Published in final edited form as:

Mol Biochem Parasitol. 2009 July ; 166(1): 81–84. doi:10.1016/j.molbiopara.2009.02.010.

EFFICIENT GENOTYPING OF *SCHISTOSOMA MANSONI* MIRACIDIA FOLLOWING WHOLE GENOME AMPLIFICATION

Claudia L. L. Valentim^{1,2}, Philip T. LoVerde², Timothy J. C. Anderson¹, and Charles D. Criscione³

¹ Department of Genetics, Southwest Foundation for Biomedical Research. P.O. Box 760549, San Antonio, Texas, 78245, USA

² Departments of Biochemistry and Pathology, Mail Code 7760, University of Texas Health Science Center, San Antonio, Texas 78229, USA

³ Department of Biology, 3258 TAMU, Texas A&M University, College Station, TX 77843, USA

Abstract

Small parasites and larval stages pose a problem for molecular analyses because limited amounts of DNA template are available. Isothermal methods for faithfully copying DNA have the potential to revolutionize studies of such organisms. We evaluated the fidelity of multiple displacement amplification (MDA) for amplifying DNA extracted from a single miracidium of *Schistosoma mansoni*. To do this we genotyped DNA extracted from 28 F1 miracidia following MDA using 56 microsatellite markers. Because these miracidia were obtained from a cross between a male and female worm of known genotypes, we were able to predict the alleles present in the progeny and quantify the genotyping error rate. We found just 8/1568 genotypes deviated from Mendelian expectations. Furthermore, because one of these resulted from a genuine mutation, the error rate due to MDA is 7/1568 (0.45%). We conclude that many hundreds of microsatellites or other genetic markers can be accurately genotyped from a single miracidium using this method, greatly expanding the scope of population genetic, epidemiological and evolutionary studies on this parasite.

The genome sequence and genetic map of *Schistosoma mansoni* [1;2] open the door to genome-wide analyses involving hundreds or thousands of genetic markers. However, adult worms live in the mesenteric veins of human hosts. Thus, only larval stages, from which limited amounts of DNA can be extracted, are available for genetic analyses when working with samples from humans. Researchers have used two approaches to get around this problem. First, eggs obtained from human hosts can be hatched in the laboratory and the miracidia used to infect the appropriate freshwater snail. Clonally produced cercariae shed from snails or adult worms generated from passage through rodents can then be used for genetic analysis. This approach is cumbersome and time consuming, and has the additional problems of creating genetic bottlenecks and possibly causing host (snail or rodent) induced selection at pertinent loci [3]. Hence, the parasites genotyped may be poorly representative of those initially sampled. Second, PCR based methods have been used to amplify and genotype DNA extracted from a single egg or miracidium [4–7]. Using this approach up to 21 loci have been genotyped from a single miracidium [6]. However, PCR genotyping of miracidia using this approach requires careful optimization and pooled PCR reactions. Observations of strong deviations from Hardy-

Weinberg proportions in the genotypes generated may indicate that template is limiting or of poor quality and suggest caution is needed when using these methods [4]. Furthermore, these methods [4,5] do not leave enough template DNA for repeat genotyping, which precludes assessment of genotyping error rates in field studies.

Isothermal amplification of template DNA using Phi-29 DNA polymerase provides a simple approach for generating microgram amounts of DNA template from a few ng of starting material [8]. This method is now widely used for copying DNA for use in human genetic studies and has been carefully validated for use in a variety of organisms [9]. This approach has also been used to generate large amounts of DNA template from single sperm [10] and from a variety of parasite species, such as *Plasmodium sp.* [11] and *Trypanosoma sp.* [12]. We have evaluated the use of Phi-29 to faithfully replicate DNA extracted from a single miracidium. To measure the reliability of this approach we genotyped F1 miracidia generated from a genetic cross between parental worms in which genotypes were known. Therefore, we were able to predict the expected genotypes at each locus genotyped and quantify genotyping error rates. We observed just 7 discrepant genotypes due to whole genome amplification among 28 miracidia genotyped at 56 microsatellites markers (1568 genotypes) giving an error rate of 0.45%. While this validation experiment is specific to *S. mansoni*, we expect that similar success will be found for other schistosome species and for helminth parasites for which only eggs or larval stages are easily accessible.

We set up a genetic cross between single genotypes of a female schistosome from the NMRI (Puerto Rico) strain with a male schistosome from the LE (Brazil) strain to generate the F1 progeny. To do this we exposed snails (*Biomphalaria glabrata*) to a single miracidium. After 28–29 days, the cercariae were collected by exposure to light. We extracted DNA from 20–50 cercariae from each shedding snail in 50 µL of 6% chelex containing 0.2 mg/mL of proteinase K. This was incubated at 56°C for 2 hours and at 100 °C for 8 minutes. We sexed cercariae by amplifying cercarial DNA using the W1 primers, which are specific for a repetitive region in the W chromosome of female worms [13]. As male worms are not expected to show any amplification, we ran a control PCR with the autosomal microsatellite locus sc18 (Forward CAGTTTGTCACAAGTAACGAG and reverse AAAAGACAAATTCGTTGTGG) in order to confirm the DNA was extracted. PCR was performed with 15 µL reactions containing 2.4 µL of extraction supernatant, 1X PCR buffer, 1.5 mM MgCl₂, 0.2 mM each of dNTP, 0.4 µM each primer, and 0.75 units (0.15 µL) Taq DNA polymerase (Takara Shuzo Co., Otsu, Shiga, Japan). PCR cycling was 95°C for 3 min, once; 94°C for 45 sec, 54°C for 30 sec, 72°C for 45 sec, 35 times; 72°C for 7 min, once. We then infected a hamster with LE male cercariae of a single clonal genotype and NMRI female cercariae of a single clonal genotype to set up the cross. After 45 days, the hamster was perfused and the parental worms collected. We removed the liver, hatched the eggs in freshwater, and collected single miracidia in 10µL H₂O in 0.5mL microcentrifuge tubes for DNA extraction.

We extracted DNA from individual adult worms and 29 F1 miracidia in 200 µL and 20 µL of 5% chelex, respectively, as described above. We then used Phi-29 polymerase to expand the DNA template available using the GenomiPhi v2 DNA Amplification Kit (GE Healthcare) according to the manufacturer's manual, using 2.5µL of the chelex extracted DNA in each reaction. The total reaction volume of 50 µL was incubated at 30°C for 4 hours and terminated at 65°C for 10 minutes. After amplification, the whole genome amplified DNA was purified through Sigma Spin Post Reaction Purification columns (Sigma, St Louis, Missouri). The final volume of the DNA product was adjusted to 150 µL with water. DNA concentration was determined using a Hoefer DyNS Quant 200 Fluorometer and diluted to a final concentration of 15 ng/µL. Two µL (30 ng) of product were used as template for genotyping of each microsatellite marker.

We used an economical method for fluorescent labeling of microsatellites [14], in which oligos are unlabeled, but attached to the 5' end of the forward primer is the M13 oligo (5'-TGTAACGACGGCCAGT-3'). The PCR cocktail contains a fluorescent labeled M13 primer, the reverse primer, and the forward primer with an M13 tail. The concept underlying this approach is simple: the first rounds create the specific product and then the last rounds label the product. This method allows one to order a single labeled primer (i.e., M13) that can be used to label multiple loci. We also "pig-tailed" the reverse primers by adding GTTCTT to the 5' ends [13] to minimize stutter bands. Amplifications were performed in 5 µL reactions containing 1X buffer, 1.5 mM of MgCl₂, 0.2 mM each of dNTP, 0.16 µM of fluorescent label M13(-21), 0.08 µM of the forward primer, 0.16 µM of the reverse primer, 0.15 units (0.03 µL) of Takara Taq (Takara Shuzo Co., Otsu, Shiga, Japan) and 30 ng of DNA template. For each reaction, the labeled M13(-21) primer was fluorescently marked with 6-FAM, VIC, NED, or PET. PCR cycling was as follows: 94°C (5 min), 30 cycles at 94°C (30 s)/56°C (45 s)/65°C (45 s), followed by 8 cycles 94°C (30 s)/53°C (45 s)/65°C (45 s), and 65°C for 10 min. The PCR products were pooled in groups of four and analyzed on an ABI 3130xl genetic automated analyzer with GENESCAN software (Applied Biosystems). Allele sizes were calculated using GENOTYPER software (Applied Biosystems) and visually confirmed.

For this study, we used 56 microsatellite markers distributed across the genome of the parasite (Supplementary Table 1). These markers have known chromosomal locations and are a subset of 243 microsatellite markers used to construct a genetic linkage map for *S. mansoni* [2]. Forty-eight of the loci have autosomal inheritance (including 8 pseudoautosomal markers on the Z and W chromosomes) while 8 loci were located in the Z-specific region, which does not recombine with the W chromosome [2]. When females are genotyped for these Z-specific loci, only the allele inherited from the male parent is present. Fifty-one loci were fully informative because the parental parasites did not share alleles. Thus, we could directly assess the genotyping error rate at these loci because at an autosomal locus, F1 offspring should have 2 alleles. Five loci were partially informative because the parents shared one of the alleles.

The final data set included parental genotypes for the 56 loci as well as for the 28 F1 miracidia, giving a total of 1568 F1 genotypes. We were able to chelex extract DNA from all 29 miracidia attempted. However, one sample contained 2 miracidia or was contaminated, because more than 2 alleles were seen at multiple loci. This sample was not included in the study. The results are summarized in Table 1. We observed a single allele from each of the two parents (or for Z-specific loci, a single allele from the male parent in F1 female offspring) as expected for Mendelian segregation for 1560/1568 genotypes, giving an overall success rate of 99.49% (see Fig. 1 for examples). This error rate is similar to that observed in studies in which the same individuals are re-genotyped to measure reproducibility [15], giving us considerable confidence in this method. Just 8/1568 (0.51%) of genotypes showed patterns that differed from those expected under Mendelian segregation. Of these 8, discordant genotypes resulting from generation of additional non-parental alleles were observed in 4/1568 genotypes (Fig. 2a). In these cases the expected parental alleles were present, but a third non-parental allele was also observed. Discordant genotypes resulting from failure to genotype one of the parental alleles were observed in 4/1568 cases. In this situation one of the parental alleles was present but the other allele was shifted by 2 bp (2 case) (Fig. 2b) or absent (2 cases). These 8 discordant genotypes could have resulted from natural mutation events (e.g., occurring during clonal expansion of sporocysts within the snail host). Alternatively they may represent mutations occurring during MDA or slippage during the PCR reaction. To differentiate between these explanations, we repeated PCR reactions using miracidial DNA pre-MDA (i.e., a PCR using the chelex extracted DNA). In 7 cases, the discordant patterns were not observed when using extracted miracidial DNA as template for PCR. Hence, in a small number of genotypes (0.45% [95% confidence intervals: 0.22–0.92%]) MDA induced artifactual genotypes. We note that in about 6% of the genotypes the shorter allele had a lower peak height than the larger allele.

These genotypes were not restricted to a subset of miracidia but were randomly distributed among individuals. This pattern appeared to result from the MDA because a repeat of the MDA from a few of these samples reversed this pattern (data not shown). A lower shorter allele could be problematic for scoring loci with a large degree of stutter. However, we feel this is not an insurmountable problem as there are methods that can reduce or eliminate microsatellite stuttering (e.g., using a polymerase that is highly processive). One could also repeat MDA on a subset of samples in order to determine if lower shorter alleles result in genotyping errors. Interestingly, one of the eight Mendelian discrepancies resulted from a genuine mutation in the miracidia genotyped. In this case both chelex extracted and WGA DNA generated genotypes with a 2bp shift relative to the allele from the LE parent (Fig. 2c). This gives an estimated mutation rate of 3.7×10^{-4} [95% confidence intervals: $1.0 \times 10^{-4} - 2.2 \times 10^{-3}$] per generation for autosomal loci [$1/(48 \text{ autosomal loci} \times 28 \text{ miracidia} \times 2 \text{ chromosomes})$], similar to that observed in humans and maize, but higher than in *Drosophila* [16]. This mutation was not included in the overall error rate, because the Mendelian discrepancy observed was not due to WGA.

Only 2 markers (sc290, and sc19; $p < 0.01$) showed a deviation from expected Mendelian ratios. This was tested by comparing observed and expected numbers of genotypes using chi-squared tests. Both markers were fully informative and all offspring had 2 alleles. Thus, these deviations are not the result of MDA. Interestingly, sc290 and sc19 are only 2.4 cM apart in the linkage map [2]. It is not uncommon to find genomic regions with segregation distortion when crossing diverged populations due to the evolution of coevolved gene complexes or of incompatible regions [16].

These results demonstrate that MDA can generate large amounts of DNA from single miracidia with high fidelity. In this work we used the high yield GenomePhi kit (GenomePhi HY DNA Amplification Kit) giving an average (± 1 s.d.) yield of $53.5\mu\text{g} \pm 3.6\mu\text{g}$. Thus, 2.5 μL of chelex extracted miracidium can provide enough DNA to run up to 1500 microsatellite genotyping reactions (assuming 30 ng of template/reaction). Hence, up to 12000 reactions could be generated from the 20 μL chelex extraction. The total amount of DNA available from a single MDA is also sufficient for SNP typing assays (e.g., Illumina Golden Gate) and next-generation sequencing using platforms such as the Illumina Genome analyzer (Solexa). An advantage of generating large amounts of DNA is that genotypes that fail to amplify at first attempt can be retried. In addition, samples can be regenotyped to provide reproducibility estimates in field-based studies. MDA currently costs between \$4 and \$16 depending on whether low or high yield is required. For most applications low yield kits generate ample DNA (approximately 5 μg). Hence the method is relatively inexpensive.

The *Schistosoma mansoni* genome sequence (v3.0 of the sequence assembly) contains approximately 16,757 di-nucleotide repeats (≥ 8 perfect repeats) and 4,976 tri (≥ 7 perfect repeats) and 1,422 tetranucleotide repeats (≥ 6 perfect repeats). Some of these have been previously used for population studies of *S. mansoni*, while 243 markers have recently been used to generate a genetic map for this organism [2]. Furthermore, multiple SNP polymorphisms are present in the genome sequence data and from examination of EST sequences [17]. MDA gives researchers the ability to genotype many hundreds of microsatellite or SNPs from single miracidia. Hence, genome-wide studies of genetic variation are now feasible. We anticipate that genotyping of miracidia after WGA will allow detailed studies of parasite population genetics at a variety of spatial scales, improving our understanding of mating systems and transmission dynamics of this parasite. Genome-wide genotyping of miracidia could also be used to systematically search for genome regions that have been exposed to strong selection, for example by drug treatment, and for genome-wide association studies of schistosome traits of biomedical importance.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Supported by NIH R21 AI072704 (TJCA), SFBR Forum Grant (CDC), NIH Training Grant D43TW006580 (PTL), and NIH schistosome supply grant AI30026. The molecular work at the Southwest Foundation for Biomedical Research was conducted in facilities constructed with support from Research Facilities Improvement Program Grant C06 RR013556 from the National Center for Research Resources, National Institutes of Health. We thank the following: the Wellcome Trust Sanger Institute for providing the genome sequence and repeat masked sequence; Matt Berriman and Najib El-Sayed for genome support; and Guilherme Oliveira for supplying the LE line.

References

1. LoVerde PT, Hirai H, Merrick JM, Lee NH, El Sayed N. Schistosoma mansoni genome project: an update. *Parasitol Int* 2004;53:183–92. [PubMed: 15081950]
2. Criscione CC, Valentim CLL, Hirai H, LoVerde PT, Anderson TJ. A genomic linkage map of the human blood fluke *Schistosoma mansoni*. 2009Submitted
3. LoVerde PT, DeWald J, Minchella DJ, Bosshardt SC, Damian RT. Evidence for host-induced selection in *Schistosoma mansoni*. *J Parasitol* 1985;71:297–301. [PubMed: 4009348]
4. Sorensen RE, Rodrigues NB, Oliveira G, Romanha AJ, Minchella DJ. Genetic filtering and optimal sampling of *Schistosoma mansoni* populations. *Parasitology* 2006;133:443–51. [PubMed: 16817994]
5. Gower CM, Shrivastava J, Lamberton PH, et al. Development and application of an ethically and epidemiologically advantageous assay for the multi-locus microsatellite analysis of *Schistosoma mansoni*. *Parasitology* 2007;134:523–36. [PubMed: 17096873]
6. Steinauer ML, Agola LE, Mwangi IN, Mkoji GM, Loker ES. Molecular epidemiology of *Schistosoma mansoni*: a robust, high-throughput method to assess multiple microsatellite markers from individual miracidia. *Infect Genet Evol* 2008;8:68–73. [PubMed: 18329981]
7. Beltran S, Galinier R, Allienne JF, Boissier J. Cheap, rapid and efficient DNA extraction method to perform multilocus microsatellite genotyping on all *Schistosoma mansoni* stages. *Mem Inst Oswaldo Cruz* 2008;103:501–3. [PubMed: 18797767]
8. Dean FB, Hosono S, Fang L, et al. Comprehensive human genome amplification using multiple displacement amplification. *Proc Natl Acad Sci USA* 2002;99:5261–6. [PubMed: 11959976]
9. Lovmar L, Syvanen AC. Multiple displacement amplification to create a long-lasting source of DNA for genetic studies. *Hum Mutat* 2006;27:603–14. [PubMed: 16786504]
10. Spits C, Le Caignec C, De Rycke M, et al. Optimization and evaluation of single-cell whole-genome multiple displacement amplification. *Hum Mutat* 2006;27:496–503. [PubMed: 16619243]
11. Wang Y, Nair S, Nosten F, Anderson T. Multiple Displacement Amplification for Malaria Parasite DNA. *J Parasitol* 2008;1.
12. Morrison LJ, McCormack G, Sweeney L, et al. Use of multiple displacement amplification to increase the detection and genotyping of trypanosoma species samples immobilized on FTA filters. *Am J Trop Med Hyg* 2007;76:1132–7. [PubMed: 17556624]
13. Gasser RB, Morahan G, Mitchell GF. Sexing single larval stages of *Schistosoma mansoni* by polymerase chain reaction. *Mol Biochem Parasitol* 1991;47:255–8. [PubMed: 1944422]
14. Schuelke M. An economic method for the fluorescent labeling of PCR fragments. *Nat Biotechnol* 2000;18:233–4. [PubMed: 10657137]
15. Hoffman JI, Amos W. Microsatellite genotyping errors: detection approaches, common sources and consequences for paternal exclusion. *Mol Ecol* 2005;14:599–612. [PubMed: 15660949]
16. Vigouroux Y, Jaqueth JS, Matsuoka Y, et al. Rate and pattern of mutation at microsatellite loci in maize. *Mol Biol Evol* 2002;19:1251–60. [PubMed: 12140237]
17. Simoes M, Bahia D, Zerlotini A, et al. Single nucleotide polymorphisms identification in expressed genes of *Schistosoma mansoni*. *Mol Biochem Parasitol* 2007;154:134–40. [PubMed: 17568698]

Figure 1. Examples of genotypes consistent with mendelian segregation

The traces generated on GENOTYPER show the alleles present in the two parents followed by the genotype observed in a miracidium. Such patterns were observed in the vast majority of genotypes. The boxes show the estimated allele size relative to an internal size standard. (a) Fully informative autosomal locus (Locus sc239). Here the parental parasites do not share any alleles and the miracidium inherits one allele from each parent. (b) Partially informative autosomal locus (Locus sc349). Here the parental parasites have one allele in common and the miracidia genotypes are consistent with one allele from each parent. (c) Z-linked locus (Locus sc16b)). In this case female miracidium (ZW) inherits a single allele from the male parent, while males (ZZ) (not shown) inherit an allele from each parent.

Figure 2. Examples of genotypes inconsistent with Mendelian segregation

A small number of genotypes (0.45%) showed discrepant patterns of segregation as illustrated below. (a) Additional non parental alleles (Locus sc40, Miracidium 21). The two parents are homozygous for alternative alleles. However an additional non-parental allele occurs following WGA. This additional allele was not present in genotypes generated from chelex extracted material prior to WGA. (b) Missing allele (Locus sc168, Miracidium 14). The two parents are homozygous for alternative alleles. However, while the NMRI allele is present in the miracidium, the LE allele has been shifted by 2bp. The expected LE allele is observed in chelex extracted miracidium prior to WGA. (c) Genuine mutation (Locus sc453, miracidium 24). Both parents are heterozygous. While the NMRI allele is present in the F1 miracidium the LE allele

has been shifted by 2bp. The shift occurs in miracidial DNA both before and after WGA demonstrating that this is a genuine mutation.

Table 1**Summary of results**

The 1568 genotypes were characterized as being either expected or discrepant with respect to Mendelian segregation. See Figures 1 and 2 for illustrative examples in each category. The italicized categories show different types of discrepant genotypes.

Category	No. of genotypes	%
Expected	1561	99.55
Discrepant	8	0.51
<i>Additional allele</i>	4	0.26
<i>Missing allele</i>	3	0.19
<i>Mutation</i>	1	0.06
		%