (ng/µL)	PCR template (µL)	qPCR signal
1	undiluted bleach	100	4.20	4.3	3	–
2	50% bleach	50	2.10	35	3	–
3	25% bleach	25	1.05	8	3	–
4	12.5% bleach	12.5	0.53	72.5	3	–
5	6.25% bleach	6.25	0.26	1.7	3	–
6	3.13% bleach	3.13	0.13	2.1	3	–
7	1.56% bleach	1.56	0.07	2.1	3	–
8	2% bleach[(a)]	2.00	0.08	11.8	3	+
9	untreated control	0	0	158.9	1	+
10	PCR positive control				1	+
11	PCR negative control				–	–

[(a)] Treatment with 2% bleach for 1 min following the current MAF Biosecurity New Zealand protocol.

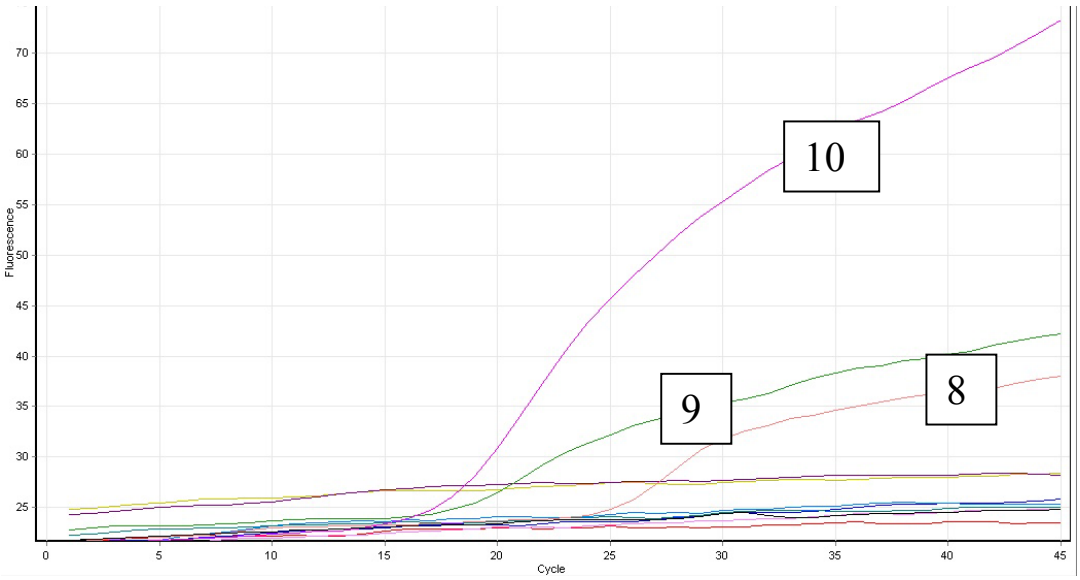


Figure 5.1. Quantitative PCR plot of amplification (fluorescence) against cycle number. Lane 10 - PCR positive control, lane 9 - untreated *Didymosphenia geminata* sample, and lane 8 – sample treated with 2% Janola for 1 min. All remaining samples did not produce amplification signals. Lanes as in Table 5.2.

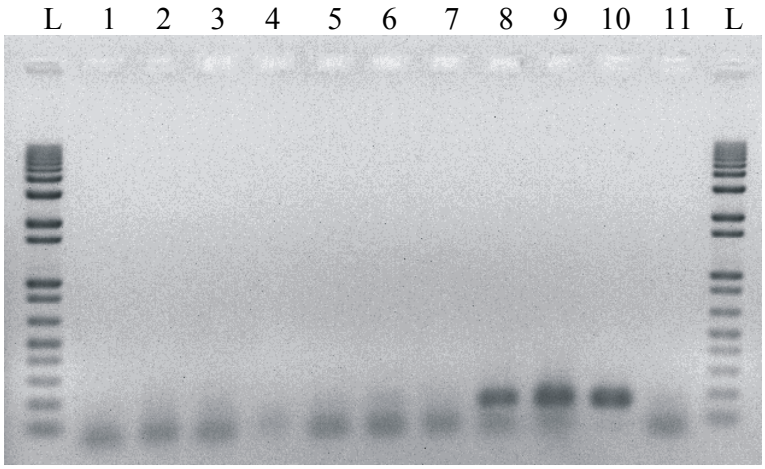


Figure 5.2. Quantitative PCR products electrophoresed through 1.5% agarose. Lanes as numbered in Table 5.2; L= 1 kb+ ladder (Invitrogen).

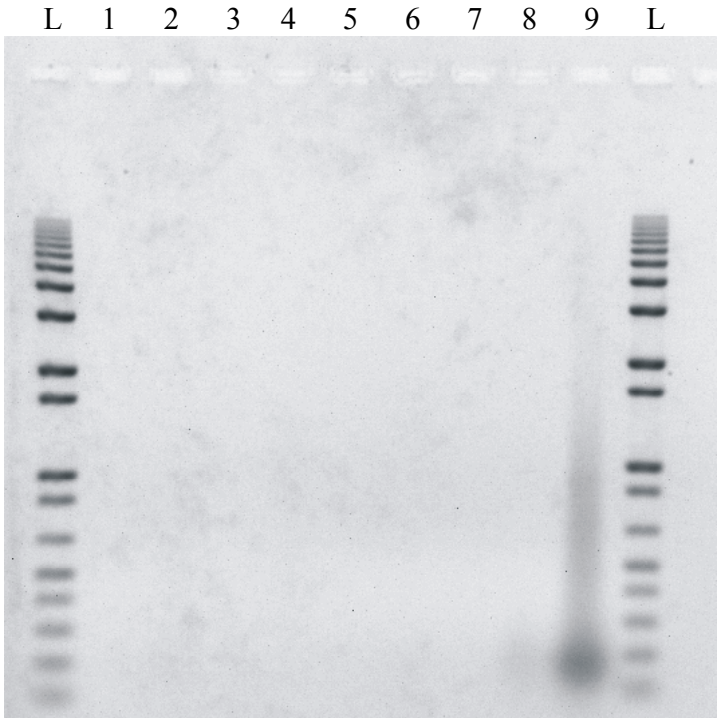


Figure 5.3. Total DNA extracts of *Didymosphenia geminata* treatments electrophoresed through 1.5% agarose. Lanes as numbered in Table 5.2**;** L= 1 kb+ ladder (Invitrogen).

Denaturation Experiment 2: For both bleaches, all Taqman qPCR results were positive for *D. geminata* except for Replicate 1 for 15 mins for Janola that showed a small amplification that did not reach the threshold C_T (Table 5.3). Controls had the highest DNA concentrations, suggesting that some denaturation did occur, but the positive amplifications show that it was incomplete. All high resolution melts (HRMs) were positive for *D. geminata*, further confirming the presence of intact *D. geminata* DNA.

Table 5.3. Results of quantitative PCR analysis of *Didymosphenia geminata* field samples after incubation with 2% Janola bleach (original concentration of sodium hypochlorite 42 g/L at the date of manufacture) and Chlor-O-Gene bleach (original concentration of sodium hypochlorite 31.5 g/L at the date of manufacture) for 1, 5 and 10 and 15 min.

Code	Time (mins)	DNA concentration (ng/μL)	QPCR result	High resolution melt
Janola				
Control	0	10.4	positive	positive
Replicate 1	1	2.9	positive	positive
Replicate 2	1	2.5	positive	positive
Replicate 1	5	0.7	positive	positive
Replicate 2	5	0.2	positive	positive
Replicate 1	10	0.5	positive	positive
Replicate 2	10	0.7	positive	positive
Replicate 1	15	0.5	below threshold	positive
Replicate 2	15	0.1	positive	positive
Generic bleach				
Control	0	116.0	positive	positive
Replicate 1	1	3.9	positive	positive
Replicate 2	1	0.6	positive	positive
Replicate 1	5	1.5	positive	positive
Replicate 2	5	1.5	positive	positive
Replicate 1	10	0.3	positive	positive
Replicate 2	10	1.4	positive	positive
Replicate 1	15	2.2	positive	positive
Replicate 2	15	2.7	positive	positive

5.3.3 Concerns over bleach strength

To investigate the concentration of sodium hypochlorite in a variety of household bleaches available in supermarkets, the manufacturers' labels of five different bleaches were carefully examined. Inconsistencies in the label information between brands was apparent, especially for the crucial "use by" date (Table 5.4). One manufacturer (Janola) reveals that loss of active ingredient occurs with storage. Though the shelf life of bleach is not entirely clear, Janola brand shows that the available chlorine is halved by the "use by" date. Janola purchased on 13 April 2007 had a "use by" date of 18 Nov 2008, suggesting that the shelf life was about 18 months.

All manufacturers gave original concentration of sodium hypochlorite at time of manufacture, but Budget and Chlor-O-Gene brands gave no indication of a "use by" date. The labels show that in addition to 42 g/L sodium hypochlorite (the highest amount of any brand), Janola contains 4 g/L sodium hydroxide, which is also known to be a powerful DNA denaturant.

Table 5.4. Brands of bleach and concentrations of active ingredient investigated in Hamilton supermarkets on 15 April 2007.

Bleach brand	Comment on label	Use by date	Sodium hypochlorite (g/L)	Sodium hydroxide (g/L)
Janola regular premium bleach	Available chlorine 4.0% W/V, 2.0% at use by date	7-Dec-08	42	4
Budget bleach extra strength		None given	36.7	0
Sophora regular	Active ingredient as at date of manufacture	25-Feb-08	25	0
Clor-o-gene		None given	31.5	0
Budget bleach regular		None given	21.5	0

5.4 Conclusions and recommendations

Mesh sizes of 75 and 100 μm caught significantly less cells of *D. geminata*. Despite the capture of sediment and other algae cells by 40 μm, we conclude that coarser meshes cannot be used, and that 40 μm is still the optimum mesh size.

Our survey of bleach brands revealed that we cannot recommend generic bleach brands because of the variation in concentration of active ingredient, and loss of active ingredient with storage. In Denaturation Experiment 1, all *D. geminata* samples treated for 15 min with 1.56 to 100% Janola produced no amplification signals, and thus Janola treatment must have degraded all DNA. The standard MAF Biosecurity New Zealand decontamination protocol, i.e., treatment with 2% Janola for 1 min produced a positive amplification signal indicating that this treatment did not degrade all DNA in the sample.

In Denaturation Experiment 2, in contrast to the first denaturation experiment, none of 2% concentrations of household bleach denatured DNA of *D. geminata*, including the 15-min incubation with 2% Janola. The reason for this difference from the first experiment is not clear, but given the reduction in active ingredient (sodium hypochlorite) with storage, the Janola used in the Experiment 2 may have lost some of its effectiveness during storage. Alternatively, there may have been more *D. geminata* cells in Experiment 2 compared to Experiment 1, or fewer inhibitors. The result demonstrates that 2% bleach is at the limit of effectiveness and cannot be relied upon.

These results indicate that our initial recommendation for the November 2006 delimiting survey (Duncan 2007, Duncan et al., 2007) to use 5% household bleach for 15 min was correct and provides a significant enough safeguard to account for most errors in dilution that might occur in the field and for differences in sodium hypochlorite concentrations in different brands of household bleach. The 15 min exposure to the household bleach is critical for complete DNA removal, considering the extreme sensitivity of our assay for *D. geminata*. Given the loss of active ingredient with storage and after exposure to air (Kilroy et al. 2006), we recommend only Janola brand be used for decontaminating equipment used in the collection of samples for subsequent analysis by Taqman QPCR. We also recommend that the bottles be un-opened at the beginning of the sample collection period, that they be tightly closed between each use, and that the "use by" date be carefully checked. "Expired" bleach according to the "use by" date or bleach which has been opened for an unknown time should not be used.

6 High throughput DNA extraction protocol

6.1 Introduction

Despite obtaining high-quality and PCR-amplifiable DNA from samples using the modified CTAB protocol, it was determined not to be a suitable method for a high-throughput detection system of *D. geminata* due to the protracted nature of the procedure and low throughput capacity. For instance, the maximum number of samples able to be extracted in a single CTAB assay was 24, due to the limitation of the table-top microcentrifuges utilized. It took approximately 6 hours to complete this number of extractions, limiting the number of samples that could be processed within 24 hrs. Additionally, protocol complexity and the potential exposure of the laboratory technician to the toxic chlorinated solvent, chloroform, were additional downfalls of the CTAB protocol.

In order to develop a highly sensitive and specific technique that can analyse large numbers of samples, and as a result, increase the efficiency of the current surveillance programme, we explored other possible DNA extraction methods that would offer a higher throughput capacity without greatly increasing costs. To facilitate this goal, we set out to evaluate commercial DNA extraction kits currently available on the market and compare their relative performance to that of the current CTAB protocol.

6.2 Objective

Our objective was to identify a commercially available DNA extraction kit that would increase sample extraction capacity, reduce extraction time, and prevent exposure to harmful chemicals. It was critical to maintain quality control and quality assurance and the method had to be economically viable.

6.3 Method

Six commercial DNA extraction kits from four different biotechnology companies were compared to our current CTAB protocol. These included four spin-column based extraction kits, and two magnetic-bead based kits. All six kits and the CTAB protocol were employed to extract DNA from duplicate aliquots of three samples from a river known to be positive for *D. geminata*. To effectively evaluate the relative capacities of each DNA extraction method, the three samples ranged in the level and type of contaminants present. Following completion of each method, the extracted DNA was quantified and visualized for integrity, before being amplified by QPCR using *D. geminata*-specific primers. The relative performance of each extraction method on a range of different quantitative and qualitative criteria was then compared in order to identify a particular extraction protocol as the forerunner.

6.3.1 Samples

The following three samples from the Buller River (West Coast, South Island) which were positively identified by microscopy to contain *D. geminata* were used:

1. BU6-75 clean sample – no organic/inorganic contaminants apparent.
2. BU1-40 inorganic sample – visibly contaminated by inorganic compounds, e.g., sand.
3. BU3-100 organic sample – visibly contaminated by organic compounds, e.g., bark.

6.3.2 Extraction protocols

A brief description of each of the protocols utilized to extract DNA from the three Buller River samples is provided below:

1. Current CTAB protocol.
   As previously described (see Appendix 2).
2. Agencourt® Genfind™ v2 Blood and Serum gDNA Isolation Kit (Cat. No. A41499)
   Magnetic bead-based technology.
3. Invitrogen Chargeswitch® Plant gDNA Kit (Cat. No. CS18000)
   Magnetic bead-based technology. Beads provide a switchable surface which is charge dependent on the pH of the surrounding buffer, to facilitate DNA purification.
4. Invitrogen Purelink™ Genomic DNA Kit (Cat. No. K1820-01)
   Spin column technology.
5. Invitrogen Purelink™ Plant Total DNA Purification Kit (Cat. No. K1830-01)
   Spin column technology.
6. Qiagen DNeasy Plant Mini Kit (Cat. No. 69106)
   Spin column technology. Based on the selective binding of DNA to a silica-based membrane in the presence of chaotropic salts.
7. Roche High Pure PCR Template Preparation Kit (Cat. No. 11796828001)
   Spin column technology.

6.3.3 Experimental Design

From each of the original three ethanol-stored samples, fourteen 1-mL aliquots were transferred into microcentrifuge tubes. Two aliquots from each sample (six in total) were extracted with each of the different extraction protocols. An extraction NTC was included with each assay in order to test for reagent contamination.

Three of the commercial DNA extraction kits did not include a Proteinase K treatment. However, our earlier observation of its importance in increasing the total DNA yield and integrity (see Section 2.4.2.7) justified our decision to add 30 µL of Proteinase K (10 mg/mL) (Roche) to these protocols. All Proteinase K treatments were standardized across the protocols to 1 hour incubations. Apart from these adjustments each extraction followed the manufacturer's protocol without modification.

The internal reference pGEM was added (14 ng per 1 mL aliquot) to each extracted sample, for all extraction protocols. By adding the same amount of pGEM to each sample, the relative extraction efficiency could be determined by its subsequent detection by QPCR. For example, a protocol with a relatively low extraction efficiency would thus reduce the absolute amount of pGEM in the sample as detected by QPCR.

Following the completion of each protocol, the extracted DNA was quantified by spectrophotometry, and approximately 150 ng analysed by agarose gel electrophoresis, ethidium bromide staining, and UV-transillumination.

6.4 Results

For each 1-mL aliquot extracted, 2 µL was used to determine the nucleic acid concentration with the Nanodrop spectrophotometer. These readings were presented in ng/µL, so to estimate the total nucleic acid yield, the concentration value obtained was multiplied by the volume of the buffer used to elute and resuspend the extracted DNA (Table 6.1). For example, using the CTAB protocol, clean sample (a) produced a concentration of 20.46 ng/µL when resuspended in 20 µL of LoTE. Therefore, the total nucleic acid yield for that sample was 0.409 µg (20.46 ng/µL x 20 µL).

The Roche kit produced the highest nucleic acid yield for the clean sample (Table 6.1). For the inorganic sample, the Invitrogen Chargeswitch and Purelink Genomic performed the best. In addition, all three of these protocols also produced the highest total yield for the organic sample. The replicate aliquots of the organic sample gave unreliable nucleic acid readings when extracted by the CTAB protocol.

Table 6.1. Total nucleic acid yield obtained by the different DNA extraction methods, as determined by spectrophotometer readings.

DNA extraction protocol	Total nucleic acid yield (µg)					
	Clean sample		Inorganic sample		Organic sample	
	(a)	(a)	(a)	(b)	(a)	(b)
1. CTAB	0.409	0.429	3.253	3.387	2.310	10.052
2. Agencourt® Genfind v2 - Blood and serum gDNA Kit	1.5	1.577	1.245	1.321	1.843	1.605
3. Invitrogen - Chargeswitch® Plant gDNA Kit	0.818	0.673	7.091	6.669	3.091	3.148
4. Invitrogen - Purelink™ Genomic DNA Kit	0.609	0.402	4.112	5.028	3.209	1.913
5. Invitrogen - Purelink ™ Plant DNA Kit	0.053	0.118	1.203	0.944	0.296	0.255
6. Qiagen - DNeasy Plant Mini Kit	0.159	0.399	1.685	1.896	1.887	2.393
7. Roche - High Pure PCR Template Preparation Kit	3.178	2.378	2.9	2.98	3.548	2.678

6.4.1 Agarose gel electrophoresis

Using the Nanodrop readings as a concentration guide, approximately 150 ng of extracted DNA was loaded into a 1.5% agarose gel with sodium boric acid (SBA) buffer. After electrophoresis at 100 V for 20 mins, the DNA was then stained with ethidium bromide and visualised by ultra-violet transillumination (Figure 6.1). This method of analysis provides information on the integrity of the genomic DNA (gDNA) extracted. In gel pictures of all seven extraction methods, the thick single bands at a high molecular weight indicated the presence of gDNA. However, the presence of degraded DNA is apparent in some samples as indicated by the smears. This was particularly obvious in the CTAB-extracted organic sample (b), e.g., CTAB Lane 4 in Figure 6.1, which corresponded with a very high total yield of 10.05 µg.

The gDNA bands produced by the samples extracted with the Invitrogen Chargeswitch Kit were significantly reduced in intensity as would have been expected according to the Nanodrop readings (Table 6.1). This may have been due to residual magnetic beads in the extraction that then absorbed at a similar wavelength when analyzed by spectrophotometry, resulting in the overestimation of the nucleic acid concentration.

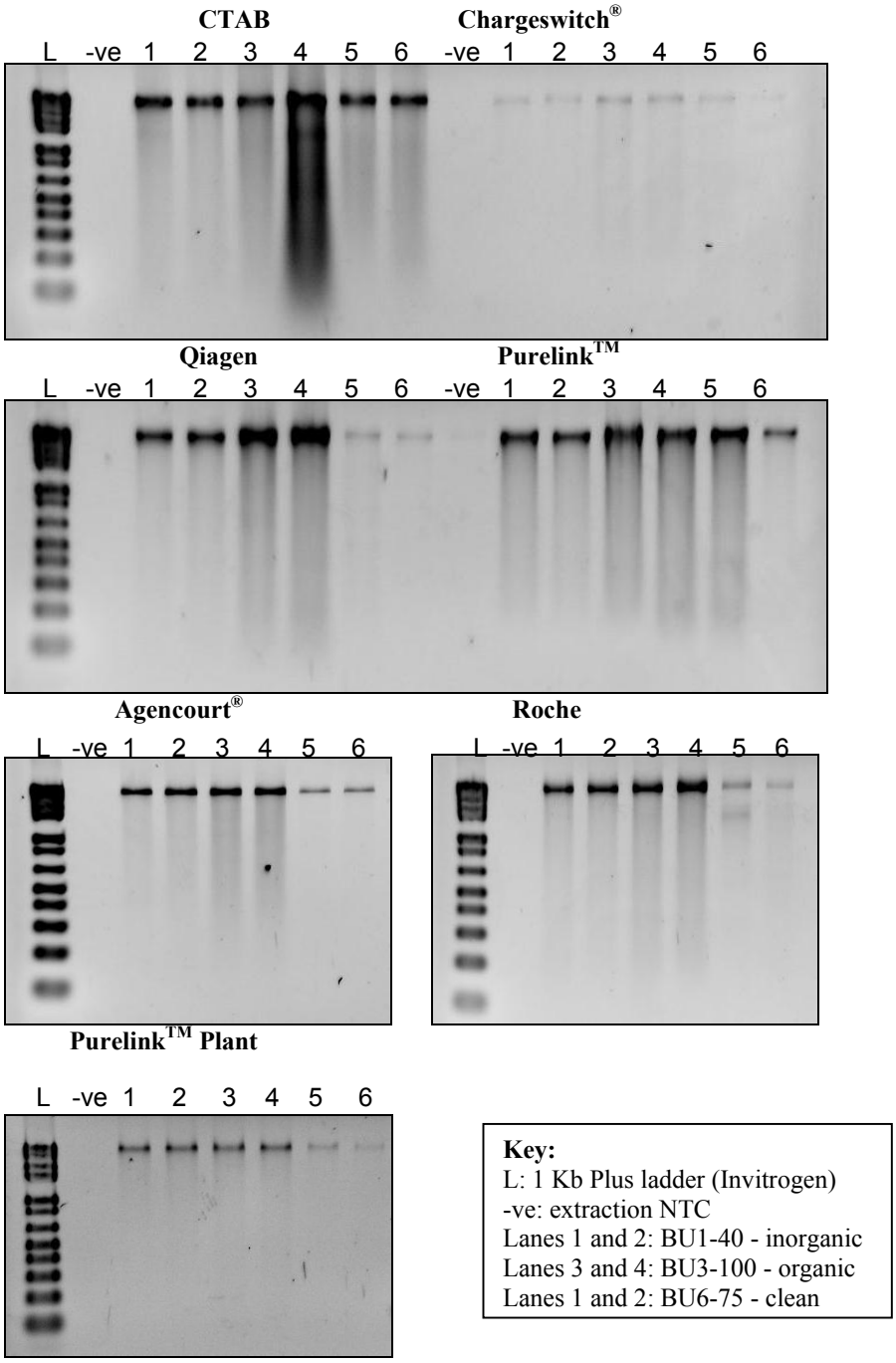


Figure 6.1. Agarose gel (1.5%) showing the intact gDNA extracted from the Buller River samples with seven different extraction methods.

6.4.2 QPCR

Extracted DNA was then diluted to 10 ng/μL and used in duplicate reactions for Taqman QPCR (see section 2.4), with the *D. geminata*-specific D602F and D753R primers, to evaluate the relative amplification success of each of the extraction methods for all three samples. Primers for amplifying pGEM (M13F and pGEMR) were also used to determine the extraction efficiency of each method. All reaction and thermal cycling parameters were as previously described.

Amplification success (above 1 pg threshold) with both D602F-D753R and M13F-pGEMR was identified (Table 6.2).

Table 6.2. QPCR amplification of the DNA extracted by different DNA extraction protocols.

DNA extraction protocol	Amplification of 20 ng of extracted DNA					
	Clean sample		Inorganic sample		Organic sample	
	C_T > threshold	pGEM	C_T > threshold	pGEM	C_T > threshold	pGEM
CTAB	Yes	Yes	No	Variable	Yes	Yes
Agencourt® Genfind v2 - Blood and serum gDNA Kit	Yes	Variable	Yes	Yes	Yes	Yes
Invitrogen - Chargeswitch® Plant gDNA Kit	Yes	Variable	No	No	No	No
Invitrogen - Purelink ™ Genomic DNA Kit	Yes	Yes	Yes	Variable	Yes	Yes
Invitrogen - Purelink ™ Plant DNA Kit	Yes	Yes	Yes	Variable	Yes	Variable
Qiagen - DNeasy Plant Mini Kit	Yes	Yes	Yes	Variable	Yes	Yes
Roche - High Pure PCR Template Preparation Kit	Yes	Variable	Yes	Yes	Yes	Yes

All of the samples extracted by Agencourt, Invitrogen Purelink™, Purelink™ Plant, Qiagen, and Roche kits were successfully amplified with D602F and D753R. However, both the CTAB and Invitrogen Chargeswitch protocols demonstrated variable success. The amplification of pGEM was successful for at least 4 of the 6 samples for all the extraction protocols, with the exception of the Invitrogen Chargeswitch Kit.

Following completion of Taqman QPCR analysis, all DNA extraction protocols were evaluated for a number of different criteria, including time, cost, yield, and QPCR amplification (Table 6.3). A basic scoring system was used for each criterion, where protocols were marked out of 5 or 10, depending on their relative success.

Table 6.3. Evaluation of the seven different DNA extraction methods.

Selection Criteria	Standard CTAB protocol	Score	Agencourt ®Genfind v2 Blood and serum gDNA kit	Score	Invitrogen Chargeswitch ®Plant gDNA Kit	Score	Invitrogen Purelink TM Genomic DNA Kit	Score	Invitrogen Purelink TM Plant DNA Kit	Score	Qiagen DNeasy Plant Mini Kit	Score	Roche High Pure PCR Template Preparation Kit	Score
96-sample capacity	No	0	Yes	10	Yes	10	Yes	10	No	10	Yes	10	No	0
Total assay time (for 6 reactions)	6.5 hours	0	4.0 hours	4	3.0 hours	8	2.5 hours	10	3.0 hours	8	3.0 hours	8	2.5 hours	10
Cost per reaction	<$2.00	10	$7.68	0	$3.13	9	$2.48	10	$5.15	5	$6.93	1	$3.33	8
Cost per reaction (high-throughput)	N/A	0	$3.83	7	$3.10	9	$2.90	9	N/A	0	$5.05	5	N/A	0
Proteinase K step included	Yes	5	Yes	5	No	0	Yes	5	No	0	No	0	Yes	5
Proteinase K incubation time	1 hour	5	1 hour	5	1 hour	5	1 hour	5	1 hour	5	1 hour	5	1 hour	5
Relative ease of protocol	Laborious	0	Moderate	5	Easy	7	Very easy	10	Very easy	10	Easy	8	Very easy	9
Extra reagents and equipment needed	CTAB, $CHCl_3$, IAA, IPA, NaCl, LoTE, 100% EtOH heat block, 1.7 mL tubes, table-top centrifuge	-	100% EtOH, TE, heat block, 2.0 mL tubes	-	100% EtOH, heat block, 1.7 mL tubes, magnetic rack	-	100% EtOH, heat block, 1.7 mL tubes, table-top centrifuge	-	100% EtOH, 1.7 mL tubes, heat block, table-top centrifuge	-	100% EtOH, heat block, 1.7 mL tubes, table-top centrifuge	-	100% EtOH, IPA, 1.7 mL tubes,heat block, table-top centrifuge	-
Comments	Exposure to toxic reagents	-10	Residual magnetic beads, poor design of magnetic rack	-5	Residual magnetic beads	-10		-	Kit was stored at RT for approx 1 week	-10	2 x spin columns	-		-
Total nucleic acid yield (µg) from clean sample	0.41-0.43	3	1.5-1.58	8	0.67-0.82	6	0.4-0.61	4	0.05-0.1	0	0.16-0.4	2	2.38-3.18	10
Total nucleic acid yield (µg) from inorganic sample	3.25-3.39	6	1.25-1.32	3	6.67-7.09	10	4.11-5.03	8	0.94-1.2	1	1.69-1.89	4	2.9-2.98	5
Total nucleic acid yield (µg) from organic sample	2.31-10.05	6	1.61-1.84	5	3.09-3.15	10	1.91-3.21	8	0.26-0.3	0	1.89-2.39	7	2.68-3.55	9
Volume of elution buffer	20 µL	-	100 µL	-	100 µL	-	50 µL	-	100 µL	-	100 µL	-	200 µL	-
DNA Purity (260:280) of clean sample	1.84-2.03	9	1.52-2.02	7	1.34-1.54	4	1.85-1.93	9	-	0	1.55-2.18	7	1.27-1.35	3
DNA Purity (260:280) of inorganic sample	1.48-1.52	5	1.45-1.55	5	1.04-1.32	2	1.37-1.39	4	1.26-1.29	3	1.43-1.48	4	1.27-1.3	3
DNA Purity (260:280) of organic sample	1.66-1.92	8	1.52-1.6	5	1.35	4	1.53-1.56	5	1.18-1.28	2	1.54-1.66	6	1.27-1.39	3
D602F-D753R QPCR of all 3 samples	~2/3	7	all	10	~1/3	3	all	10	all	10	all	10	all	10
M13F-pGEMR QPCR of all 3 samples	~2.5/3	8	~2.5/3	8	~0.5/3	2	~2.5/3	8	~2/3	7	~2.5/3	8	~2.5/3	8
Total Score	**52**		**75**		**61**		**96**		**46**		**79**		**80**	

The cumulative score derived for each extraction protocol then facilitated the identification of the top three extraction methods:

1st	Invitrogen Purelink ™ Genomic DNA Kit	– score 96
2nd	Roche High Pure PCR Template Preparation Kit	– score 80
3rd	Qiagen DNeasy Plant Mini Kit	– score 79

The success of the Invitrogen Purelink™ Genomic DNA Kit was largely attributed to the relatively short assay time, easy protocol, QPCR success, and low cost. The Roche High Pure PCR Template Preparation Kit almost had equal success, although the inability to convert to a high throughput protocol was a major disadvantage. The relatively high cost per extraction contributed to the lower placement of Qiagen's DNeasy Plant Mini Kit. The presence of residual magnetic beads reduced the success of the two magnetic bead-based protocols.

6.5 Conclusions and Recommendations

Using the basic scoring system described above, it was determined that the extraction method that produced the best yield, quickest assay time, highest QPCR amplification, and was most cost effective was the Invitrogen Purelink™ Genomic DNA kit. This kit is available in a 96-extraction format, which would potentially increase the current daily maximum samples able to be extracted from 24 to 192. Thus, the employment of such an extraction method for the current project would increase the sample turnaround time, maintain quality control and assurance, and reduce the costs associated with each extraction and decrease the amount of technician time required.

7 Technology transfer

7.1 Introduction

Over the past eighteen months, we have successfully developed an efficient and cost-effective molecular-based protocol capable of detecting *D. geminata* in environmental samples with extreme sensitivity. If employed as standard procedure, this highly sensitive and specific technique could potentially increase our ability to screen large numbers of samples in a short period of time and to detect the alga at very low levels. This would increase the efficiency of the *D. geminata* surveillance programme, without greatly affecting the costs. However, in order to establish the new procedure, we understand the importance of transferring this information and technology to all those who are directly or indirectly involved in *D. geminata* surveillance efforts in the laboratory and field in New Zealand. To facilitate this technology transfer, we have designed a 1 day workshop that will train invited end users from various government and regional organisations on all aspects of the DNA detection system.

7.2 Objective and workshop details

Our objective is to host the workshop at the request of MAF Biosecurity New Zealand and/or Didymo Long-Term Management Partners (local government councils, Fish and Game councils, Department of Conservation, industry and iwi) involved in surveillance activities for *D. geminata.* We will provide an extensive manual in order to train and educate end users and organisations on the sampling procedures required for use with the QPCR-based procedure we have developed for the detection of *D. geminata*. The agenda of these workshops will focus primarily on the sampling procedure (including net design, net deployment, decontamination, DNA denaturation, quality control, sample preparation, sample collection, sample packaging, and sample transport).

7.3 Workshop Manual

A specifically designed sampling procedure manual will be provided to all workshop attendees. This operational manual will contain comprehensive step-by-step guides for the sampling procedures.

7.4 Conclusions and recommendations

Upon completion of the workshop, attendees will be knowledgeable on the sampling protocols and technologies relevant to the highly-sensitive DNA detection system designed to improve the efficiency, cost-effectiveness, and sensitivity of current *D. geminata* surveillance efforts.

8 Rapid Response Protocol, Risk Assessment, Quality Control and Quality Assurance

8.1 Rapid Response Protocol

The power of the Taqman QPCR method lies in its negative predictive value. The sensitivity and specificity of the method indicates that any negative result means that the water body is highly likely free of *D. geminata*. However, continued surveillance with the method in the North Island will inevitably produce suspect weak positive results (at or below the BT) that will require further verification and validation. In most cases, we believe that the built-in quality control and quality assurance (QC/QA) protocols (see Section 8.2 below) will provide the necessary information to immediately verify the results as being a true positive or a false positive. However, there will be cases where the results remain ambiguous needing further validation.

We propose the following rapid response to any sample that shows a positive or weak amplification at or below the BT from areas previously thought to be free of *D. geminata*.

1. Sample DNA will be immediately re-run (QPCR, gel, HRM) following the normal prescribed analysis pipeline (24 hours). If the results remain suspect, proceed to step 2.

2. Original samples will be re-extracted and run through the normal sample pipeline (QPCR, gel, HRM) again (48 hours). If the results remain suspect, proceed to step 3.

3. The suspect DNA will be amplified for the full *D. geminata* 18S rRNA and the resulting amplicon sequenced. The resulting sequence will be compared with our NZ diatom 18S rDNA alignment database (48 hours). If the resulting sequence aligns with *D. geminata*, proceed to step 4.

4. Request an immediate re-sampling of the original site. This would entail multiple sampling at the original site as well as upstream and downstream using the drift net technique. We will also request that benthic swab samples, as described by Cary *et al.,* (2006), be taken at each of these locations. We would also preserve drift net samples for immediate microscopy. Collectively, these new samples will be run through the normal QPCR sample analysis pipeline (48 hours). If these were to present positive or weak positive results, we would then notify MAF Biosecurity New Zealand and the appropriate regional authorities.

8.2 Risk Assessment, Quality Control and Quality Assurance

Although every effort has been made to assure the accuracy and sensitivity of the DNA method, it remains, and will remain, fallible. Though we believe that it overcomes some limitations of microscopy at low levels of *D. geminata* abundance, it is important to recognise the potential sources for failure of the DNA method, and so we have summarised these risks in Table 8.1. In addition to the possibility that in the wild there exists an unknown "*D. geminata*-like" organism (see Section 3. Validating the DNA method), there are several ways through the sample collection and preparation or in the analytical process itself where "contamination" can be introduced that can result in either false positives or negatives. We have developed a battery of

controls and the use of an internal reference standard that are integrated into the method to alert us if any of these are affecting a given sample.

Table 8.1 lists known and hypothesised entry points for contamination and inhibitors in the sample collection and analysis pipeline, the effect of these contaminants on the QPCR, and the mechanisms used to detect and control for these risks. In each case, the controls are performed at least in duplicate on each sample or sample series. The result is that there are 4 QPCR reactions run for every sample (two for *D. geminata* and two for the pGEM control) and an additional 12 controls (positive and negative reagent and template controls) for each experiment. Because the DNA method is so sensitive, we must maintain this high level of control to assure QC/QA. We believe that the sample procedures and processing pipeline developed for *D. geminata* are exceptionally robust providing the highest level of QC/QA possible. Below we address several specific areas of control and concern in more detail. We must reiterate here that the DNA method was designed as a very sensitive tool for surveillance to complement the current microscopy method by greatly extending our range of detection.

8.2.1 Redundancy in sample archiving

When each sample is received into the laboratory, it is given its own reference number, logged into the database and stored at -20°C until processing. When processing commences, the original sample is first split and a portion (~2 mL) immediately archived at -20°C. An additional sample of the initial sample digestion and lysis solution (with the pGEM standard) is also archived and stored at -20°C. The inclusion of the lysate in the archive means that the sample can easily be re-extracted quickly, with a minimum of effort. When the final DNA is obtained and quantified, it is also archived at -20°C (see Figure 2.1). If at any time the quality/result of any sample is questioned, we can return to the original process stream and resample if needed. This 3-fold archiving protocol builds in another level of QC/QA insurance to the procedure.

8.2.2 Environmental contamination and the pGEM standard

Environmental samples often pose a sizable challenge to molecular based techniques in that common contaminants can cause problems with the enzymes and reactants used in the PCR process. Aquatic samples are notorious for containing plant-derived inhibitors (humics, tannins, etc.) that will co-purify with many DNA extraction protocols. The composition and diversity of these compounds are unknown but the effects have been well documented. Many of these compounds will inhibit or diminish the efficiency of the DNA polymerase used in the PCR reaction resulting in a false negative or a reduced target signal. We have included an internal reference standard (pGEM) with our CTAB lysis buffer that we allow to co-purify with the environmental DNA during the extract process. In a previous study (Coyne *et al.,* 2005) we developed a Taqman QPCR assay specifically for the pGEM standard. The pGEM QPCR is run side by side with the *D. geminata* QPCR on each sample. If the amount of pGEM added to the sample is known at the onset of the extraction process, by measuring the pGEM after extraction we can deduce the overall extraction efficiency and presence of any inhibitors. Analysis of the pGEM standard QPCR provides not only an indication of the extraction efficiency in that particular sample, but also a sensitive indicator of possible environmental inhibitors. In the event the pGEM fails or is dramatically reduced, we will assume inhibitors are present and rerun the

QPCR with additives that enhance the reaction in the presence of inhibitors or re-extract the sample using alternative, more time-consuming procedures.

8.2.3 Uncontrollable sources of contamination.

Given that we are working with environmental samples, multiple ways exist to contaminate waterways and our samples that are beyond our control. While such contamination can be prevented in most cases, the sensitivity of the DNA method requires that we recognise such contamination is possible and conduct our sampling in a way to minimize its effect.

The DNA method detects the presence of *D. geminata* DNA down to levels equivalent to a single cell per mL of sample – which could be representative of thousands of litres of water when using the drift nets. The technique does not differentiate between live or recently dead *D. geminata*. A probable source of error for the DNA method is that sampling crews might enter the river upstream of the drift net sampling and introduce higher concentrations of benthic inhibitors, or worse, contaminant dead *D. geminata* cells from previous sites. It is also conceivable, however extremely unlikely, that dead cells brought into a river by a recreational user or water fowl might be picked up down stream and eventually be detected as positive by the DNA method. We must emphasise the word "unlikely" because of the nature of how the drift net is deployed, what it actually samples, and the dilution factors of the river itself, which mean the chances of this happening are very slim. In any case, if a high risk area were to show up positive by the DNA method, our procedure would be to alert MAF Biosecurity New Zealand and request that our end users resample the river in the same place immediately.

Table 8.1. Summary of the sources of risk for the DNA method.

Point of contamination	Origin of contamination	Effect on process	Probability of risk	Control (s) implemented
Sample collection				
	Dead *D. geminata*	False positive	Extremely low	Dead cells introduced by wildfowl, upstream samplers, or recreational users. Make sure that net sampling takes place upstream of all additional sampling practices. Resample river at a later time.
	Drift net and associated gear	False positive	Low	DNA denaturation treatment for nets developed and tested - removes all ampifiable traces of *D. geminata* DNA (Section 5).
	Field personnel	False positive	Medium	Comprehensive training and instruction for field personnel, inclusion of disposable plasticware for each sample, decontamination procedures.
	Laboratory	False positive	Low	Each sample logged on arrival, sub-sampled (3X) at each step in extraction process and archived.
DNA Extraction				
	Extraction supplies	False positive	Extremely low	Each sample isolated from others, use all new sterile plasticware with barriers
	Extraction reagents	False positive	Low	Included 2 extraction reagent “no-sample” controls for each set of extractions
QPCR				
	QPCR reagents	False Positive	Low	Included 3 no-template controls for each QPCR experiment performed
	QPCR instrument	False Positive	Low	Wiped out before each experiment, QPCR negative controls
	Environmental inhibitors	False negative	Medium	Include 2 pGEM internal standard controls for each sample run.

9 Overall conclusions and recommendations

9.1 Summary of conclusions

We have reported the development and validation of a highly sensitive and specific genetic detection method for *D. geminata.* This method has been developed specifically to complement current routine surveillance efforts in New Zealand. A more cost-effective SYBR Green I assay was investigated and found capable of detecting *D. geminata* over eight orders of magnitude, making it an ideal method for *en masse* sampling and routine monitoring of rivers already known to contain *D. geminata*, should this be warranted for future ecological or control studies. The more sensitive Taqman method was found capable of detecting less than a single cell per mL of an ethanol-preserved drift net sample and is the method of choice for surveillance of *D. geminata* in areas where it has not been previously detected.

The Taqman assay was stringently validated using over 100 environmental samples and found to be specific to *D. geminata*. The assay continues to be tested and updated as new rivers are examined. The Taqman method also appears capable of detecting and enumerating all international samples collected to date with the same specificity and sensitivity. In order to fully embrace the sensitivity of the Taqman assay, we were required to develop new sampling protocols for the field and laboratory that assure that *D. geminata* DNA contamination is not introduced anywhere in the sample collection, preparation, and analysis pipeline. This required the design of a new drift net, protocols to decontaminate sampling gear, and the development of strict QC/QA procedures to assure the samples remain untainted until fully analysed. The analysis pipeline includes a 3-fold redundancy in archiving to assure that in the event any result is questioned, we can quickly return to that sample for a second round of analysis.

We also reported our preliminary work on the phylogeography of *D. geminata*. Our current analysis of population level genetic markers has shown that the ITS region of the 18S rRNA gene has the variability necessary to differentiate at the population level. Preliminary analysis of the ITS revealed that, after limited sampling, one New Zealand strain is more closely related to the Montana, USA sample compared to any of the other international samples in our collection, suggesting that the original introduction may have come from the USA. However, much more work is needed to strengthen this preliminary conclusion and to improve the capability and power of the phylogeographic analysis.

9.2 Recommendations and the future

After 18 months of research and development, we believe the DNA detection method has been adequately tested and validated and is now ready for full implementation. The method, while extremely robust, will continue to be tested and assessed to maintain the highest level of quality control and quality assurance possible. Because many high-risk waterways in New Zealand have yet to be tested with the DNA method, we plan to continue to fully validate (via electrophoresis gel, HRM, sequence) any QPCR positive sample we obtain. As new sequences from more closely related diatoms are obtained by us or through international databases, we plan to reassess our current method, and adapt it if necessary.

We recommend that the Taqman QPCR DNA detection method become the primary surveillance tool for the North Island of New Zealand, where *D. geminata* has yet to be detected. The power of the Taqman QPCR method lies in its negative predictive value. The sensitivity and specificity of the method means that any negative result indicates that the water body is highly likely to be free of *D. geminata*. Also, any positive results with the DNA detection method can be validated with other molecular methods (electrophoresis gel,

HRM, sequence) or by the microscopic method, to rule out contamination and provide high confidence in the extreme low level detection capability.

Detecting *D. geminata* as early as possible, should it spread to the North Island, will increase the chance of a successful elimination attempt, should a control tool be developed. Early detection will also enable containment measures to achieve their maximum effectiveness by facilitating early communication of the positive status of a river. Having the knowledge that a river is affected before the organism is visible should enable MAF Biosecurity New Zealand and their *D. geminata* long-term management partners to target their social marketing campaigns towards freshwater users who are especially at risk of spreading the organism, and whose actions to slow the spread will reduce impacts to other river systems.

10 Acknowledgements

We especially thank Susie Wood of Cawthron Institute, Nelson, for her independent review and assistance with cell counts. We also thank Emily Atkinson, DOC, Southland, for running the denaturation experiments in the field. Eric Edwards of DOC, Southland, also provided critical support. We thank Cathy Kilroy and Maurice Duncan of NIWA, Christchurch, for their unfailing cooperation and sample provision. We owe a major debt of gratitude to Christina Vieglais (MAF Biosecurity New Zealand; MAFBNZ) for her unending support and enthusiasm for this work, and Frances Velvin (MAFBNZ) for her guidance in the final stages of report production. MAFBNZ funded the project. DNA sequencing was carried out by the Waikato DNA Sequencing Unit, Department of Biological Sciences, University of Waikato. We also thank Naomi Crawford for assisting at various stages of the project when more hands were needed.

11 References

Antoine, S.E. and Benson-Evans, K. 1984. Morphological variation in six populations of *Didymosphenia geminata* (Lyngbye) M. Schmidt of Great Britain. Nova Hedwigia 40: 141-156.

Behnke, A., T. Friedl, V. A. Chepurnov, and D. G. Mann. 2004. Reproductive compatibility and rDNA sequence analyses in the *Sellaphora pupula* species complex. (Bacillariophyceae). Journal of Phycology 40: 193-208.

Beszteri, B., E. Acs, J. Makk, G. Kovács, K. Márialigeti, and K. T. Kiss. 2001. Phylogeny of six naviculoid diatoms based on 18S rDNA sequences. International Journal of Systematic and Evolutionary Microbiology 51: 1581-1586.

Bornet, B. E. Antoine, S. Françoise, and C. Marcaillou-Le Baut. 2005. Development of sequence characterized amplified region markers form intersimple sequence repeat fingerprints for the molecular detection of toxic phytoplankton *Alexandrium catenella* (Dinophyceae) and *Pseudo-nitzschia pseduodelicatissima* (Bacillariophyceae) from French coastal waters. Journal of Phycology 41: 704-711.

Bowers, H. A., T. Tengs, H. B. J. Glasgow, J. M. Burkholder, P. A. Rublee and D. W. Oldach. 2000. Development of real-time PCR assays for rapid detection of *Pfiesteria piscicida* and related dinoflagellates. Applied and Environmental Microbiology 66:4641-4648.

Cary, S. C., B. J. Hicks, N. J. Crawford and K. Coyne. 2006. A sensitive genetic-based detection capability for *Didymosphenia geminata*. Interim Report. CBER contract Report 45. Prepared for MAF Biosecurity New Zealand. 69 p.

Coyne, K. J., Handy, S. M., Demir, E., Whereat, E. B., Hutchins, D. A., Portune, K. J., Doblin, M. A. and Cary, S. C. 2005. Improved quantitative real-time PCR assays for enumeration of harmful algal species in field samples using an exogenous DNA reference standard. Limnology and Oceanography Methods 3: 381-391.

Créach, V., A. Ernst, K. Sabbe, B. Vanelslander, W. Vyverman, and L.J. Stal. 2006. Using quantitiative PCR to determine the distribution of a semicryptic benthic diatom, *Nivicula phyllepta* (Bacillariophyceae). Journal of Phycology 42: 1142-1154.

Demir, E., Coyne, K. J., Doblin, M. A., Handy, S. M. and Hutchins, D. A. Microzooplankton grazing rates on *Heterosigma akashiwo* in the Delaware Inland Bays: A species-specific approach combining quantitative real-time PCR (QPCR) and dilution methods. In Press, Microbial Ecology.

Duncan, M.J. 2007. New Zealand-wide surveys in November 2006, February 2007 and May 2007 for the presence of the non-indigenous freshwater diatom *Didymosphenia geminata*

in high sites. NIWA Client Report CHC2006-053. For MAF Biosecurity New Zealand. 129 p.

Duncan, M.J., Kilroy, C., and Vieglais, C.M.C. 2007. Protocol for the collection of samples for delimiting surveys for *Didymosphenia geminata* for microscopic and DNA analysis. NIWA Client Report CHC2007-102. Prepared for MAF Biosecurity New Zealand. 22 p.

Ehara, M., Y. Inagaki, K. I. Watanabe, and T. Ohama. 2000. Phylogenetic analysis of diatom *coxI* genes and implications of a fluctuating GC content on mitochondrial genetic code evolution. Current Genetics 37: 29-33.

Godhe, A., M. R. McQuoid, I. Karunasagar, I. Karunasagar, and A.-S. Rehnstam-Holm. 2006. Comparison of three common molecular tools for distinguishing among geographically separated clones of the diatom *Skeletonema marinoi* Sarno et Zingone (Bacillariophyceae). Journal of Phycology 42: 280-291.

Handy, S. M., Cary, S. C., Hutchins, D. A., and Coyne, K. J. 2006. Simultaneous enumeration of multiple raphidophyte species by quantitative real-time PCR: capabilities and limitations. Limnology and Oceanography Methods 3:193-204.

Handy, S. M., Coyne, K. J., Portune, K. J., Demir, E., Doblin, M. A., Hare, C. E., Cary S. C., and Hutchins, D. A. 2005. Evaluating vertical migration behavior of harmful raphidophytes in the Delaware Inland Bays utilizing quantitative real-time PCR. Aquatic Microbial Ecology 40:121-132.

Heid, C. A., J. Stevens, K. J. Livak, and P. M. Williams. 1996. Real time quantitative PCR. Genome Research 6: 986-994.

Kilroy, C. 2004. A new alien diatom, *Didymosphenia geminata* (Lyngbye) Schmidt: its biology, distribution, effects and potential risks for New Zealand fresh waters. NIWA Client Report CHC2004-128. For MAF Biosecurity New Zealand. 34 p.

Kilroy, C. and Dale, M. 2006. A comparison of sampling methods for the detection of the invasive alga *Didymosphenia geminata* in New Zealand. NIWA Client Report CHC2006-078. For MAF Biosecurity New Zealand. 48 p.

Kilroy, C., Lagerstedt, A., Davey, A and Robertson, K. 2006. Studies on the survivability of the invasive diatom *Didymosphenia geminata* under a range of environmental and chemical conditions. NIWA Client Report CHC2006-116. For MAF Biosecurity New Zealand. 121 p (Revised May 2007).

Klein, D. 2002. Quantification using real-time PCR technology: applications and limitations. Trends in Molecular Medicine 8: 257-260.

Kooistra, W.H.F.C. 1992. Biogeography of *Cladophopsis membranacea* (Chlorophyta) based on comparisons of nuclear rDNA ITS sequences. Journal of Phycology 28: 660-668.

Lasken, R. S. and Egholm, M. 2003. Whole genome amplification: abundant supplies of DNA from precious samples or clinical specimens. Trends in Biotechnology 21:531-535.

Livak, K. J. and Schmittgen, T. D. 2001. Analysis of relative gene expression data using RT PCR and the 2-ddCt method. Methods 25: 402-408.

Lovatt, A. 2002. Applications of quantitative PCR in the biosafety and genetic stability assessment of biotechnology products. Reviews in Molecular Biotechnology 82: 279-300.

Metzeltin, D. and H. Lange-Bertaolt. 1995. Kritische wertung der taxa in *Didymosphenia* (Bacillariophyceae). Nova Hedwigia 60: 381-405.

Popels, L. C., S. C. Cary, D. A. Hutchins, R. Forbes, F. Pustizzi, C. J. Gobler, and K. J. Coyne. 2003. The use of quantitative polymerase chain reaction for the detection and enumeration of the harmful alga Aureococcus anophagefferens in environmental samples along the United States East Coast. Limnology and Oceanography Methods 1: 92-102.

Provan, J., S. Murphy, and C. A. Maggs. 2004.Universal plastid primers for Chlorophyta and Rhodophyta. European Journal of Phycology 39:43-50.

Provan, J., S. Murphy, and C. A. Maggs. 2005. Tracking the invasive history of the green alga *Codium fragile* spp. *tomentosides*. Molecular Ecology 14: 189-194.

Ririe,K.M., Rasmussen,R.P. and Wittwer,C.T. 1997. Product differentiation by analysis of DNA melting curves during the polymerase chain reaction. Analytical Biochemistry 245: 154-160.

Rynearson, T. A. and E. V. Armbrust. 2004. Genetic differentiation among populations of the planktonic marine diatom *Ditylum brightwellii* (Bacillariophyceae). Journal of Phycology 40: 34-43.

Saito, K., T. Drgon, J. Robledo, D. Krupatkina and G. Vasta. 2002. Characterization of the rRNA locus of *Pfiesteria piscicida* and development of standard and quantitative PCR-based detection. Applied Environmental Microbiology 68: 5394-5407.

Stevens, M.I., S. A. Hunger, S. F. Hills, C. E. C. Gemmill. 2007. Phantom hitch-hikers mislead estimates of genetic variation in Antarctic mosses. Plant Systematics and Evolution 263: 191–201.

Tebbe, C. C. and W. Vahjen. 1993. Interference of humic acids and DNA extracted directly from soil in detection and transformation for recombinant DNA from bacteria and a yeast. Applied Environmental Microbiology 59: 2657-2665.

The Taxonomicon. http://taxonomicon.taxonomy.nl/ downloaded December 9, 2006.

Utermöhl, H., 1958. Zur Vervollkommung der quantitativen Phytoplankton Methodik (Towards a perfection of quantitative phytoplankton methodology). Mitteilungen der Internationale Vereinigung für Theoretische und Angewandte Limnologie 9: 1-38.

White, T. J., T. Bruns, S. Lee and J. Taylor. 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. Pp. 315-322 *in* M. Innis, D. Gelfand, J. Sninsky & T. White (editors), PCR Protocols: A Guide to Methods and Applications. Academic Press, San Diego, California.

Wilson, I. G. 1997. Inhibition and facilitation of nucleic acid amplification. Applied Environmental Microbiology 63: 3741-3751.

Zechman, F. W., E. A. Zimmer and E. C. Theriot. 1994. Use of ribosomal DNA internal transcribed spacers for phylogenetic studies in diatoms. Journal of Phycology 30: 507-512.

Appendix 1. QPCR master mix for *Didymosphenia geminata* using pGEM as internal reference standard.

Each sample should be run with *D.geminata*-primers and probe in duplicate, plus a calibrator sample and no template controls (NTC). Also, each sample, calibrator and NTCs needs to be run in duplicate with pGEM primers and probe. A calibrator sample is one that has been quantified for *D.geminata* microscopically. The current calibrator sample, 40B (Buller River), is known to have 61,718.34 cells/mL. For quantification of unknown samples, this calibrator must be included in each QPCR run.

Note: Every time a new CTAB+pGEM buffer (100 mM Tris-HCl (pH 8), 1.4 M NaCl, 20 mM EDTA, 2% (w/v) cetyltrimethylammonium bromide (CTAB), 0.4% (v/v) ß-mercaptoethanol, 1% (w/v) polyvinylpyrollidone, 20 ng/mL pGEM) is made up, a new calibrator sample or the original calibrator sample must be re-extracted with the new buffer, because the amount of pGEM will vary slightly.

All samples must be extracted with 700 µL of CTAB+pGEM buffer. This amount must be kept constant, so measure carefully! All downstream dilutions will not affect the ratio of pGEM to the template DNA, so dilute the samples to 10 ng/µL in LoTE (3mM Tris-HCl pH 7.5, 0.2mM EDTA pH 8.0).

Samples

In a 72-reaction run (Table A1-1), 13 unknown samples can be run in duplicate for both sets of primers/probe. The run also includes a 5-point standard curve with the Calibrator DNA, and NTCs, as outlined below:

Table A1-1. Examples of samples in a 72-reaction QPCR.

1	40B 10 ng/µL	11	Unknown (6) 10 ng/µL
2	40B 1 ng/µL	12	Unknown (7) 10 ng/µL
3	40B 100 pg/µL	13	Unknown (8) 10 ng/µL
4	40B 10 pg/µL	14	Unknown (9) 10 ng/µL
5	40B 1 pg/µL	15	Unknown (10) 10 ng/µL
6	Unknown (1) 10 ng/µL	16	Unknown (11) 10 ng/µL
7	Unknown (2) 10 ng/µL	17	Unknown (12) 10 ng/µL
8	Unknown (3) 10 ng/µL	18	Unknown (13) 10 ng/µL
9	Unknown (4) 10 ng/µL	19	No template control (NTC)
10	Unknown (5) 10 ng/µL		

Reaction set-up

Prepare a master mix for each primer/probe set in a PCR cabinet, as instructed in Table A1-2. Keep all the reagents on ice, and make sure the final volume is well mixed. Table A1-2 shows the amount of each reagent used per 12.5 µL reaction, and how much to prepare for the whole 72-reaction assay. It is highly recommended that approximately 10% more master mix is made than needed to make up for any pipetting errors that may occur.

Table A1-2. Master mix composition for a 72-reaction experiment.

Master mix reagents	*D.geminata* (602F-753R)		pGEM (M13F-pGEMR)	
	Vol. for single 12.5 µL reaction	**Total vol. needed (µL)**	**Vol. for single 12.5 µL reaction**	**Total vol. needed (µL)**
$MgCl_2$ (50 mM)	1.5	64.5	1.5	51
10x PCR buffer (-$MgCl_2$)	1.25	53.75	1.25	42.5
dNTPs (2 mM)	1.25	53.75	1.25	42.5
Hot-start Taq (5 U/µL)	0.1	4.3	0.1	3.4
Forward Primer (6 µM)	0.1	4.3	0.63	21.42
Reverse Primer (6 µM)	0.63	27.09	1.13	38.42
Probe (X µM)	0.25	10.75	0.25	8.5
BSA (0.2 mg/mL)*	1.5	64.5	1.5	51
MQ-H_2O	4.93	211.99	3.9	132.6

*non-acetylated BSA should be used

Of the 72 reactions, 40 will prepared with the *D.geminata*-primers/probe master mix and 32 with the pGEM primers/probe master mix.

Pipette 11.5 µL of the *D.geminata* master mix into the 0.1 mL labelled 1 – 40. Immediately add 1 µL of sterile MQ-H_2O into tubes 38 – 40 and securely cover with lids. These three reactions are the NTCs for the *D.geminata* master mix.

Pipette 11.5 µL of the pGEM master mix into tubes 41 – 72. Immediately add 1 µL of sterile MQ-H_2O into tubes 70 - 72 and securely cover with lids. These three reactions are the NTCs for the pGEM master mix.

Mix the diluted DNA samples well and then aliquot 1 µL in the pattern described in Table A1-3. Cap the tubes securely immediately after the addition of DNA.

Note – the standard curve (Samples 1 - 5) is only performed with the *D.geminata* master mix, and Sample 5 is in triplicate reactions. With the pGEM master mix, only Sample 1 (40B 10 ng/µL) is used, and in triplicate reactions.

Table A1-3. Reaction set-up for 72-reaction QPCR with the Rotor-Gene 6000. NTC = no-template contols.

A. *D. geminata* reactions

1	Sample 1	9	Sample 5	17	Sample 8	25	Sample 12	33	Sample 16
2	Sample 1	10	Sample 5	18	Sample 9	26	Sample 13	34	Sample 17
3	Sample 2	11	Sample 5	19	Sample 9	27	Sample 13	35	Sample 17
4	Sample 2	12	Sample 6	20	Sample 10	28	Sample 14	36	Sample 18
5	Sample 3	13	Sample 6	21	Sample 10	29	Sample 14	37	Sample 18
6	Sample 3	14	Sample 7	22	Sample 11	30	Sample 15	38	NTC
7	Sample 4	15	Sample 7	23	Sample 11	31	Sample 15	39	NTC
8	Sample 4	16	Sample 8	24	Sample 12	32	Sample 16	40	NTC

B. pGEM reactions

41	Sample 1	49	Sample 8	57	Sample 12	65	Sample 16
42	Sample 1	50	Sample 9	58	Sample 13	66	Sample 17
43	Sample 1	51	Sample 9	59	Sample 13	67	Sample 17
44	Sample 6	52	Sample 10	60	Sample 14	68	Sample 18
45	Sample 6	53	Sample 10	61	Sample 14	69	Sample 18
46	Sample 7	54	Sample 11	62	Sample 15	70	NTC
47	Sample 7	55	Sample 11	63	Sample 15	71	NTC
48	Sample 8	56	Sample 12	64	Sample 16	72	NTC

Place the tubes into the Rotor-Gene 6000 QPCR machine. The FAM-BHQ probe is run on the Green Channel, with a gain of 8. This should be all set up in the "Didymo Taqman" program. Follow the instructions in Appendix 5.

Analysis

1. After run, click on "Analysis" button on tool bar. In the analysis window, identify the run and click "Show".

2. In the Quantitation Analysis window, click on the "Dynamic Tube" button and the "Slope Correct" button.

3. Set the threshold line for C_T calculation. Make sure to view the data in "Log View" and set the C_T line in the linear portion of the curve (the exponential phase of the amplification).

4. Calculate the average C_T for each sample for the *D. geminata* run and for the pGEM run.

5. To determine the number of cells in the sample:

a. For each sample and the calibrator, subtract the average C_T for the pGEM from the average Ct for *D. geminata* to get the ΔC_T.
b. Subtract the ΔCt of each unknown sample from the ΔC_T of the calibrator sample to get – $\Delta\Delta C_T$ for each unknown.
c. Calculate $2^{-\Delta\Delta C_T}$ for each sample to get the ratio of unknown cells to the number of cells in the calibrator sample.
d. Multiply the ratio for each unknown by the cells/mL for the calibrator (i.e. in this study calibrator sample 40B = 61,718cells).
e. Multiply the number of cells by the total volume of the sample (70% ethanol) divided by the volume extracted to get the number of cells per sample.

6. To get litres of water sampled, multiply the flow rate (m/s) by the area (m^2) of the net opening and then multiply this number by the number of seconds the net was in the water. This will give m^3 of river water that flowed through the net. Since 1 dm^3 = 1 L, and there are 1,000 dm^3 in a m^3, multiply the m^3 of water going through the net by 1,000 to get the number of litres sampled.

7. Finally, divide the number of cells for each sample (5e. above) by the number of litres of water sampled (6 above) to get cells/L.

8. This is summarized in the following equation:

$$\text{Cells/m}^3 = [\Delta\Delta C_T \text{ (unknown)}] * (\text{cells in 2 mL of ethanol-fixed calibrator}) * (\text{total volume of ethanol/2}) / [\text{water velocity (m/s)} * \text{area of collection net (m}^2\text{)} * \text{time for collection (s)}]$$

Appendix 2. CTAB and Proteinase K DNA extraction protocol.

1. Mix the samples well and transfer half of each sample into a newly labeled sterile 14 mL Falcon tube. Make a note of the volumes in each tube. Store the original tubes with the remaining volume of sample in a -20°C freezer and archive. Make sure the lids are on tightly, and are sealed with a small strip of Parafilm.
2. In addition to your samples, perform an extraction on an empty 1.7 mL microcentrifuge tube as an extraction 'no template control' (NTC), to ensure reagents are free of contamination.
3. Centrifuge the new tubes for 4 minutes at 8,000 x *g*. Pipette off the 70% ethanol, and replace with approximately the same volume of sterile milliQ water. Vortex to mix, and then centrifuge for 4 minutes at 8,000 x *g*.
4. Pipette off the water, careful not to disturb the sample pellet formed at the bottom of the tube. Add 700 µL of CTAB + pGEM (20 ng/mL) buffer for every mL of original sample. For example, if using a 5 mL aliquot from a 10 mL sample, you would add 3.5 mL (0.7 mL x 5) of CTAB + pGEM buffer.
5. In a fume hood, add 0.4% of the CTAB buffer volume of β-mercaptoethanol (βME) to each sample e.g., add 14 µL of βME if you added 3.5 mL of CTAB + pGEM buffer. Add 3% of the CTAB + pGEM volume of Proteinase K (PK; 10 mg/mL) e.g., add 105 µL of PK to a sample resuspended in 3.5 mL CTAB + pGEM buffer. Vortex to mix well.
6. Incubate in a thermomixer at 55°C for 1 hour, mixing at 1,000 rpm.
7. Increase temperature to 65°C and incubate for a further 20 minutes.
8. Aliquot off 700 µL of each sample into sterile 2 mL microcentrifuge tubes. Store the remaining volume at -20°C freezer as an additional archive.
9. To each microcentrifuge tube, add 700 µL of chloroform: isoamyl alcohol (24:1), and vortex for approximately 30 seconds to form an emulsion. Place samples on a rocker or rotator wheel for 20 minutes.
10. Centrifuge at 16,000 rpm for 15 minutes.
11. Using a 1 mL pipette filter-tip, set at 600 µL pipette as much of the upper phase as possible into a newly labeled sterile 1.7 or 2 mL microcentrifuge, paying special attention to not disturb the interphase or bottom phase.
12. Add 300 µL of 5M NaCl and 600 µL of ice-cold isopropanol. Invert to mix and place in -20°C freezer for at least 20 minutes, allowing DNA precipitation. Samples may be left overnight at this point.
13. Remove from the freezer and centrifuge for at 13,000 x *g* 20 minutes.
14. Carefully decant off the liquid and add 1 mL of 70% ethanol. Invert to mix and then centrifuge for at 13,000 x *g* 20 minutes.
15. Remove the ethanol with a sterile pipette tip, taking extreme care not to dislodge the white/brown DNA pellet at the bottom of the tube.
16. Dry the pellets at room temperature in a fume hood for 10 minutes or use a Speedvac at a low temperature. Resuspend the pellet in 20 µL LoTE. Mix well.
17. Measure the DNA concentration by using 2 µL of the extracted DNA for analysis by a spectrophotometer. If there is enough DNA, approximately 150 ng can be electrophoresed in a 1.5% agarose gel, stained with ethidium bromide and visualized by UV transillumination.

Appendix 3. Long sequencing protocol for validation of positive QPCR results.

1. In a PCR cabinet, set-up and label XX (+ 2) number of 0.2 mL thin walled tubes.

2. Each 50 µL reaction will contain:

Reagent	Volume in 50 µL reaction (µL)	Final concentration
$MgCl_2$ (25 mM)	5.0	2.5 mM
10X PCR Buffer (-$MgCl_2$)	5.0	1X
dNTPs (2 mM)	5.0	0.2 mM
Taq (1U/µL)	2.0	2 Units
Didymo 602F (10 µM)	1.0	0.2 µM
Didymo 1670R (10 µM)	1.0	0.2 µM
BSA (0.2 mg/mL)*	1.2	0.24 µg
MQ-H_20	25.8	-
Genomic DNA (10 ng/µL)	4.0	-

*non-acetylated BSA should be used

3. In a sterile 1.7 mL microcentrifuge, prepare a master-mix by adding XX (+ 3) times each of the above reagents, except the gDNA. Mix gently, and aliquot 46 µL into each of the labeled 0.2 mL tubes.

4. Add 4 µL of thoroughly mixed gDNA (10 ng/µL) to each corresponding tube. In one of the two remaining 0.2 mL tubes, add 4 µL of *Didymosphenia geminata* plasmid or an environmental sample known to be positive for *D. geminata,* this will be the positive control. Add 4 µL of sterile MQ-H_2O to the last tube which will serve as the negative template control (NTC).

5. Close the lids as firmly as possible and place into a Peltier Thermal Cycler (PTC). Expose the reactions to the following thermal cycling parameters:
   i. Initial denaturation of 94°C for 2 minutes
   ii. 40 cycles of 94°C for 20 seconds
   55°C for 1 minute
   72°C for 2 minutes
   iii. Final extension of 72°C for 5 minutes
   iv. Incubation at 4°C until removed from PTC.

6. Mix 5 µL of each reaction with 3 µL of gel loading buffer (0.0083% bromophenol blue, 2.5% ficol), and then pipette into each well of a 1% agarose gel with sodium boric acid (SBA) buffer. Include 5 µL of 1 Kb Plus DNA ladder (Invitrogen) to the outer lane of each gel to aid as a base-pair size estimate of the amplification products.

7. Run the gel at 100 V for 20 minutes. Remove gel and stain with ethidium bromide solution (0.1 mg/L) for 20 minutes, followed by 20 minutes in water to destain.

8. Using ultra-violet transillumination, take a photo to document the gel and identify successful reactions. The presence of a distinct band at approximately 1 Kb indicates

that the 18S rRNA 602 – 1670 bp genomic region of *D. geminata* was successfully amplified.

9. Using a PCR purification kit (e.g., Genscript), purify the remaining volume of the PCR products. Quantify using a spectrophotometer, such as the Nanodrop 1000.

10. Submit the resulting purified PCR products for direct sequencing, along with the D602F and D1670R primers at 5 μM.

11. Upon receipt of sequencing results, directly import the sequences into a DNA sequence alignment tool, such as DNAstar, and compare to the known *D. geminata* sequence. No significant differences should be apparent.

Appendix 4. Preliminary alignment of 1621 bp of partial 18S rDNA and ITS of New Zealand *D. geminata*. Boulder 5, Boulder Creek, Montana, USA; LW1, Lower Waiau, Southland, New Zealand; N2, Nidelva, Norway; OR3, Oreti River, Southland, New Zealand; UKC1, River Coquet, United Kingdom.

```
BC5             CGATGAATGTATTAGCATGGAATAATAAGATAGGACCTTTGTACTATT
LW1             CGATGAATGTATTAGCATGGAATAATAAGATAGGACCTTTGTACTATT
N2              CGATGAATGTATTAGCATGGAATAATAAGATAGGACCTTTGTACTATT
OR3             CGATGAATGTATTAGCATGGAATAATAAGATAGGACCTTTGTACTATT
UKC1            CGATGAATGTATTAGCATGGAATAATAAGATAGGACCTTTGTACTATT

BC5             TTGTTGGTTTGTGTATAGAGGTAATGATTAAAAGGAACAGTTGGGGGT
LW1             TTGTTGGTTTGTGTATAGAGGTAATGATTAAAAGGAACAGTTGGGGGT
N2              TTGTTGGTTTGTGTATAGAGGTAATGATTAAAAGGAACAGTTGGGGGT
OR3             TTGTTGGTTTGTGTATAGAGGTAATGATTAAAAGGAACAGTTGGGGGT
UKC1            TTGTTGGTTTGTGTATAGAGGTAATGATTAAAAGGAACAGTTGGGGGT

BC5             ATTTGTATTCCATTGTCAGAGGTGAAATTCTTGGATTTTTGGAAGACA
LW1             ATTTGTATTCCATTGTCAGAGGTGAAATTCTTGGATTTTTGGAAGACA
N2              ATTTGTATTCCATTGTCAGAGGTGAAATTCTTGGATTTTTGGAAGACA
OR3             ATTTGTATTCCATTGTCAGAGGTGAAATTCTTGGATTTTTGGAAGACA
UKC1            ATTTGTATTCCATTGTCAGAGGTGAAATTCTTGGATTTTTGGAAGACA

BC5             AACTACTGCGAAAGCATTTACCAAGGATGTTTTCATTAATCAAGAACG
LW1             AACTACTGCGAAAGCATTTACCAAGGATGTTTTCATTAATCAAGAACG
N2              AACTACTGCGAAAGCATTTACCAAGGATGTTTTCATTAATCAAGAACG
OR3             AACTACTGCGAAAGCATTTACCAAGGATGTTTTCATTAATCAAGAACG
UKC1            AACTACTGCGAAAGCATTTACCAAGGATGTTTTCATTAATCAAGAACG

BC5             AAAGTTAGGGGATCGAAGATGATTAGATACCATCGTAGTCTTAACCAY
LW1             AAAGTTAGGGGATCGAAGATGATTAGATACCATCGTAGTCTTAACCAT
N2              AAAGTTAGGGGATCGAAGATGATTAGATACCATCGTAGTCTTAACCAT
OR3             AAAGTTAGGGGATCGAAGACGATTAGATACCATCGTAGTCTTAACCAT
UKC1            AAAGTTAGGGGATCGAAGATGATTAGATACCATCGTAGTCTTAACCAT
                                                               *

BC5             AAACTATGCCAACAAGGGATTGGTGGGGTTTCGTAATGTCCCCATCAG
LW1             AAACTATGCCAACAAGGGATTGGTGGGGTTTCGTAATGTCCCCATCAG
N2              AAACTATGCCAACAAGGGTTTGGGGGGGTTTCCTAATGTCCCCATCAG
OR3             AAACTATGCCAACAAGGGATTGGTGGGGTTTCGTAATGTCCCCGTCAG
UKC1            AAACTATGCCAACAAGGGATTGGTGGGGTTTCGTAATGTCCCCATCAG

BC5             CACCTTAGGAGAAATCAAAAGTTTTTGGGTTCCGGGGGGAGTATGGTC
LW1             CACCTTAGGAGAAATCAAAAGTTTTTGGGTTCCGGGGGGAGTATGGTC
N2              CACCTTAGGAGAAATCAAAAGTTTTTGGGTTCCGGGGGGAGTATGGTC
OR3             CACCTTAGGAGAAATCAAAAGTTTTTGGGTTCCGGGGGGAGTATGGTC
UKC1            CACCTTAGGAGAAATCAAAAGTTTTTGGGTTCCGGGGGGAGTATGGTC

BC5             GCAAGGCTGAAACTTAAAGAAATTGACGGAAGGGCACCACCAGGAGTG
LW1             GCAAGGCTGAAACTTAAAGAAATTGACGGAAGGGCACCACCAGGAGTG
N2              GCAAGGCTGAAACTTAAAGAAATTGACGGAAGGGCACCACCAGGAGTG
OR3             GCAAGGCTGAAACTTAAAGAAATTGACGGAAGGGCACCACCAGGAGTG
UKC1            GCAAGGCTGAAACTTAAAGAAATTGACGGAAGGGCACCACCAGGAGTG
```

```
BC5             GAGCCTGCGGCTTAATTTGACTCAACACGGGGAAAYTTACCAGGTCCR
LW1             GAGCCTGCGGCTTAATTTGACTCAACACGGGGAAACTTACCAGGTCCA
N2              GAGCCTGCGGCTTAATTTGACTCAACACGGGGAAACTTACCAGGTCCA
OR3             GAGCCTGCGGCTTAATTTGACTCAACACGGGGAAACTTACCAGGTCCA
UKC1            GAGCCTGCGGCTTAATTTGACTCAACACGGGGAAACTTACCAGGTCCA

BC5             GACATAGTGAGGATTGACAGATTGAGAGCTCTTTCTTGATTCTATGGG
LW1             GACATAGTGAGGATTGACAGATTGAGAGCTCTTTCTTGATTCTATGGG
N2              GACATAGTGAGGATTGACAGATTGAGAGCTCTTTCTTGATTCTATGGG
OR3             GACATAGTGAGGATTGACAGATTGAGAGCTCTTTCTTGATTCTATGGG
UKC1            GACATAGTGAGGATTGACAGATTGAGAGCTCTTTCTTGATTCTATGGG

BC5             TGGTGGTGCATGGCCGTTCTTAGTTGGTGGAGTGATTTGTCTGGTTAA
LW1             TGGTGGTGCATGGCCGTTCTTAGTTGGTGGAGTGATTTGTCTGGTTAA
N2              TGGTGGTGCATGGCCGTTCTTAGTTGGTGGAGTGATTTGTCTGGTTAA
OR3             TGGTGGTGCATGGCCGTTCTTAGTTGGTGGAGTGATTTGTCTGGTTAA
UKC1            TGGTGGTGCATGGCCGTTCTTAGTTGGTGGAGTGATTTGTCTGGTTAA

BC5             TTCCGTTAACGAACGAGACCGCTGCCTGCTAAATAGACTAGTTAGTGA
LW1             TTCCGTTAACGAACGAGACCGCTGCCTGCTAAATAGACTAGTTAGTGA
N2              TTCCGTTAACGAACGAGACCGCTGCCTGCTAAATAGACTAGTTAGTGA
OR3             TTCCGTTAACGAACGAGACCGCTGCCTGCTAAATAGACTAGTTAGTGA
UKC1            TTCCGTTAACGAACGAGACCGCTGCCTGCTAAATAGACTAGTTAGTGA

BC5             TTTTCACTGATCAGGTCTTCTTAGAGGGACGTGCATTTTATTAGATGC
LW1             TTTTCACTGATCRGGTCTTCTTAGAGGGACGTGCATTTTATTAGATGC
N2              TTTTCACTGATCAGGTCTTCTTAGAGGGACGTGCATTTTATTAGATGC
OR3             TTTTCACTGATCAGGTCTTCTTAGAGGGACGTGCATTTTATTAGATGC
UKC1            TTTTCACTGATCAGGTCTTCTTAGAGGGACGTGCATTTTATTAGATGC

BC5             AGGAAGATAGCGGCAATAACAGGTCTGTGATGCCCTTAGATGATCTGG
LW1             AGGAAGATAGCGGCAATAACAGGTCTGTGATGCCCTTAGATGATCTGG
N2              AGGAAGATACCGGCAATAACAGGTCTGGGATGCCCTTAAATGATCTGG
OR3             AGGAAGATAGCGGCAATAACAGGTCTGTGATGCCCTTAGATGATCTGG
UKC1            AGGAAGATAGCGGCAATAACAGGTCTGTGATGCCCTTAGATGATCTGG

BC5             GCCGCACGCGCGCTACACTGATGTACTCAACGAGTTTTTCCTTGGCTG
LW1             GCCGCACGCGCGCTACACTGATGTACTCAACGAGTTTTTCCTTGGCTG
N2              GCCGCACGCGCGTTACACTGATGTACTCAACGAGTTTTTCCTTGGCTG
OR3             GCCGCACGCGCGCTACACTGATGTACTCAACGAGTTTTTCCTTGGCTG
UKC1            GCCGCACGCGCGCTACACTGATGTACTCAACGAGTTTTTCCTTGGCTG

BC5             AGAAGCCTGGGTAATCTTTTAAACTTGCATCGTGATAGGGATAGATTA
LW1             AGAAGCCTGGGTAATCTTTTAAACTTGCATCGTGATAGGGATAGATTA
N2              AGAAGCCTGGGTAATCTTTTAAACTTGCATCGTGATAGGGATAGATTA
OR3             AGAAGCCTGGGTAATCTTTTAAACTTGCATCGTGATAGGGATAGATTA
UKC1            AGAAGCCTGGGTAATCTTTTAAACTTGCATCGTGATAGGGATAGATTA

BC5             TTGCAATTATTAATCTTGAACGAGGAATTCCTAGTAAACGCAGATCAT
LW1             TTGCAATTATTAATCTTGAACGAGGAATTCCTAGTAAACGCAGATCAT
N2              TTGCAATTATTAATCTTGAACGAGGAATTCCTAGTAAACGCAGATCAT
OR3             TTGCAATTATTAATCTTGAACGAGGAATTCCTAGTAAACGCAGATCAT
UKC1            TTGCAATTATTAATCTTGAACGAGGAATTCCTAGTAAACGCAGATCAT

BC5             CAGTCTGCATTGATTACGTCCCTGCCCTTTGTACACACCGCCCGTCGC
LW1             CAGTCTGCAATGATTACGTCCCTGCCCTTTGTACACACCGCCCGTCGC
N2              CAGTCTGCATTGATTACGTCCCTGCCCTTTGTACACACCGCCCGTCGC
OR3             CAGTCTGCATTGATTACGTCCCTGCCCTTTGTACACACCGCCCGTCGC
UKC1            CAGTCTGCATTGATTACGTCCCTGCCCTTTGTACACACCGCCCGTCGC
```

```
BC5             ACCTACCGATTGAATGGTCCGGTGAAGGCTCGGGATTGCAGCTAATGC
LW1             ACCTACCGATTGAATGGTCCGGTGAAGGCTCGGGATTGCAGCTAATGC
N2              ACCTACCGATTGAATGGTCCGGTGAAGGCTCGGGATTGCAGCTAATGC
OR3             ACCTACCGATTGAATGGTCCGGTGAAGGCTCGGGATTGCAGCTAATGC
UKC1            ACCTACCGATTGAATGGTCCGGTGAAGGCTCGGGATTGCAGCTAATGC

BC5             CTTTACTGGTGTTGGTTTCAAGAACTTGTCTAAACCTTATCATTTAGA
LW1             CTTTACTGGTGTTGGTTTCAAGAACTTGTCTAAACCTTATCATTTAGA
N2              CTTTACTGGTGTTGGTTTCAAGAACTTGTCTAAACCTTATCATTTAGA
OR3             CTTTACTGGTGTTGGTTTCAAGAACTTGTCTAAACCTTATCATTTAGA
UKC1            CTTTACTGGTGTTGGTTTCAAGAACTTGTCTAAACCTTATCATTTAGA

BC5             GGAAGGTGAAGTCGTAACAAGGTTTCCGTAGGTGAACCTGCGGAAGGA
LW1             GGAAGGTGAAGTCGTAACAAGGTTTCCGTAGGTGGACCTGCGGAAGGA
N2              GGAAGGTGAAGTCGTAACAAGGTTTCCGTAGGTGAACCTGCGGAAGGA
OR3             GGAAGGTGAAGTCGTAACGAGGTTTCCGTAGGTGAACCTGCGGAAGGA
UKC1            GGAAGGTGAAGTCGTAACAAGGTTTCCGTAGGTGAACCTGCGGAAGGA
```

Appendix 5. Manual for operating the Corbett Life Science Rotor-Gene 6000.

The Rotor-Gene 6000 is a centrifugal rotary based real-time quantitative PCR (QPCR) machine specifically designed to meet the requirements of real-time analyses.
Most conventional QPCR instruments use a 96-well heat block thermocycler with each site having an individual optical excitation and detection system. This design leads to inconsistent thermal and optical conditions across the cycling block (Hermann *et al,*. 2006). This well-to-well variability can have an adverse affect on the quantification results and sample-to-sample comparison. To overcome these problems the reaction tubes in the Rotor-Gene are arranged in a circular rotor, which is situated in a heating/cooling chamber allowing thermal uniformity. Optical detection is also uniform since all samples are measured using the same excitation and detection light source.

The Rotor-Gene 6000 used for this project.

Preparing a Run

Rotor type
The Rotor-Gene 6000 comes with three different rotor types: 36-well rotor, 72-well rotor and a Gene Disc (either 72 or 100 reaction tubes).

The 36-well rotor and locking ring requires 0.2 ml tubes and for most experiments uses a reaction volume of 25 µl. There is no need to ensure the tubes contain optical clear caps as the Rotor-Gene 6000 reads fluorescence from the bottom of the tubes rather than at the top.

The 72-well rotor and locking ring uses 0.1 ml tubes, which come in strips of four. The optimal reaction volume in these tubes is 12.5 µl.

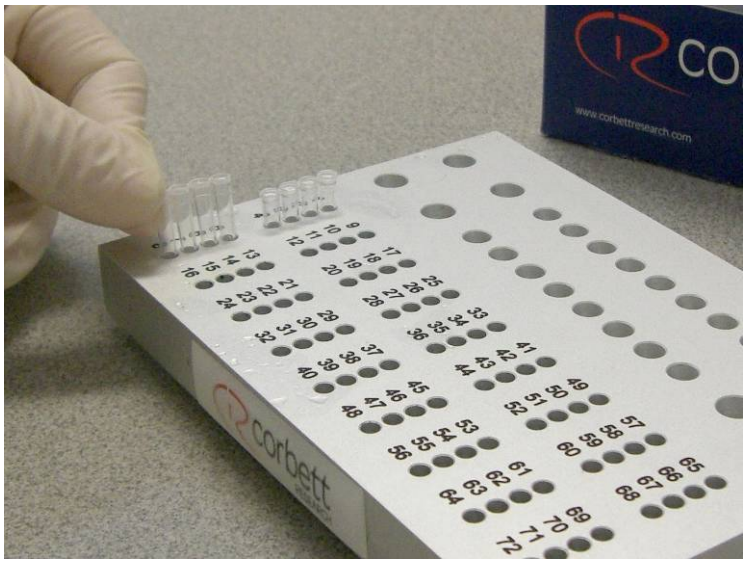
For automated reaction set up allowing high throughput analysis Gene Discs, which hold either 72 or 100 reaction tubes, are available. The Gene Discs tubes come in one

structure, which is sealed on top with a clear polymer film providing a strong, durable and tamper-proof seal.

Reaction Setup

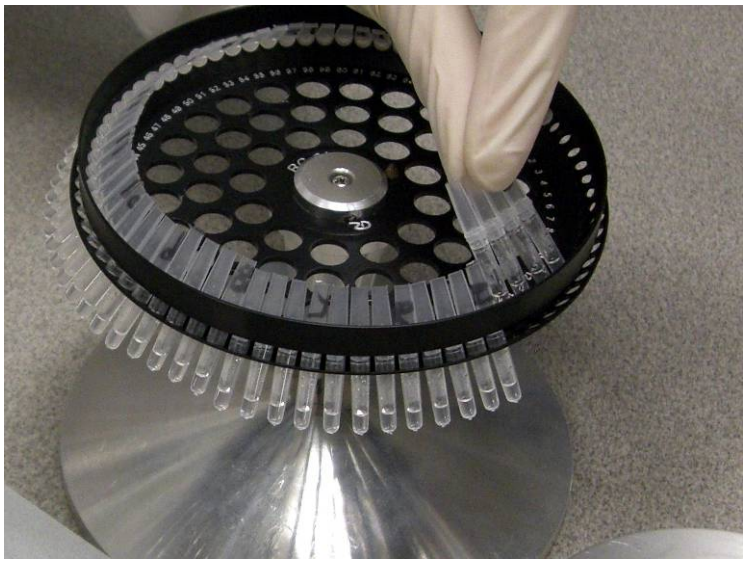
The following protocol describes the setup of a QPCR run on a Rotor-Gene 6000 using a 72-well rotor and 0.1 ml tubes. When using the 36-well rotor 0.2 ml tubes would be used.

1. Insert the 0.1 ml tubes (as strips of four) into the loading block. Note: the 0.2 ml tubes used for the 36-well rotor are single tubes. The block is made of aluminum and should be pre-cooled at 4°C. Aliquot the PCR master mixture and the template DNA into the tubes.

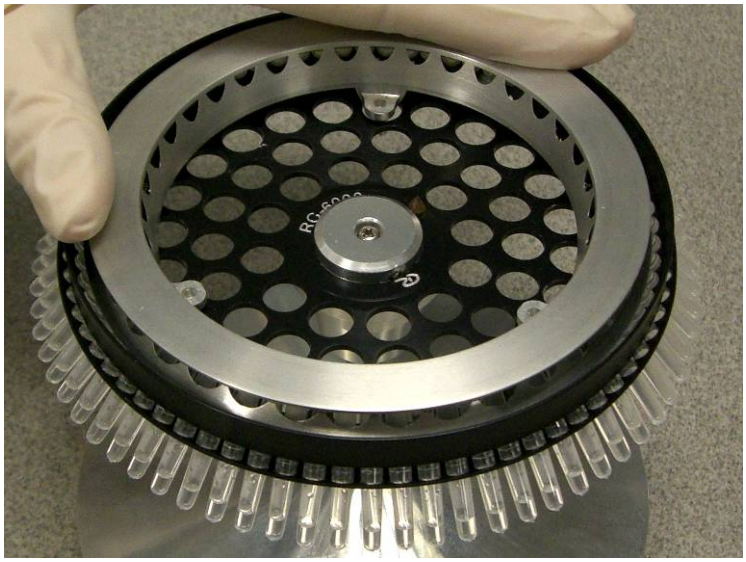


2. Seal the tubes tightly with the caps (as strips of four) to prevent any evaporation during thermocycling.

3. Insert the reaction tubes into the 72-well rotor hub and ensure that each tube sits correctly into place. Note: All 72 wells in the rotor must be filled (fill with empty tubes if less than 72 samples are to be run).



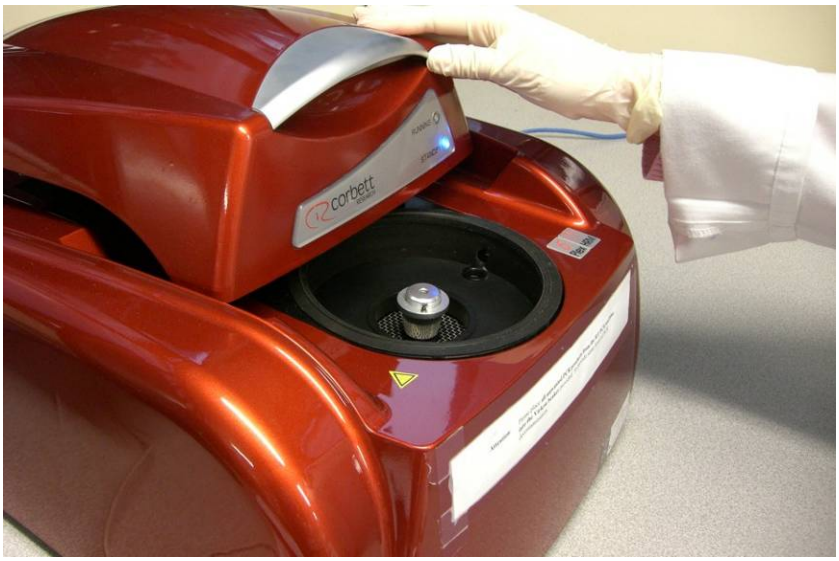
4. Press the 72-well locking ring on top of the rotor to secure tubes.



Loading the Rotor-Gene 6000

1. Turn the Rotor-Gene 6000 and computer on. The switch for the Rotor-Gene 6000 is located on the back right hand side.

2. Open the instrument lid by gently sliding the lid up and back exposing the instruments themocycling chamber.

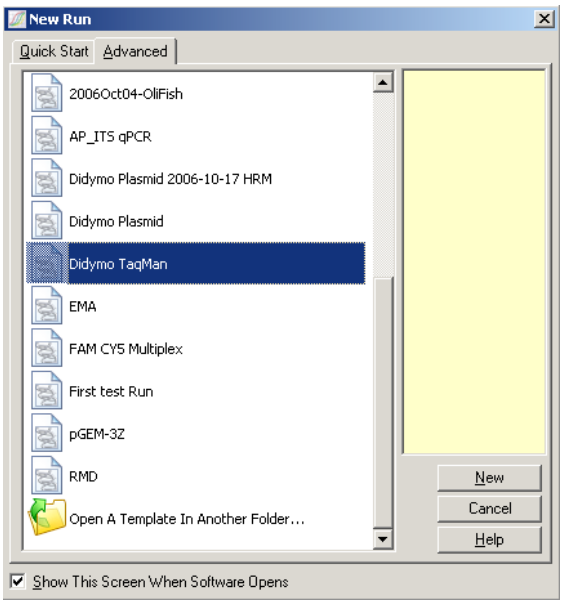


3. Insert the 72-well rotor into the thermocycling chamber. Align the positioning pin with the "position one" hole in the 72-well rotor disc and click into place. Close the instrument lid.

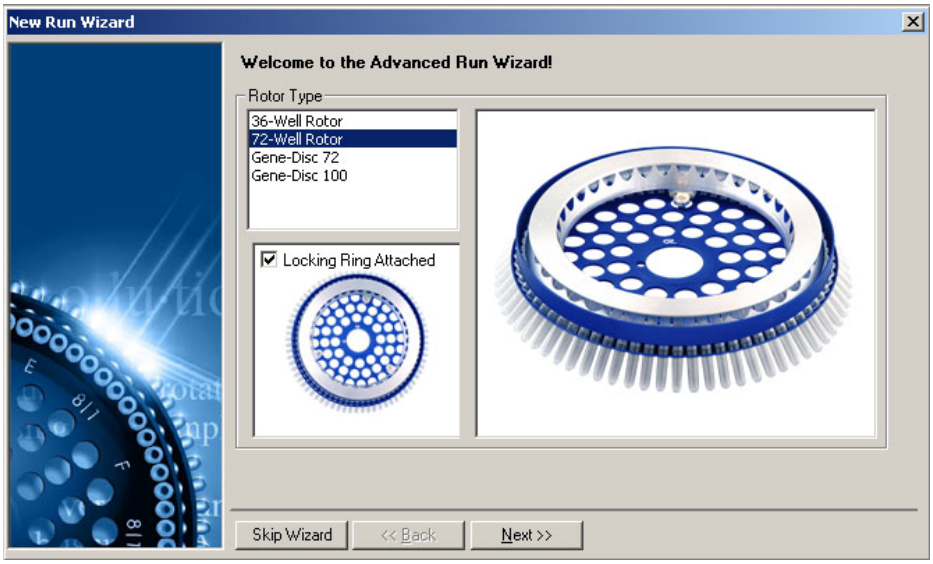
Starting a QPCR Run

1. Open the Rotor Gene Software by double-clicking the Rotor-Gene Desktop icon.

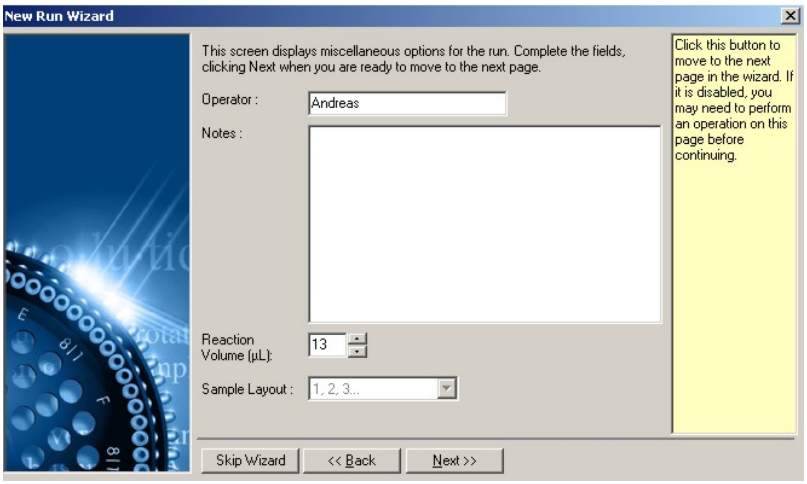
2. Create a Run: Open the "New Run" window by clicking the "New Run" icon and select "Didymo Taqman" by double-clicking. When setting up a new run with reactions that have not been previously run on the Rotor-Gene then the thermal cycling conditions for the SYBR Green I and the TaqMan assay given in sections 2.3.1 and 2.4.2.2 have to be set manually.



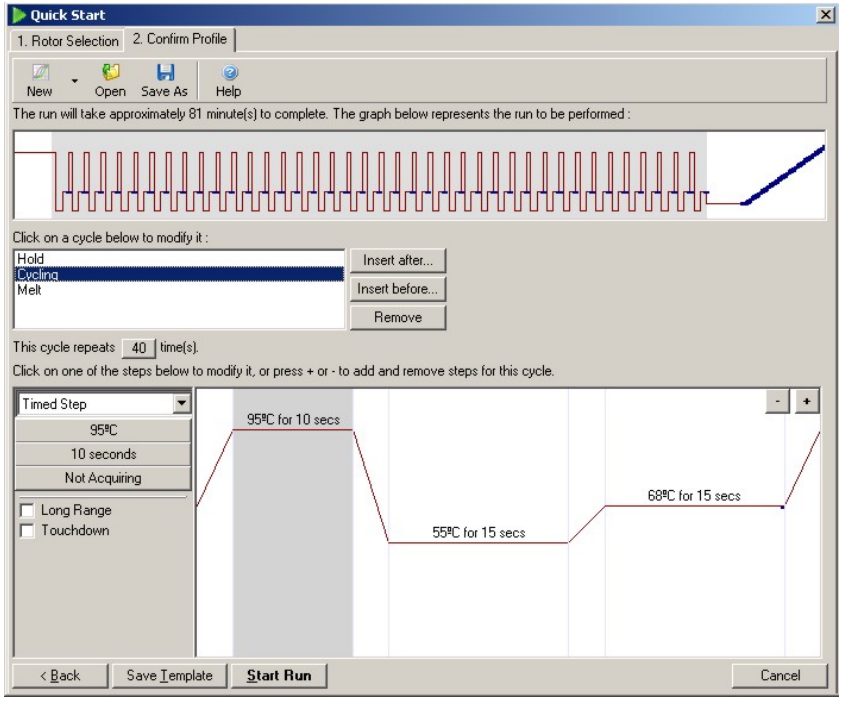
3. Rotor Selection: Select "72-well Rotor", check you have the lock ring in place and then tick the "Locking Ring Attached" box.



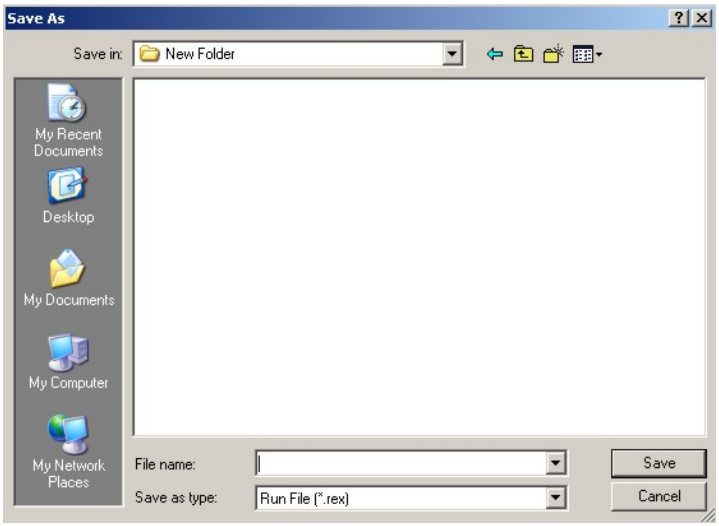
4. Operator name and sample volume: In the next display the name of the operator and notes about the run can be entered. The reaction volume must be entered.



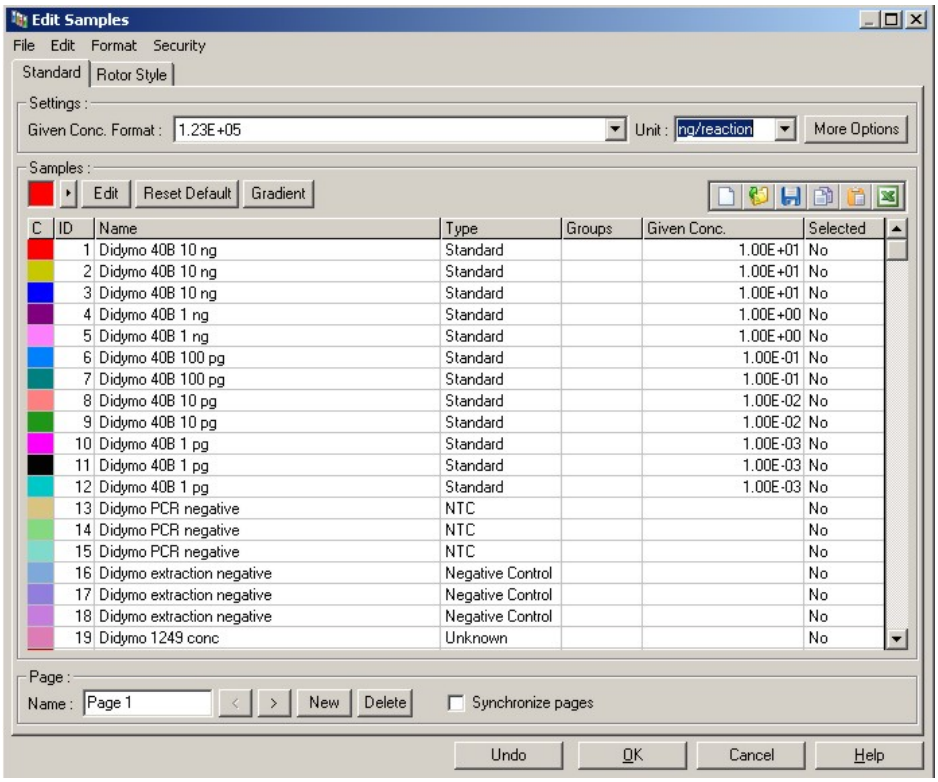
5. Confirm Profile: The "Didymo Taqman" template chosen above imports the cycling conditions, gain settings and acquisition channel, e.g., green channel for SYBR Green and FAM TaqMan probes. When setting up a new run with reactions that have not yet been previously run on the Rotor-Gene then the gain setting can automatically be aquired by using the Auto-Gain Optimization Setup. For more details please refer to the Rotor-Gene 6000 manual book.



6. Start Run: Clicking the "Start Run" button brings up a "Save As" window. The run needs to be saved now by entering a filename. Once a filename has been entered click "Save". This will initiate the start of the QPCR reaction.



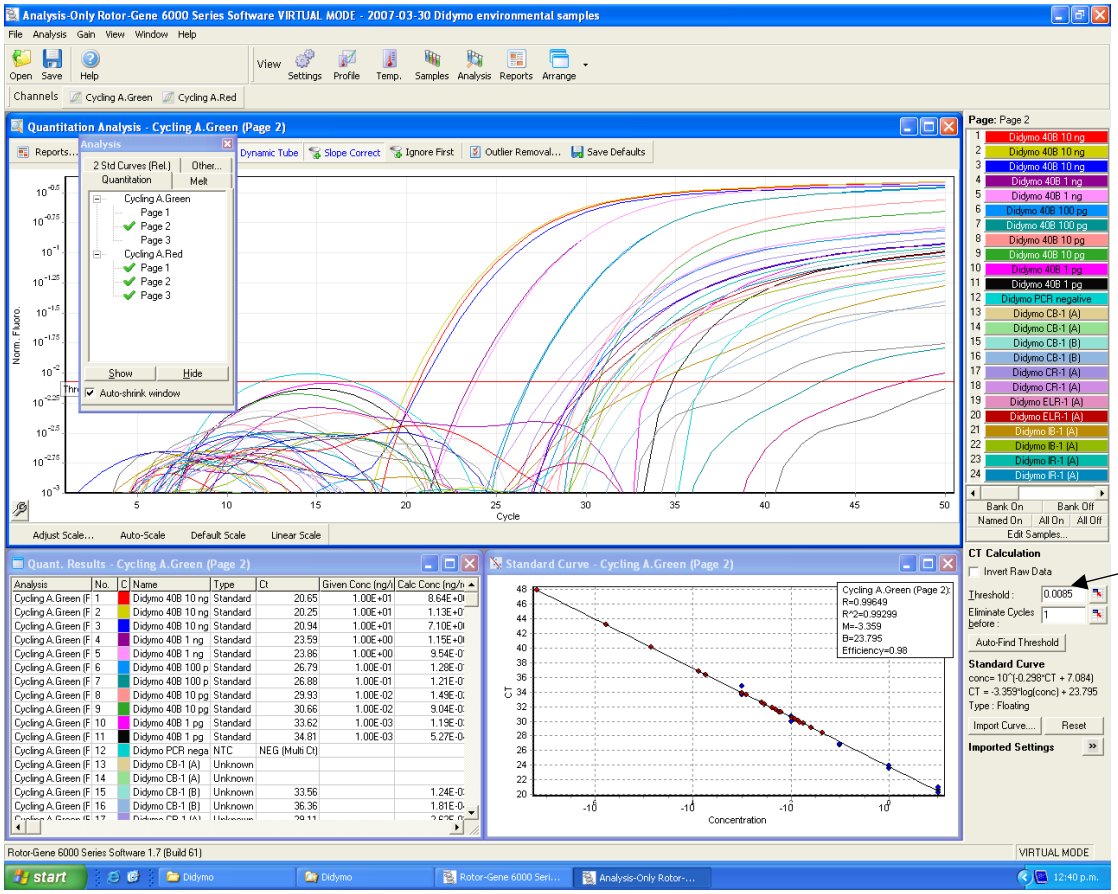
7. Sample Setup: Once the run has been started it will automatically load the sample editor; this allows the user to define sample types and names. Sample names are entered into the "Name" column. The identity of the sample can be chosen in the "Type" column. For instance, treatment samples are selected as "Unknown" and calibration samples are selected as "Standard". Negative control samples are selected as "NTC". The column "Given Conc." is used only for calibration samples and indicates the amount of template DNA used.



C	ID	Name	Type	Groups	Given Conc.	Selected
	1	Didymo 40B 10 ng	Standard		1.00E+01	No
	2	Didymo 40B 10 ng	Standard		1.00E+01	No
	3	Didymo 40B 10 ng	Standard		1.00E+01	No
	4	Didymo 40B 1 ng	Standard		1.00E+00	No
	5	Didymo 40B 1 ng	Standard		1.00E+00	No
	6	Didymo 40B 100 pg	Standard		1.00E-01	No
	7	Didymo 40B 100 pg	Standard		1.00E-01	No
	8	Didymo 40B 10 pg	Standard		1.00E-02	No
	9	Didymo 40B 10 pg	Standard		1.00E-02	No
	10	Didymo 40B 1 pg	Standard		1.00E-03	No
	11	Didymo 40B 1 pg	Standard		1.00E-03	No
	12	Didymo 40B 1 pg	Standard		1.00E-03	No
	13	Didymo PCR negative	NTC			No
	14	Didymo PCR negative	NTC			No
	15	Didymo PCR negative	NTC			No
	16	Didymo extraction negative	Negative Control			No
	17	Didymo extraction negative	Negative Control			No
	18	Didymo extraction negative	Negative Control			No
	19	Didymo 1249 conc	Unknown			No

8. Run analysis: Clicking the Analysis menu brings up the Analysis window. This window allows you to create a new analysis or display an existing one. Once you have done this, choose the channels and sample page to be analysed. To analyse a channel or to view an existing channel, simply double-click on the channel to be viewed. This will open a specific analysis window.

With the samples already type-selected, i.e. standards, unknowns, NTC, etc. you can quantify your samples by either clicking on "AutoFind Threshold" or by manually selecting the threshold. To do this click on the icon (grid with red arrow) and drag the threshold baseline (red line) to the desired level or entering a value in the threshold box (see below).



9. Export run data into Excel: The results obtained from the run are summarized in a table. Clicking the right mouse button, and selecting "Export to Excel" will automatically export the table into Excel.

High Resolution Melt Analysis

1. Once the QPCR run has been analyzed, remove the rotor with tubes still in place, and store at 4°C temporarily.

2. Using sterile milliQ water, make up a 1/200 dilution of the stock 5 mM Syto 9 green fluorescent nucleic acid stain, (Invitrogen).

3. Remove rotor and tubes from storage, and open the lids. If using the 0.1 mL tubes in strips label the lids and do not throw away once removed, as you will use them again.

4. Add 5 µL of the diluted Syto 9 solution to 25 µL QPCR reaction. If using different QPCR volumes, alter the amount of Syto 9 added appropriately.

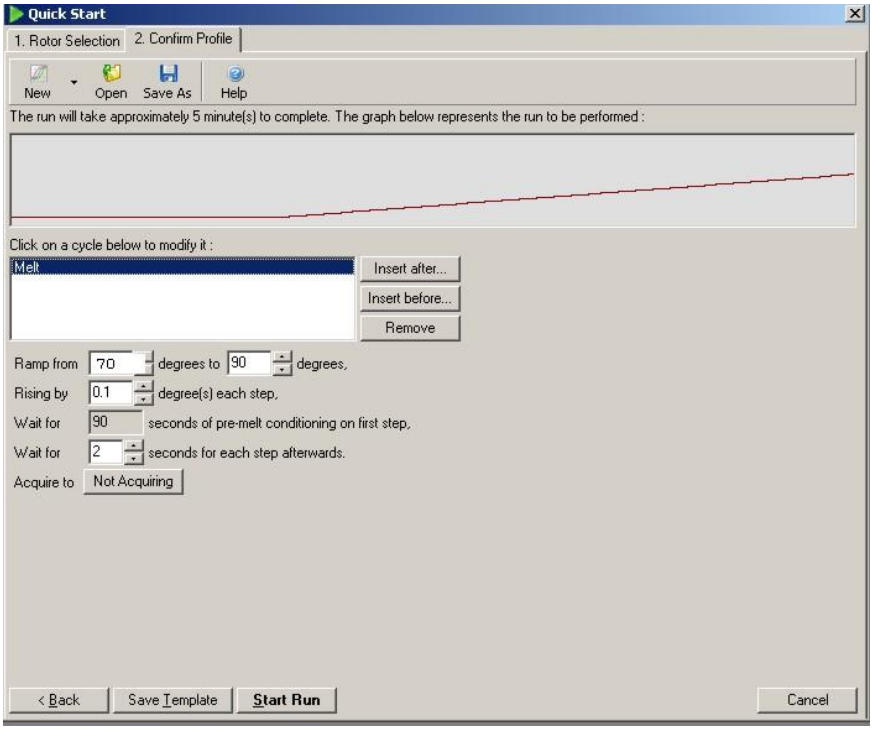
5. Close the lids again, making sure they are sealed tightly. Mix well, and ensure the contents of each tube are at the bottom and not on the sides.

6. Place the rotor and tubes back into the Rotor-Gene 6000. Close the instrument lid.

7. Open the Rotor-Gene software and choose a new 'HRM Analysis' run.

8. Enter the appropriate rotor-type and reaction volume as described previously. However, remember that the reaction volume has now increased since the addition of Syto 9.

9. Perform the melt from 70 - 90°C, in 0.1°C increments.



10. Assign a file name for the HRM run and save to the appropriate location. The run will now start.

11. Once completed, click on the 'Analysis' button. Click on the 'Melt' tab and then highlight 'HRM Analysis on HRM channel'. Click 'Show'.

12. Two windows will now appear: one illustrating the melt curves for each reaction and the other showing the temperature of each peak shown (set Threshold first).

13. Using the melt curves produced by the positive control reactions (calibrator 40B), compare any peaks produced by the unknown samples. The 151 bp product amplified by D602F-D753R should produce a melt curve with a peak (T_m) at approximately

83°C. Melt curves around 75 - 80°C are produced by primer-dimers and do not indicate a positive amplification of *D.geminata* 18S rDNA.

Reference

Hermann, M. G., J. D. Durtschi, L. K. Bromle, C. T. Wittwer and K. V. Voelkerding. 2006. Amplicon DNA Melting Analysis for Mutation Scanning and Genotyping: Cross-Platform Comparison of Instruments and Dyes. Clinical Chemistry 53(3), 494-503.