A rapid and cost-effective quantitative microsatellite genotyping protocol to estimate intraspecific competition in protist microcosm experiments

Ewan J. A. Minter¹,², Chris D. Lowe²,³, Michael A. Brockhurst⁴ and Phillip C. Watts²,⁴

¹Department of Biology, University of York, Wentworth Way, York, Yorkshire YO10 5DD, UK; ²Institute of Integrative Biology, University of Liverpool, Biosciences Building, Crown Street, Liverpool L69 7ZB, UK; ³Centre for Ecology and Conservation, College of Life and Environmental Sciences, University of Exeter, Cornwall Campus, Falmouth TR10 9FE, UK; and ⁴Department of Biology, University of Oulu, PO Box 3000, FI-90014 Oulu, Finland

Summary

1. High levels of intraspecific variation are commonly observed in natural microbial populations, yet the consequences of this variation for ecological and evolutionary processes remain poorly understood. Protists are excellent experimental models for investigating fundamental and applied questions in ecology and evolution, but studying intraspecific variation remains a challenge due to a lack of molecular resources to aid in quantifying and distinguishing strains during experiments.

2. Here we present a molecular method, quantitative microsatellite genotyping, to accurately quantify strain-specific frequencies from microcosm experiments of the marine flagellate Oxyrrhis marina, both between many pairs of strains and between strains in a multistrain mixture.

3. We find that for pairs of strains, the method is effective for relative frequencies as low as 0.02 and with around 99% accuracy. The method is able to quantify four strains reasonably well, though less accurate than for pairs (range 92–97% accuracy).

4. This makes accessible a cheap and easy-to-implement method for quantifying strain (or allele) frequencies and is suitable for use in a broad range of single-celled eukaryotes (protists) where copy number should correlate well with number of individuals (i.e. cells). This opens up the possibility of examining the role of intraspecific variation using experimental protist microcosms.

Key-words: competition, experiments, frequency, intraspecific, microsatellites, molecular assay, Oxyrrhis, selection

Introduction

There is a pressing need to understand the functional ecological consequences of intraspecific genetic variation. Intraspecific genetic diversity is predicted to have a fundamental impact upon a range of ecological and evolutionary processes (Bijlsma & Loeschcke 2005; Hughes et al. 2008; Bolnick et al. 2011), including population dynamics (Lankau 2009), ecosystem resilience to disturbance (Hughes & Stachowicz 2004) and the impact of pathogens upon their hosts (Balmer et al. 2009).

High genetic diversity can protect populations from extinction by facilitating evolutionary responses to environmental change (Markert et al. 2010); therefore, intraspecific diversity could play a crucial role in conservation (Reed & Frankham 2003; Frankham 2005). Our understanding of the impact of intraspecific genetic diversity on such population processes is, however, limited by a lack of experimental data: existing experimental studies use either low or unquantified levels of intraspecific genetic diversity, and predictions (i.e. model parameterization) tend to be based on the response of just one or relatively few genotypes (e.g. Yang et al. 2013).

Protists (eukaryotic microbes) are well suited for experimental studies of ecological processes due to their fast growth rates, small size and the ease with which one may manipulate and control experimental conditions (e.g. Petchey et al. 1999; Collins & Bell 2004; Friman et al. 2008). Moreover, protists are ecologically significant, underpinning numerous ecological processes, particularly in aquatic systems where protists drive primary production, nutrient cycling and recycling carbon via the microbial loop (Pernthaler 2005). High levels of intraspecific diversity are a feature of many natural protist populations (Evans, Kühn & Hayes 2005; Logares et al. 2009; Harnstrom et al. 2011; Lowe et al. 2011d) and could be as important for protist ecology as interspecific variation (Weisse 2002); however, a lack of morphological features and small size makes distinguishing among protist strains within mixed cultures impossible without recourse to a molecular assay. Correspondingly, one aspect that has been largely overlooked in protist experimental systems is the precise role of intraspecific genetic variation. The individual contributions of separate strains to the experimental outcome are either not quantified or has to be
estimated using a considerable resampling effort that imposes time and cost constraints, for example re-isolating many single cells/colonies from the experimental system to establish monocultures that are then individually genotyped to estimate strain frequencies (e.g. Roger, Godhe & Garnfeldt 2012).

In principle, the genetic composition of a sample can readily be obtained by genotyping a pool of DNA samples within a single PCR, whereby the frequency of each allele indicates the amount of genetic material that each sample contributed to the DNA pool (Gruber, Colligan & Wolford 2002). For most purposes, DNA pooling is used to rapidly estimate population genetic parameters (e.g. differences in allele frequencies and diversity) at many loci and among several, large samples. Hence, contemporary protocols for genotyping DNA pools generally assess variation at single-nucleotide polymorphisms (SNPs) (e.g. Gruber, Colligan & Wolford 2002; Lavebratt & Sengul 2006; Doostzadeh et al. 2008) because of convenient genotyping methodology and as SNPs are the most abundant genetic marker in the genome (Schlotterer 2004). For model protists where whole-genome data are available for multiple strains (e.g. Saccharomyces, Chlamydomonas), it may be straightforward to identify many SNPs. Of course, next-generation sequencing technology offers reasonable prospects for identifying genomic SNPs in non-model species (Nielsen et al. 2011). However, where a study aims to follow the fate of many different genotypes during the course of an experiment, as opposed to quantifying population genomic parameters, it is typically difficult to identify short genomic regions (i.e. distinct and easily genotyped loci) that possess sufficient SNPs to be able to differentiate among more than just a few strains. For protist species lacking genomic resources, therefore, there is a need to develop a cost-effective and rapid molecular genetic assay that can quantify the contributions of different strains to ecological outcomes from within experimental populations.

One such genotyping approach is to take advantage of the typically high polymorphism exhibited by microsatellite loci (Ellegren 2004; Schlotterer 2004); for example, in the eastern North Atlantic, samples of the marine flagellate Oxyrrhis marina had between 6 and 36 different alleles at 9 microsatellite loci (Lowe et al. 2011d), but just three different haplotypes at two loci (Lowe et al. 2012). Indeed, microsatellite-based approaches to genotype DNA pools were developed to quantify diversity and differentiation among metazoan samples (Khatib et al. 1994; Hillel et al. 2003; Skalski et al. 2006), with the amount of PCR product of every allele (estimated from electropherogram peak heights) reflecting the amount of respective template DNA. Hence, accurate genotyping of metazoan samples requires rigorous standardization of sample quality and quantity, so that each sample makes a tangible contribution during PCR. For protist samples, the allele frequencies per se should indicate the relative numbers of cells of each genotype (strain) within the mixed sample (e.g. see Meyer et al. 2006; Takala et al. 2006; Cruz et al. 2010).

To study the dynamics of individual protist strains within mixed experimental cultures, we (i) describe the development of new microsatellite loci that can differentiate between multiple strains of a protist (the flagellate Oxyrrhis marina) and then (ii) demonstrate how allele-specific quantitative PCR (AsQ-PCR) microsatellite genotyping (Meyer et al. 2006) can accurately determine the relative abundance of different strains (i.e. different alleles) from within mixed cultures and (iii) demonstrate proof of principle by quantifying multistrain dynamics in growth experiments. This easy, fast and cost-effective technique opens up novel avenues for future research using protists to study intraspecific variation in experimental microcosms.

Materials and methods

CELL CULTURING AND DNA EXTRACTION

We examined strains of the marine flagellate Oxyrrhis marina (Dujardin, 1841), a protist that is widely used in experimental studies (Lowe et al. 2011a; Yang et al. 2013). Seven strains of O. marina were isolated from the North East Atlantic Ocean: BGN01, BOD02, TM001, PLY01, ROS03, FAR01 and EST02 (Table 1). Stock cultures were maintained at 16°C in 32 PSU sterile-filtered artificial sea water (ASW) enriched with f/2 (Sigma Aldrich, St Louis, MO, USA) and fed on the green alga Dunaliella primolecta CCAP11/34 (Chlorophyta) at a cell density of approximately 3 × 10⁵ cells mL⁻¹, and maintained at a light intensity of approximately 80 μmol photons m⁻² s⁻¹ on a light:dark cycle of 14 h:10 h. Stock cultures were maintained by subculturing once per month. Approximately 2 weeks prior to DNA extraction, strains were subcultured in the dark and fed on heat-killed E. coli (Lowe et al. 2011b) to deplete D. primolecta (i.e. reduce the amount of non-target eukaryotic DNA).

To generate DNA for optimizing microsatellite loci PCR conditions, O. marina cultures (400–500 mL) were grown to mid-log-phase (5000–10 000 cells mL⁻¹) and harvested by centrifugation at 4500 g for 20 min. Total genomic DNA was extracted from cell pellets using a Qiagen Genomic-tip 20/G (Qiagen, Hilden, Germany) and the standard protocol for cell cultures. DNA concentrations were quantified

| Strain ID | Sampling Location | Latitude | Longitude | Sampling Date |
|-----------|-------------------|----------|-----------|---------------|
| TM001     | Tromso, Norway    | 69°37′48″ | 18°54′36″ | 15/08/2011    |
| BOD02     | Bodo, Norway      | 67°16′33″ | 14°34′14″ | 25/06/2011    |
| BGN01     | Bergen, Norway    | 60°14′8″  | 5°11′8″   | 22/07/2009    |
| PLY01     | Plymouth, UK      | 50°21′48″ | −4°8′21″  | 25/04/2008    |
| ROS03     | Roscoff, France   | 48°43′39″ | −3°59′22″ | 17/05/2009    |
| EST02     | Estoril, Portugal | 38°42′7″  | −9°23′35″ | 09/07/2009    |
| FAR01     | Faro, Portugal    | 37°1′1″   | −7°59′36″ | 21/01/2008    |
using a Qubit dsDNA BR assay (Life Technologies, Gaithersburg, MD, USA), and samples were diluted to approximately 12.5 ng μL⁻¹ and the DNA concentrations requantified in triplicate to ensure accuracy.

MICROSATELLITE LOCI DEVELOPMENT

Microsatellite loci and corresponding primer sequences were obtained from an *O. marina* (reference strain PLY01) RNAseq data set (generated using Roche 454 sequencing – Low et al. 2011c) using QDD2 (Meglécz et al. 2010). We identified all unique loci that contained trinucleotide repeats, since this class of microsatellite loci typically suffers less from stutter banding than dinucleotide repeats (Guichoux et al. 2011) and was more frequent than tetra- or pentanucleotide repeats (data not shown). We then tested loci for PCR amplification success and for polymorphisms using an M13-tailed primer method (Schuelke 2000) and by genotyping multiple strains of *O. marina*. Loci were amplified in 10 μL reaction volumes that contained Green GoTaq® reaction buffer (pH 8.5) (Promega, Madison, WI, USA), 200 μM each dNTP, 1.5 mM MgCl₂, 10 μg BSA, 0.25 units GoTaq® DNA polymerase (Promega), 0.3 pmol of forward primer (whose 5’ end incorporated an M13 primer tail sequence), 0.3 pmol reverse primer and 0.3 pmol of an M13 primer that was 5’-labelled with either 6-FAM, NED, PET or VIC fluorophores (Life Technologies, Carlsbad, CA, USA) and approximately 25 ng target DNA. Thermal cycling conditions were as follows: 94°C 3 min, 35 × [94°C 30 s, 58°C 30 s, 72°C 45 s] and 72°C 4 min. PCR products were pooled with GENESCAN 500 LIZ (Life Technologies) and then separated by capillary electrophoresis on an AB3130xl (Life Technologies).

QUANTITATIVE MICROSATellite-GENOTYPING CALIBRATION

DNA from *O. marina* strains was mixed into a series of DNA pools to provide a combination of eight pairs of strains, each with twelve ratios of DNA (that varied between 50:1 and 1:50). The following strain combinations were tested: (i) FAR01 & BOD02; (ii) ROS03 & EST02; (iii) ROS03 & BGN01; (iv) EST02 & BOD02; (v) PLY01 & BGN01; (vi) BGN01 & TM001; (vii) BGN01 & ROS03; and (viii) TM001 & PLY01. Expected DNA ratios were calculated using the mean DNA concentrations and the proportions in which the DNAs were pooled. All test pools of DNA were PCR-amplified (n = 8 replicates) using one of four microsatellite PCRs, the forward primers were labelled with a 5’ fluorescent dye (Table 2) and the PCR used the standard two (forward and reverse) primers. Loci were amplified as 10-μL PCRs as described above [i.e. with GoTaq® DNA polymerase (Promega)], but with 0.3 pmol of forward primer (5’ labelled with either 6-FAM, NED, PET or VIC, Life Technologies) and 0.3 pmol reverse primer. Thermal cycling conditions were as follows: 95°C for 1 min, 35 × [95°C 30 s, 58°C 30 s, 72°C 45 s] and 72°C 10 min, where N represents the number of cycles (Table 2). PCR products were pooled with GENESCAN 500 LIZ size standard (Life Technologies) and separated by electrophoresis on an AB3130xl (Life Technologies).

Further calibrations were performed using DNA extracted with Chelex-100™ (Walsh, Metzger & Higuchi 1991), a fast and inexpensive DNA extraction method. Here, the target DNA was a series of mixed cultures to simulate a protist microcosm experiment: *O. marina* cultures were diluted to 1500 cells mL⁻¹ the strain pairs (i) ROS03 & EST02, (ii) EST02 & BOD02, (iii) PLY01 & BGN01 and (iv) BGN01 & TM001, then mixed in at four ratios of 1:24, 1:4, 4:1 and 24:1 to a total volume of 50 mL. Samples were centrifuged and genomic DNA extracted from the cell pellet by adding 50 μL Chelex-100™ (Bio-Rad Laboratories, Herceules, CA, USA) solution (5% w/v) and then incubating the samples at 95°C for 1 h. PCR amplification of four microsatellite loci (MS_23, MS_30, MS_37 and MS_41) was performed for each sample (and in replicate n = 8) as described above, but with 2 μL of the Chelex-100™ (Bio-Rad Laboratories)-extracted DNA as a template DNA. Genotyping was performed as described above. Expected ratios of the pair of strains were calculated from the mix ratio and a triplicate cell density count of each population, performed using a Sedgewick-Rafter (Hydro-Bios, Kiel-Holtenau, Germany) chamber.

MULTISTRAIN GENOTYPING CALIBRATION

We also tested the ability of microsatellite genotyping to quantify the relative abundance of four strains of *O. marina* from a mixed population. Stock cultures of BGN01, TM001, ROS03 and PLY01 were diluted to approximately 1000 cells mL⁻¹ and mixed in twelve combinations of four ratios (25:13:8:3) to a total volume of 50 mL, where at any given ratio for a specific clone there were three different combinations of the other clones’ specific ratio. Cells were harvested by centrifugation, and DNA was extracted using Chelex-100™ (Bio-Rad Laboratories). Microsatellite alleles at locus MS_41 were amplified by PCR and genotyped using the same methods as described for pairs of strains.

EXAMPLE COMPETITION EXPERIMENT

As proof of principle, *O. marina* strains ROS03 and BGN01 were used in a competition experiment to illustrate how AsQ-PCR microsatellite genotyping could infer the relative contribution of two clones in a mixed culture. Prior to experiments, both strains were grown at 16°C in 32 PSU ASW and fed heat-killed *E. coli* at a light intensity of approximately 80 μmol photons m⁻² and on a light/dark cycle of 14 h:10 h. An experimental population was established by diluting both strains to approximately 500 cells mL⁻¹ (i.e. total population density ~ 1000 cells mL⁻¹) with the addition of heat-killed *E. coli* as food and the antibiotics gentamycin and penicillin/streptomycin, in one large 2-L culture. To establish experimental microcosms, this 2-L culture was randomly divided into six 300-mL microcosms. These microcosms were divided among either high or low pCO₂ treatments that were aerated with either 400 or 1153 μatm CO₂/air, respectively, directly from
gas cylinders (BOC Special Products, Guildford, UK). All other culture conditions were as described above. Mixed culture samples were taken after approximately 0, 3, 6, 9 and 13 days by manually agitating the microcosms before removing 50 mL into centrifuge tubes. Of this sample, 5 mL was fixed in 2% (v/v) Lugol's iodine and the total population size enumerated with a Sedgewick-Rafter chamber in a Zeiss Axiovert A1 (Carl Zeiss, Oberkochen, Germany) inverted microscope. The remaining 45 mL was centrifuged and DNA extracted using Chelex as previously. AsQ-PCR was performed as above ($n = 4$ for each sample) at the locus MS_23 to estimate the peak height ratio for alleles corresponding to ROS03 and BGN01.

**DATA ANALYSIS**

Allele sizes and peak heights were determined using Genemapper v.3.0 (Life Technologies). These data were then transformed to a peak height ratio (PHR) – the ratio between the peak heights of the alleles generated by the two strains – which was natural log-transformed prior to analysis in all cases. Calibration curves were created using linear least-squares regression (Levenberg-Marquardt Algorithm). We also performed the same analyses for peak area ratio, though the results were almost identical (i.e. difference in $R^2$ values $\sim 0.001$). Multistrain calibration was performed using multiple regression, with one model for each strain. Additional terms were included for ratios of non-target strains, followed by stepwise removal of non-significant terms until a minimal model was found. To estimate the frequencies of ROS03 and BGN01 in competition observed PHR peak height ratios were transformed using the calibration curve provided in Fig. 1 (see Results). Clone-specific cell densities were calculated by multiplying the frequency of each clone by the total population density. All statistical analyses were performed using R v.2.15.3 (R Core Development Team, 2013).

**Fig. 1.** Calibration curves for the pairwise quantitative microsatellite-genotyping assay, each strain pair and the locus used are shown above each panel, data for calibrations performed using mixed DNA are in black and data for Chelex extracted mixed cell populations are shown in red, solid lines are regression fits and dashed line is $x = y$. © 2014 The Authors. *Methods in Ecology and Evolution* published by John Wiley & Sons Ltd on behalf of British Ecological Society., *Methods in Ecology and Evolution*, 6, 315–323
Results

We identified some 130 microsatellite loci from Lowe et al.’s (2011c) transcriptome data for *O. marina*. Out of the 38 trinucleotide microsatellite loci (for which 47 primer pairs were designed), ten loci amplified clean, single peaks for each allele and presented no PCR failures, potential null alleles or substantial stutter banding (i.e. extra PCR products that can be produced by DNA polymerase during PCR, and which differ in size from the main allele by a multiple of the repeat unit; see Selkoe & Toonen (2006) and Guichoux et al. (2011) for detailed discussions about quality of microsatellite genotyping). Out of these ten loci, four were homozygous and polymorphic in at least some of the strains tested (i.e. loci exhibited just one allele within strains, but with alleles differing among strains), and thus suitable to quantify strain abundance in mixed cultures of *O. marina*. The final panel of loci used for quantitative genotyping (referred to as MS_23, MS_31, MS_37 and MS_41) yielded 3, 4, 2 and 7 alleles, respectively, in the eight strains of *O. marina* that we tested.

AasQ-PCR microsatellite genotyping effectively recovered the known genotype frequencies from the test pools of extracted DNA (Fig. 1, Table 3). Calibration curves fit the extracted DNA data very well in most cases ($R^2 > 0.97$), with the regression slopes usually close to 1 (range $0.84–1.18$) and the intercepts mostly close to 0, albeit with some statistically significant differences from expected values (Table 3). Detection and quantification of alleles was possible at nDNA frequencies as low as 0.02 (i.e. 1:50 ratio of cell abundance). Based on analyses of residuals, we found no systematic bias in the error distributions of the fitted linear models to the data. Calibrations using DNA extracted with Chelex-100™ (Bio-Rad Laboratories) were successful (Fig. 1, Table 3), also producing values of $R^2 > 0.99$ for all four trials. In addition, data for peak height ratio matched those obtained from pooled DNA calibrations relatively well; hence, this microsatellite genotyping technique works with potential low-quality DNA samples that are obtained when using a cheap and fast extraction method.

Multistrain calibration was successful, but required multiple-regression estimated calibration coefficients that took account of the ratios of non-target alleles in the data (see Appendix S1 in the Supporting Information). Regression models fitted the data well, with $R^2$ values of 0.96, 0.93, 0.92 and 0.96 for BGN01, TMO01, ROS03 and PLY01, respectively.

In the competition experiments, strain ROS03 outperformed BGN01, irrespective of the level of $p$CO$_2$, and ‘won’ the competition in all six cases (Fig. 2). BGN01 tended to show an initial increase while the total population density was still low, but after 13 days ROS03 was at a density of around 2000 cells mL$^{-1}$. BGN01 tended to be below its initial density of 500 cells mL$^{-1}$. Overall, the microcosms displayed a high level of consistency between replicates and across treatments. The variance within samples (mean SE across all experiments 64 cells mL$^{-1}$) from the genotyping assay was well within the limits to distinguish each strain with confidence (Fig. 2), and for four out of six replicates it recovered an excellent estimate of the starting density (i.e. 500 cells mL$^{-1}$ for each strain). In the other two cases, the genotyping error was very small and the dynamics appear distinct from other replicates suggesting that the estimates of frequency reasonably reflect the strain composition in these microcosms.

Discussion

Species naturally exhibit more genetic diversity than can be examined during laboratory and field experiments. Yet, understanding the consequences of how the variation that exists within and between populations may interact is of fundamental importance to understanding numerous ecological and evolutionary processes (Bijlsma & Loeschcke 2005; Hughes et al. 2008). Here we demonstrate how AasQ-PCR microsatellite genotyping can be used to determine the relative abundance of multiple protist strains that are interacting within a mixed population, without a need for high-quality DNA extraction, by testing against known ratios of cells and from mixed strain culture experiments. Once microsatellite loci are available, therefore, this method provides a fast, cheap and easy-to-implement way of studying intraspecific dynamics of many, genetically distinct protist strains within experimental microcosms.

Microsatellite loci are good genetic markers to differentiate among strains of protists for a number of reasons. First, they are present in the genomes of nearly all known eukaryotes (Schlotterer 2004; Aishwarya, Grover & Sharma 2007). Secondly, microsatellites are often highly polymorphic (Elle gren 2004; Schlotterer 2004; Guichoux et al. 2011), which

Table 3. Regression statistics, strains used and allele information from the calibrations between 8 pairs of *Oxyrrhis marina* used in quantitative microsatellite-genotyping calibration

| Strain A  | Strain B  | Marker | A size (bp) | B size (bp) | $R^2$ | Slope | Intercept |
|----------|-----------|--------|-------------|-------------|-------|-------|-----------|
| FAR01    | BOD02     | MS_23  | 338         | 344         | 0.984 | 0.990 | –0.384   |
| ROS03    | EST02     | MS_23  | 341         | 344         | 0.997 | 1.180 | 0.174    |
| ROS03    | BGN01     | MS_23  | 341         | 344         | 0.973 | 0.933 | 0.487    |
| EST02    | BOD02     | MS_30  | 162         | 168         | 0.988 | 0.957 | 0.686    |
| PLY01    | BGN01     | MS_37  | 136         | 142         | 0.960 | 0.841 | 1.411    |
| BGN01    | TMO01     | MS_41  | 92          | 101         | 0.994 | 1.031 | 0.083    |
| BGN01    | ROS03     | MS_41  | 92          | 113         | 0.993 | 0.927 | 1.435    |
| TMO01    | PLY01     | MS_41  | 101         | 119         | 0.970 | 0.925 | 0.229    |

© 2014 The Authors. Methods in Ecology and Evolution published by John Wiley & Sons Ltd on behalf of British Ecological Society., Methods in Ecology and Evolution, 6, 315–323
makes them capable markers of intraspecific diversity; for example, studies typically find between 10 and 20 alleles at most microsatellite loci during field surveys of protist intraspecific diversity (e.g. Evans, Kühn & Hayes 2005; Harnstrom et al. 2011; Lowe et al. 2011d). Nonetheless, some pertinent issues when developing microsatellites should be noted. We targeted trinucleotide microsatellite loci, as they were predicted to suffer less from stutter banding than dinucleotide repeats (Guichoux et al. 2011). However, stutter banding is neither limited to, nor an inherent problem of, dinucleotide microsatellite loci; indeed, Reininger, Grunig & Sieber (2011) estimated fungal biomass from allele frequencies at dinucleotide microsatellite loci. Thus, other studies may use dinucleotide loci as a means of finding greater numbers of presumably more polymorphic loci. For practical purposes, we found that using peak height ratio gave slightly better estimates than peak area ratio (data not shown), but this difference was negligible (see Materials and methods).

We were able to accurately estimate the relative frequencies of multiple strains within mixed pools using a simple Chelex-100™ protocol (Walsh, Metzger & Higuchi 1991) to extract DNA from a mixed cell culture, just as well as using a more time-consuming and expensive DNA extraction kit; Chelex-100™ (Bio-Rad Laboratories) extracted DNA was used in the AsQ-PCR protocol of Meyer et al. (2006). We used the Chelex-100™ protocol (Bio-Rad Laboratories) only because it is rapid and inexpensive; however, other DNA extraction methods may be suitable for quantitative microsatellite genotyping depending on laboratory preferences. Extracting DNA directly from a mixed sample, rather than having to re-isolate and individually genotype cells (cf. Evans, Kühn & Hayes 2005; Harnstrom et al. 2011; Roger, Godhe & Gamfeldt 2012) saves time and money, thereby opening the possibility of tracking the dynamics of specific protist strains in experimental microcosms that contain high standing diversity. The next step will be to examine how AsQ-PCR can be used to differentiate among 10s of strains (using multiple markers) and then quantify the relevance of naturally high levels of protist diversity typical of natural populations. By extension, it might be informative to use this genotyping method to assess allelic diversity and differences between samples of natural populations. In addition to characterizing changes in intraspecific diversity, AsQ-PCR could be used conveniently to track changes in species’ composition when the microsatellite loci amplify alleles in different species. It is notable, however, that there are few reports of the same microsatellite loci being used successfully in different protist species or clades (see e.g. Bay, Howells & van Oppen 2009 for clades of the endosymbiotic dinoflagellate *Symbiodinium*) and thus other methods are likely to be better suited to characterize diversity within a single species.

An assumption of this methodology is that number of genome copies of the genetic marker(s) correlates with numbers of individual cells, and thus there may be problems in calibrating loci that have been duplicated (e.g. Meglecz et al. 2007) or in species where the DNA content per cell varies among individuals (Parfrey, Lahr & Katz 2008). Nonetheless, our study highlights an obvious time and cost benefit associated with using a protist model where the focus is on single cells. While there is no need to identify an appropriate reference strain for this method (Reininger, Grunig & Sieber 2011), it is advisable to
initially culture and genotype individual strains so that standard pools of DNA (or mixed cells) can be used to calibrate and validate the assay. An additional advantage is that suitably polymorphic microsatellite loci are available for many protists. If unavailable, they can be easily found by mining expressed sequence tag (EST) or other genomic data sets, of which many exist for model and non-model protist species.

With this in mind, it is important to remember that microsatellite loci are not necessarily neutral markers, with divergence among natural populations at some loci driven by selection (Vasemagi, Nilsson & Primmer 2005; Watts et al. 2008) through linkage with another locus or directly. For example, the length of the microsatellite tract affects transcription at several loci in yeast, and selection can be quite efficient in generating new alleles of different lengths (Vinces et al. 2009). This raises the presently unexplored possibility that, by targeting microsatellite loci associated with genes (e.g. in promoter regions or within ESTs), AsQ-PCR could be used to quantify selection on key cellular processes in non-model protists, on standing genetic variation and/or on new mutations. Indeed, that many microsatellites have a relatively high mutation rate (approximately 10−6/locus/cell division in yeast; Lynch et al. 2008) may impact some long experiments, as under some circumstances the actions of mutation (from ‘one strain to another’) and selection may be non-negligible. By genotyping strains at multiple, independent loci, it should be possible to identify discrepancies in ‘strain frequency’ as potential instances of mutation selection acting on a specific locus. Additionally, sex could recombine alleles resulting in de novo heterozygous strains during culture experiments, thus impacting the reliability of our genotyping assays. For many protists, including Oxyrrhis, sex is either absent or at least infrequent under standard culture conditions and can often be environmentally controlled (Ducks & Roger 1999), negating this possibility.

Microsatellite-based genotyping of DNA pools (e.g. Khatib et al. 1994; Hillel et al. 2003; Skalski et al. 2006) has largely been replaced by other methods to estimate genotype frequencies from mixed populations of cells, including qPCR melt curve analysis (Cruz et al. 2010), SNP-based pyrosequencing (Wasson et al. 2002) and FREQ-Seq (Chubiz et al. 2012). While these methods typically present high levels of accuracy and the possibility to detect alleles at a lower frequency than the microsatellite method, they all suffer the same problems when working in non-model organisms that have large and complex genomes (i.e. most protists): they all require identification of SNPs between the genotypes that are to be quantified, and finding suitable loci with limited genomic resources is both time-consuming and expensive.

One important application of this method is to better understand natural selection – the dynamic process by which a particular genotype (or allele) increases in relative frequency in a population due to higher fitness (Chevin 2011). While the outcome of selection, and hence evolution, is determined by competition between genotypes of the same species, the contribution that apparently simple parameters, such as growth rate (r) and carrying capacity (K), make to fitness and selection surprisingly remains widely untested by empirical research (Chevin 2011). Here we have shown that in a simple two-strain system, one genotype out competes the other and, in a constant environment, natural selection should favour the fittest genotypes, yet natural populations often harbour considerable genotypic diversity (Lowe et al. 2011d). Using protist microcosms and AsQ-PCR to investigate intraspecific diversity and dynamics could help to reconcile gaps in our knowledge of evolutionary ecology.

The ability to infer fitness from experimental evolution/selection experiments has wider application, especially in the fields of global change biology (e.g. Lohbeck, Riebesell & Reusch 2012), drug resistance evolution (e.g. Melnyk, Wong & Kassen 2014) and agricultural pest management (e.g. Lopes et al. 2008). Traditionally, these fields have employed prokaryote, plant or animal models, but in all three areas wider taxonomic coverage and more functionally relevant studies should encompass a greater emphasis on protists. This is especially true when considering marine planktonic ecosystems, where single-celled eukaryotes account for the majority of primary productivity and play a considerable role in the major biogeochemical cycles. Protists exhibit substantial functional and genetic diversity (Kremp et al. 2012), but understanding the relevance of this diversity in mixed populations remains a challenge. Using microsatellites to follow the success and demise of pre-defined genotypes (strains) under competition will allow us to make inferences about how populations with high standing genetic variation may adapt to future environmental change, and thus shape the future biosphere.

Acknowledgements

EJAM was supported by a NERC research studentship (NE/H025472/2) as part of the UK Ocean Acidification Research Programme, cofunded by the Natural Environment Research Council (NERC), the Department for Environment, Food and Rural Affairs (DEFRA) and the Department of Energy and Climate Change (DECC). We also thank two anonymous reviewers, whose suggestions improved the content of this manuscript. The authors declare no conflict of interest.

Data accessibility

Data available from the Dryad Digital Repository: doi:10.5061/dryad.de7s2.

References

Aishwarya, V., Grover, A. & Sharma, P.C. (2007) EuMicroSatIdb: a database for microsatellites in the sequenced genomes of eukaryotes. BMC Genomics, 8, 225.

Balmer, O., Stearns, S., Schotzau, A. & Brun, R. (2009) Intraspecific competition between co-infecting parasite strains enhances host survival in African trypanosomes. Ecology, 90, 3367–3378.

Bay, L.K., Howells, E.J. & van Oppen, M.J.H. (2009) Isolation, characterisation and cross amplification of thirteen microsatellite loci for coral endosymbiotic dinoflagellates (Symbiodinium clade C). Conservation Genetics Resources, 1, 199–203.

Bijlama, R. & Loseshcke, V. (2005) Environmental stress, adaptation and evolution: an overview. Journal of Evolutionary Biology, 18, 744–749.

Bolick, D.I., Amarasekare, P., Araujo, M.S., Burger, R., Levine, J.M., Novak, M. et al. (2011) Why intraspecific trait variation matters in community ecology. Trends in Ecology and Evolution, 26, 183–192.

Chevin, L.M. (2011) On measuring selection in experimental evolution. Biology Letters, 7, 210–213.

© 2014 The Authors. Methods in Ecology and Evolution published by John Wiley & Sons Ltd on behalf of British Ecological Society, Methods in Ecology and Evolution, 6, 315–323
Wasson, J., Skolnick, G., Love-Gregory, L. & Permutt, A. (2002) Assessing allele frequencies of single nucleotide polymorphisms in DNA pools by pyrosequencing technology. *BioTechniques*, **32**, 1144–1152.

Watts, P.C., O’Leary, D., Cross, M.C., Coughlan, J., Dillane, E., Kay, S.M. *et al.* (2008) Contrasting levels of genetic differentiation among putative neutral microsatellite loci in Atlantic herring *Clupea harengus* populations and the implications for assessing stock structure. *Hydrobiologia*, **606**, 27–33.

Weisse, T. (2002) The significance of inter- and intraspecific variation in bacti- rous and herbivorous protists. *Anthonie van Leeuwenhoek*, **81**, 327–341.

Yang, Z., Lowe, C.D., Crowther, W., Fenton, A., Watts, P.C. & Montagnes, D.J. (2013) Strain-specific functional and numerical responses are required to evaluate impacts on predator-prey dynamics. *ISME Journal*, **7**, 405–416.

Received 10 September 2014; accepted 18 November 2014

Handling Editor: M. Gilbert

Supporting Information

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

Supporting Information

Appendix S1. Multiple Regression results.

Appendix S2. Method flow chart.