Ministério da Saúde
Fundação Oswaldo Cruz
Instituto René Rachou
Programa de Pós-graduação em Ciências da Saúde

Caracterização genética de populações de campo de *Schistosoma mansoni* com o uso de microssatélites.

Nilton Barnabé Rodrigues

Belo Horizonte

Julho/2007

Ministério da Saúde
Fundação Oswaldo Cruz
Instituto René Rachou
Programa de Pós-graduação em Ciências da Saúde

Caracterização genética de populações de campo de *Schistosoma mansoni* com o uso de microssatélites.

por

Nilton Barnabé Rodrigues

Tese apresentada com vistas à obtenção do Título de Doutor em Ciências da Saúde na área de concentração de Biologia Celular e Molecular.
Orientação: Dr. Guilherme Correa de Oliveira
Orientação: Dr. Alvaro José Romanha

Belo Horizonte

Julho/2007

Catalogação-na-fonte
Rede de Bibliotecas da FIOCRUZ
Biblioteca do CPqRR
Segemar Oliveira Magalhães CRB/6 1975

R696c 2007

Rodrigues, Nilton Barnabé.

Caracterização genética de populações de campo de *Schistosoma mansoni* com o uso de microssatélites / Nilton Barnabé Rodrigues. – Belo Horizonte, 2007.

xiii, 128 f.: il.: 210 x 297mm.

Bibliografia: f. 99 - 110

Tese para obtenção do título de Doutor em Ciências pelo Programa de Pós - Graduação em Ciências da Saúde do Centro de Pesquisas René Rachou. Área de concentração: Biologia Celular e Molecular.

1. Microssatélites marcadores 2. Genética Populacional 3. *Schistosoma mansoni* I. Título. II. Oliveira, Guilherme Correa (Orientação). III. Romanha, Álvaro José (Co-orientação)

CDD – 22. ed. – 614.553

Dedico este trabalho à minha família
e aos meus amigos, que sabem, tanto
quanto eu, quantas barras tive de carregar,
e sabem, cada um deles, quanto é que
colaboraram nesta caminhada.

A vocês, muito abrigado.

Agradecimentos

Em primeiro lugar agradeço aos meus orientadores, Dr. Guilherme Corrêa de Oliveira e Dr. Alvaro José Romanha pela amizade, disponibilidade, ensinamentos e principalmente pela coragem e confiança que demonstraram me permitindo realizar este trabalho.

Às instituições financeiras, Organização Mundial da Saúde, FIRCA, FAPEMIG, FIOCRUZ, que possibilitaram a realização deste trabalho.

Aos pacientes das áreas estudadas, pessoas que em sua doença e sofrimento, me possibilitaram ter material para realizar um trabalho, o qual espero um dia, possa vir a colaborar, para pelo menos a minimização destes males, muito abrigado.

Bem, depois de tantos anos por aqui, agradeço à instituição, ao Instituto René Rachou, nas pessoas de seus diretores e chefes de laboratórios que sempre tiveram suas portas abertas para mim.

Agradeço a todos aqueles com quem tive oportunidade de conviver em bons e maus momentos (a grande maioria, sem dúvida, bons momentos).

Não vou me arriscar a agradecer nominalmente a nenhum deles, pois, correria o risco de ser injusto, e mesmo indelicado, ao priorizar algum ou ao esquecer outros.

Todos aqueles que estavam por aqui quando cheguei, passaram por aqui durante este tempo em que aqui estive e os que ainda estão, de alguma forma contribuíram para que as coisas tenham chegado onde estão agora, desta forma, agradeço a todos com o meu mais sincero muito obrigado.

Sumário	Página

	Lista de Abreviaturas e Símbolos
µg	Microgramas = 10^{-6} gramas
µl	Microlitros = 10^{-6} litros
µM	Micromolar = 10^{-6} Molar
ALF	Automated Laser Fluorescence
AMOVA	Analysis of Molecular Variance
BAC	Bacterial Artificial Chromosome
CPqRR	Centro de Pesquisas René Rachou
DALYs	Disability Adjusted Life Years
DNA	Ácido Desoxirribonucléico
dNTP	Desoxirribonucleotídeos Tri Fosfato
EDTA	Ácido etilenodiaminotetracético
EST	Expressed Sequence Tag
FAPEMIG/MCT/CNPq	Fundação de Amparo à Pesquisa do estado de Minas Gerais/Ministério de Ciência e Tecnologia/Conselho Nacional de Desenvolvimento Científico e Tecnológico
FAPESP/MCT/CNPq	Fundação de Amparo à Pesquisa do estado de São Paulo/Ministério de Ciência e Tecnologia/Conselho Nacional de Desenvolvimento Científico e Tecnológico
FCT	Índice de Fixação que indica a correlação de pares aleatórios de haplótipos de um grupo de infrapopulações com pares aleatórios de haplótipos da população total.
FIOCRUZ	Fundação Oswaldo Cruz
FIS	Índice de Fixação que indica a correlação de haplótipos de cada indivíduo com a média de cada infrapopulação.
FIT	Índice de Fixação que indica a correlaciona de pares aleatórios de haplótipos de cada indivíduo com pares aleatórios de haplótipos dentro da população total
fmol	Fentomoles = 10^{-15} moles
FSC	Índice de Fixação que indica a correlação de pares aleatórios de haplótipos de uma infrapopulação com pares aleatórios de haplótipos dentro da media dos grupos de populações
FST	Índice de Fixação que indica a correlação de pares aleatórios de haplótipos de uma infrapopulação com pares aleatórios de haplótipos dentro da população total
GSS	Genome Survey Sequences
He	Heterozigosidade Esperada
Ho	Heterozigosidade Observada
Kb	Kilobases = 10^{3} pares de bases
M	Molar
Mb	Megabases = 10^{6} pares de bases
MG	Minas Gerais
min	Minutos
ml	Mililitros = 10^{-3} litros
mM	Milimolar = 10^{-3} Molar

mRNA	RNA mensageiro
mtDNA	DNA mitocondrial
ng	Nanogramas = 10^{-9} gramas
nt	Nucleotídeos
OMS	Organização Mundial de Saúde
ORESTES	Open reading frames EST sequences
ORF	Open reading frames
p/v	Peso por volume
pb	Pares de base
PCR	Polymerase Chain Reaction
PCR-multiplex	PCR em múltiplos *loci*
pg	Picogramas = 10^{-12} gramas
pH	Potencial hidrogeniônico
pmol	Picomoles = 10^{-12} moles
PRE	Polymorphic Repetitive Element
PVPP	polivinilpolipirrolidone
RAPD	Random Amplified Polymorphic DNA
rDNA	DNA ribosomal
RFLP	Restriction Fragment Lenght Polymorphism.
RNA	Ácido ribonucléico
ROSE	Rapid One Step Extraction
rRNA	RNA ribossômico
R_{ST}	Índice de Fixação análogo ao F_{ST}, porém utilizado para microssatélites.
S	Coeficiente de sedimentação
SAGE	Serial Analysis of Gene Expression
SDS	Duodecil sulfato de sódio
seg	Segundos
SmAE	*Schistosoma mansoni* assembled ESTs
SNP	Single Nucleotide Polymorphism
SRS	Simple Repetitive Sequences
SSR	Simple Sequence Repeats
STR	Short Tandem Repeats Sequences
Taq	*Thermus aquaticus*
TBE	Tris Borato EDTA
TDR	The Special Programme for Research and Training in Tropical Diseases
TE	Tris EDTA
TIGR	The Institute for Genomic Reasearch
Tm	Temperatura média de anelamento dos iniciadores
Tris	Tri-hidroximetil amino metano
U	Unidade
UPGMA	Unweighet Pair Group Method Analysis

v/v	Volume por volume
VNTR	Variable Number of Tandem Repeats
WGS	Whole Genome Shotgun
WHO	World Health Organization

RESUMO

Apresentamos a utilização de 4 bibliotecas genômicas enriquecidas, como fonte de microssatélites, para o estudo da estrutura genética de populações de *Schistosoma mansoni*. De 382 seqüências obtidas, 250 (65,4%) apresentaram *loci* de microssatélites. Onze destes *loci* foram polimórficos quando testados em 100 vermes, com 2 a 19 alelos por *loci*. A heterozigosidade esperada (He) foi de 0,79 e a observada (Ho) de 0,59. Apresentamos ainda um *locus* de minissatélites, com uma porção interna variável composta por um microssatélite "CA". Este minissatélite e outros 3 *loci* de microssatélites, anteriormente descritos, mostraram sucesso na diferenciação de cepas de *S. mansoni* e de 9 diferentes espécies do gênero *Schistosoma*. Apresentamos também, a utilização de ovos individualizados de *S. mansoni* como fonte de DNA para PCR, um novo protocolo, que utiliza resina de troca iônica, para extração do DNA e a utilização de PCR-multiplex na genotipagem de ovos e vermes adultos de *S. mansoni*. O número de alelos por *locus* não diferiu entre ovos e vermes e observamos ainda uma redução no número de genótipos homozigotos nos vermes, em relação aos ovos. Tanto os ovos quanto os vermes utilizados foram provenientes das amostras de fezes coletadas de moradores da área endêmica de Virgem das Graças – Minas Gerais. Analisamos ainda, entre 10 e 22 ovos, de outras 53 populações de parasitos, da mesma área, para a determinação de sua variabilidade genética. Para se testar uma possível estruturação geográfica destas populações, estas foram divididas em 5 grupos, de acordo com sua origem geográfica dentro do município. Destas populações, 33 foram coletadas de amostras de fezes pré-tratamento e 20, pós-tratamento quimioterápico. Nove populações coletados antes e depois do tratamento foram comparadas para o estudo da influência da quimioterapia sobre variabilidade genética. Foram utilizados nestas análises 5 *loci* de microssatélites por apresentarem resultados consistentes com DNA de ovos e em PCR-multiplex. Entre as 53 populações observamos que o número de alelos por *locus* variou de 2 a 13 e não houve correlação entre variações no número de alelos por *locus* e o tratamento. A heterozigozidade observada variou de 0,0 a 1,0. Com exceção de 13 populações, Ho foi sempre menor que He. Em 39 das 53 populações, em pelo menos um *locus*, Ho foi significativamente diferente de He ($p<0,05$) representando desvios do equilíbrio de Hardy-Weinberg. Não se observou nenhum tipo de estruturação geográfica destas populações quando comparadas par a par, em uma ou ambas as coletas, ou quando comparadas pré e pós-tratamento quimioterápico. Este trabalho abre perspectivas promissoras no estudo e no entendimento de vários mecanismos biológicos envolvidos nas interações parasito-hospedeiro, na patogenia e epidemiologia e até numa futura quimioterapia da esquistossomose, doença que ainda hoje afeta, de maneira séria, e em certos casos, fatal, cerca de 200 milhões de pessoas em mais de 70 paises.

ABSTRACT

We describe the development and use of genomic microsatellite markers for *Schistosoma mansoni* genetic studies. Microsatellites were developed from microsatellite enriched genomic libraries. We obtained sequence data from 382 clones and detected microsatellites in 250 (65.4%) of the sequences. Eleven out of these *loci* were polymorphic and presented between 2 to 9 alleles per *loci* when tested in 100 individual worms. The average values of expected (He) and observed (Ho) heterozygosities were 0.79 and 0.59, respectively. We also identified a minisatellite *locus*, which contained an internal CA repeat. This minisatellite together with 3 previously described microsatellites *loci*, showed success in the differentiation of *S. mansoni* strains and when applied in the differentiation of 9 *Schistosoma* species. We also developed a new DNA extraction protocol that uses *S. mansoni* eggs as source of DNA. The genotyping reactions were conducted in a multiplex format. A comparison between the use of adult worms or egg DNA indicated that allele number per *locus* was the same. We also observed a reduction in homozygous genotypes in worms relative to the eggs. To investigate the genetics of a parasite population from an endemic site we obtained biological material from the village of Virgem das Graças, Minas Gerais. Ten to 22 eggs from other 53 parasite populations from the same area were analyzed in order to determine its genetic variability. The 53 populations were divided into 5 groups in accordance with their geographic origin in the village. Thirty-three out of these populations came from samples collected before chemotherapy treatment and 20 after the treatment. In order to evaluate the effects of chemotherapy in the genetic variability, 9 populations collected two times, before and after treatment, were compared. Five microsatellite markers, that present the consistent results in eggs and in multiplex-PCR, were selected to use in these analyses. Among the 53 populations, allele number per *locus* varied from 2 to 13 and no correlation could be established between this variation and the chemoterapic treatment. Observed heterozygosity varied from 0.0 a 1.0. In all but 13 populations Ho was always lower than He. In 39 out of 53 populations, Ho was significantly different from He for at least 1 *locus* ($p<0.05$) showing no Hardy-Weinberg deviation. No genetic structuring was observed between populations when comparing in pairs, or in one collect relative to other, neither when comparing before and after chemoterapic treatment. This work opens interesting perspectives for the study and knowledge of various biological points in parasite-host interaction, pathogeny, and epidemiology and even in the chemotherapy of schistosomiasis, a diseases affecting about 200 million people in more than 70 countries around the world.

I - INTRODUÇÃO:

1 – INTRODUÇÃO

1.1 - Aspectos gerais da esquistossomose

A esquistossomose é uma doença parasitária crônica, debilitante e, em alguns casos, fatal. A doença afeta indivíduos principalmente em áreas rurais, sendo endêmica em 74 países tropicais e subtropicais, atingindo cerca de 200 milhões de indivíduos, 170 milhões na África Sub-Saariana e outros 30 milhões no Norte da África, Ásia e América do Sul. Estima-se que no mundo 600 milhões de pessoas estejam sob risco de infecção, (WHO, 2002), sendo entre 6 e 12 milhões somente no Brasil (Katz & Peixoto, 2000).

Segundo a Organização Mundial de Saúde (OMS), a esquistossomose é a terceira doença tropical de maior importância sócio-econômica e de saúde pública do mundo, figurando entre as doenças de categoria 2 no *portfolio* do TDR/OMS, ou seja, aquelas para as quais existem estratégias de controle, mas o peso da doença para a humanidade persiste, (Remme et al., 2002). Por ano ocorrem aproximadamente 11 mil mortes diretamente relacionadas à esquistossomose. Seqüelas clínicas tardias e a mortalidade/morbidade indireta causadas pela esquistossomose estão relacionadas a 200 mil mortes por ano (TDR Staff, 2005). O número de "*Disability Adjusted Life Years*" (DALYs) provenientes da esquistossomose atinge 1,7 milhões. Um trabalho mais recente, sobre doenças negligenciadas, aponta uma taxa de mortalidade de 280 mil e um DALYs de 4,5 milhões somente na África (Hotez et al., 2006).

A esquistossomose é causada por um trematódeo do gênero *Schistosoma*. Merecem destaque, por sua importância médica, as espécies *Schistosoma mansoni, S. haematobium* e *S. japonicum.* O *S. mansoni* é a espécie com maior distribuição geográfica e a única espécie presente no Brasil (Johnston et al., 1993).

O controle da doença atualmente consiste na quimioterapia, seja para os grupos de risco (escolares, trabalhadores rurais) ao nível das comunidades ou de indivíduos com diagnostico de infecção. Quando medidas efetivas de diagnóstico e quimioterapia são aliadas à educação para a saúde, medidas sanitárias, distribuição de água potável e controles de caramujos, a transmissão da doença pode ser eficientemente eliminada, como no Japão, partes da China e do Brasil. Em contraste, na África sub-Saariana, estima-se que a esquistossomose perca apenas para a malária como causa de morbidade entre as doenças tropicais (TDR Staff, 2005).

O tratamento atual para as esquistossomoses é baseado principalmente no praziquantel, que foi introduzido aproximadamente há 25 anos e se apresenta seguro, bem tolerado e efetivo em dose única (Cioli & Pica-Mattoccia, 2003; Doenhoff & Pica-Mattoccia, 2006; Gryseels et

al., 2006). Entretanto, falhas no tratamento têm sido reportadas e creditadas ao grande número de parasitos imaturos, à rapidez da reinfecção (Cioli, 2000; Gryseels et al., 2001), e à possível resistência dos parasitos à droga (Cioli & Pica-Mattoccia, 2003; Gryseels et al., 2006). A oxaminiquina pode ser usada como droga alternativa, mas, não apresenta efeitos contra o *S. haematobium*, pode provocar efeitos colaterais pronunciados, tem um custo bem mais elevado que o do praziquantel (Gryseels et al., 2006) e também já foi demonstrada a existência de parasitos com resistência a esta droga (Cioli et al., 1992). Infecções repetidas induzem algum grau de imunidade em humanos, assim, uma vacina efetiva poderia oferecer uma alternativa promissora à quimioterapia ou ser usada conjuntamente a esta. Porém, esta vacina ainda não está disponível (TDR Staff, 2005).

1.2 - O parasito

O *Schistosoma mansoni* é um parasito que apresenta dimorfismo sexual na fase adulta, e um ciclo com alternância de gerações entre um hospedeiro intermediário, molusco do gênero *Biomphalaria,* e os hospedeiros definitivos vertebrados, incluindo o homem. A transmissão da doença ocorre quando do contato do homem com águas infestadas por larvas do parasito, as cercárias. O ciclo de vida do parasito se inicia com a penetração das cercárias na pele do hospedeiro definitivo, onde se transformam em esquistossômulos que migram para os pulmões, aonde chegam cerca de 7 dias após a penetração. Do pulmão os esquistossômulos migram para o sistema porta-hepático onde se transformam em vermes que se acasalam para terminar a maturação. Após a maturação os vermes adultos acasalados movimentam-se contra a corrente sanguínea e se alojam no plexo mesentérico, produzindo dezenas de ovos por dia (Pellegrino & Coelho, 1978; Valadares et al., 1981). Parte dos mais de 300 ovos produzidos diariamente por cada par fica retida na mucosa intestinal e nos capilares do sistema porta do hospedeiro, sendo estes os principais causadores da patogenia da doença. Ocasionalmente os ovos podem se alojar em outros tecidos e causar doenças como a neuroesquistossomose (Ferrari, 2004; Araujo et al., 2006). O restante dos ovos é eliminado com as fezes e em contato com a água eclodem liberando miracídios. Estes infectam novos caramujos do gênero *Biomphalaria*, nos quais, cada miracídio se transforma em esporocisto I. Cada esporocisto I, por poliembrionia, origina 150 a 200 esporocistos II. Estes migram para as glândulas digestivas e ovoteste do caramujo, originando as cercárias que serão liberadas na água fechando assim o ciclo evolutivo do parasito (Figura 1).

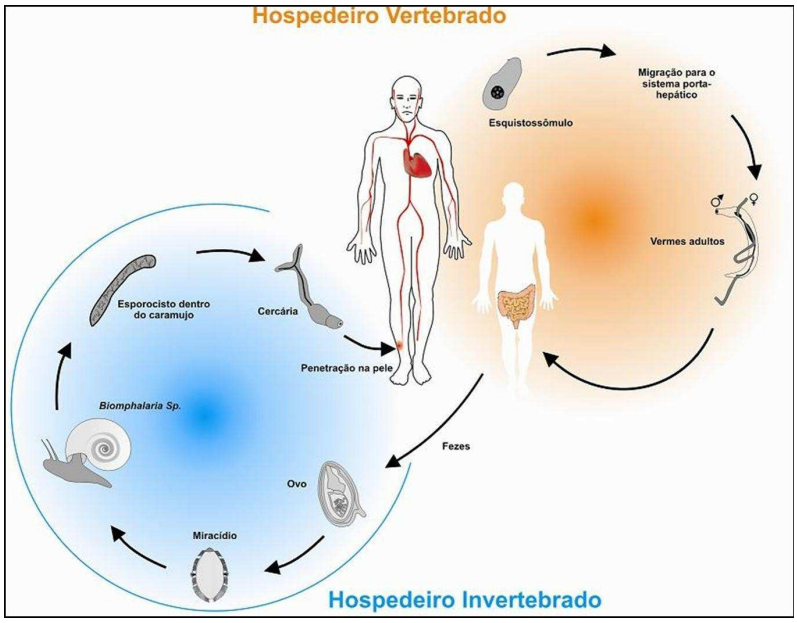


Figura 1 - Ciclo de vida do *S. mansoni*. (http://rgmg.cpqrr.fiocruz.br/)

1.3 - Genoma do *S. mansoni*

S. mansoni tem um genoma diplóide com 7 pares de cromossomos autossômicos e um par de cromossomos sexuais (Short et al., 1979). O gênero *Schistosoma* juntamente com os outros da família Schistosomatidae foram os primeiros, dentre os platelmintos parasitos, a desenvolverem dimorfismo sexual e cromossomos sexuais heteromórficos (Johnston et al., 1993). A fêmea é heterogamética, possuindo o par de cromossomos sexuais ZW e o macho o par ZZ (Short & Menzel, 1960; Short & Grossman, 1981). Os cromossomos variam em tamanho de 18 a 73 Megabases (Mb) e podem ser distinguidos também pela forma e padrão de bandas C (Short & Grossman, 1981). O tamanho do genoma haplóide é de aproximadamente 270Mb com um conteúdo CG de 29,4% (Simpson et al., 1982; Marx et al., 2000). A metilação do DNA genômico não é observada no parasito adulto (Fantappie et al., 2001). O genoma do parasito é constituído de 4% a 8% de seqüências de DNA altamente repetitivas (>1.000 cópias), 35% a 40% de seqüências de média repetitividade (~100 cópias) e

60% de seqüências correspondentes a famílias de genes ou regiões de cópia única (Simpson et al., 1982).

A primeira região de DNA repetitivo descrita em *S. mansoni* consistiu do complexo gênico que codifica o RNA ribosomal (rDNA) (Simpson et al., 1984). Este complexo ocorre como uma unidade de 10 Kilobases (Kb), repetida em tandem, com uma abundância em torno de 100 unidades por genoma haplóide. Cada unidade repetitiva codifica para as três espécies conservadas de rRNA de eucariótos: 5,8S, 18S, e 28S, separadas por duas regiões espaçadoras (Simpson et al., 1984; van Keulen et al., 1985). A família de rDNA é polimórfica com aproximadamente 10% das unidades repetitivas mostrando heterogeneidade de tamanho (Simpson et al., 1984). A análise de seqüências de rDNA nuclear (5,8S, 18S, e 28S) e rDNA mitocondrial (16S rDNA) de várias espécies de *Schistosoma* permitiu estudos filogenéticos de grande importância, assim como a discriminação entre espécies, cepas, sexo (McCutchan et al., 1984; Ali et al., 1991; Despres et al., 1992; Johnston et al., 1993; McManus & Hope, 1993) e até mesmo entre populações de *S. mansoni* sensíveis e resistentes ao hicantone ou à oxaminiquina (McManus & Hope, 1993).

Várias outras seqüências de DNA repetitivo foram descritas em *S. mansoni*. Algumas estão presentes somente em fêmeas, localizando-se no cromossomo W (Spotila et al., 1987; 1989; 1991; Webster et al., 1989; Gasser et al., 1991; Schwarzenbach et al., 2004). A seqüência repetitiva de W1 consiste de uma seqüência degenerada de 476 pares de bases (pb), pode estar presente em até mais de 500 cópias no cromossomo W. Esta região foi utilizada como sonda (Webster et al., 1989) e posteriormente, como alvo para a reação em cadeia da polimerase (PCR) (Dias-Neto et al., 1993a), na sexagem de cercárias.

Seqüências repetitivas, distribuídas de maneira aleatoria ou em arranjos pelo genoma, foram encontradas como parte da estrutura de RNA mensageiros (mRNA), sendo até mesmo traduzidas em proteínas (Spotila et al., 1991; Smith et al., 1992). Uma delas é o elemento repetitivo polimórfico (PRE) de 62 pb que está presente em transcritos de vários tamanhos em vermes adultos. Um destes mRNAs corresponde ao gene SM750. O transcrito maduro possui na sua extremidade 3', anterior à cauda poli A, 5 cópias diretas do PRE, sendo que parte da primeira repetição está contida dentro da única janela de leitura aberta (ORF), composta de 57 aminoácidos (Spotila et al., 1991). Curiosamente, um minissatélite polimórfico encontrado no DNA mitocondrial (mtDNA) do parasito possui algumas regiões idênticas à SM750, que contém várias cópias em tandem do PRE. O significado deste achado ainda é desconhecido,

mas foi sugerida uma possível transferência desta seqüência do núcleo para a mitocôndria (Pena et al., 1995).

O estudo do transcriptoma de *S. mansoni* foi uma iniciativa brasileira, financiada por agências locais, e que teve início em 1992 (Franco et al., 2000). A partir de 1994, a OMS iniciou o financiamento da Rede Genoma de *Schistosoma* para a descoberta de novos genes com o objetivo final de identificar novos alvos para o desenvolvimento de drogas e vacinas. Durante este período a comunidade científica mundial produziu aproximadamente 16.000 etiquetas de seqüências expressas (ESTs) (Oliveira & Johnston, 2001). Posteriormente, 2 outros grandes projetos de seqüenciamento do transcriptoma do *S. mansoni* foram realizados (Franco et al., 2000; Verjovski-Almeida et al., 2003; Oliveira & Bahia, 2004).

Um dos projetos, financiado pela FAPESP/MCT/CNPq, utilizou uma biblioteca normalizada de verme adulto e minibibliotecas de ORESTES (sigla em inglês para *Open reading frames EST sequences*) de 6 estágios do ciclo de vida do parasito (Verjovski-Almeida et al., 2003). O projeto gerou 124.681 ORESTES de *S. mansoni*. As seqüências foram agrupadas resultando em 30.988 "*Schistosoma mansoni* assembled ESTs" (SmAE), correspondendo a aproximadamente 92% do transcriptoma. Dentre as SmAEs anotadas, 23% encontraram identidade com outras seqüências de *S. mansoni* já depositadas, tendo 2% delas identidade com genes conhecidos e 21% com ESTs depositadas no dbest. Os outros 77% não encontraram identidade com seqüências de *S. mansoni,* sendo descritas como novos genes relatados de *S. mansoni*. Destas, 1% apresentou identidade com proteínas de *S. mansoni* conhecidas, sendo descritos como novos parálogos, 20% descritos como ortólogos de genes de outros organismos, e 55% sem relação com nenhum outro gene já relatado de outras espécies (Verjovski-Almeida et al., 2003). A comparação das ESTs de *S. mansoni* com seqüências dos bancos de dados permitiu identificar um grande número de genes de interesse. Foi possível classificar por "Gene Ontology" 8.001 ESTs, obtendo-se principalmente proteínas envolvidas em metabolismo e processos biológicos (Verjovski-Almeida et al., 2003).

O segundo projeto, financiado pela FAPEMIG/MCT/CNPq, consistiu de uma rede genômica formada por instituições do Estado de Minas Gerais que teve como objetivo caracterizar o transcriptoma de diferentes estágios de desenvolvimento do parasito, a partir da geração de 70 mil ESTs convencionais. Esse projeto contribuirá na complementação do projeto desenvolvido em São Paulo e também, será indispensável em estudos de proteoma e SAGE (sigla em inglês para *Serial Analysis of Gene Expression*) (Oliveira & Bahia, 2004;

Shaikenov et al., 2004). Até o final de 2004, foram geradas 74 mil ESTs das diferentes fases do parasito, ainda não disponíveis nos bancos de dados públicos (rgmg.cpqrr.fiocruz.br). Até o momento, mais de 150.000 ESTs foram depositadas pela comunidade científica na divisão dbest do GenBank. O baixo nível de redundância indica que ainda existam muitos novos genes para serem gerados pelas bibliotecas disponíveis (Oliveira & Bahia, 2004).

Os institutos TIGR (*The Institute for Genomic Reasearch*) e Sanger foram os responsáveis pelo seqüenciamento em larga escala do genoma de *S. mansoni,* gerando juntos, uma cobertura igual a 9 vezes o tamanho do genoma completo, estimando-se que menos de 0,5% do genoma não tenha sido seqüenciado. A estratégia de seqüenciamento utilizada foi de "Whole Genome Shotgun" (WGS), por seleção aleatória de clones e com seqüenciamento de ambas as extremidades. Estas seqüências estão atualmente sendo montadas e anotadas (Oliveira & Bahia, 2004; El Sayed et al., 2004) e um banco de dados genômico está em fase de construção pelo nosso grupo (www.schistodb.net) alem do Genedb (www.genedb.org).

Para uma revisão mais completa sobre genoma e genômica de *Schistosomas* ver Oliveira e cols., (2004) anexo I, e Verjovski-Almeida e cols., (2007) A contribuição brasileira para estas pesquisas será publicada em um número da Acta Tropica ainda em 2007.

Atualmente a abordagem mais promissora para a identificação de novos alvos de drogas, vacinas e reagentes para diagnóstico, assim como para a compreensão da resistência às drogas, diversidade antigênica, infectibilidade e patologia, consiste em compreender e decifrar, com o auxilio das ferramentas de bioinformática, as informações nos genomas dos parasitos (Remme et al., 2002; TDR Staff, 2005).

1.4 - Variabilidade fenotípica do *S. mansoni*

Características como variações morfológicas, diferenças de infectividade em moluscos e/ou vertebrados, resistência às drogas, já foram citadas como possível conseqüências da variabilidade genética do *S. mansoni*. Em 1963, Paraense e Corrêa, publicam ao mesmo tempo, dois trabalhos nos quais relatam ligações entre variações nas taxas de infecção de diferentes cepas *Biomphalaria glabrata* por uma cepa de *S. mansoni*, (Paraense & Correa, 1963b), e diferenças nas taxas de susceptibilidade de cepas de *B. tenagophila* (Paraense & Correa, 1963a), a diferentes cepas de *S. mansoni*. Em 1988 Dias e cols., mostram diferenças nas taxas de infecção de duas linhagens de *Biomphalaria glabrata* por cepas resistentes e susceptíveis de *S. mansoni*, com as cepas resistentes sendo menos infectivas. Em 1973, Katz e

cols., após estudarem 2 pacientes que permaneciam infectados mesmo depois de passarem por três tratamentos consecutivos em um período de 15 meses, publicam, pela primeira vez na literatura, trabalho demonstrado a existência de cepas resistentes de *S. mansoni* provenientes de pacientes tratados. Resultados semelhantes são apresentados por Dias e cols., (1982) e Coles e cols., (1986; 1987) e por Drescher e cols., (1993), ao estudar cepas provenientes de pacientes previamente tratados com oxaminiquina e/ou praziquantel. Trabalho de 1980, publicado por Araújo e cols., mostra as diferenças no percentual de queda na produtividade de ovos em diferentes cepas de *S. mansoni* após tratamento com hicantone ou oxaminiquina. Em todos estes trabalhos as variações fenotípicas apresentadas pelo parasito são atribuídas à sua variabilidade genética (McManus & Hope, 1993). Neste sentido, o conhecimento da variabilidade genética intra e interespecífica do parasito é considerado de fundamental importância para a compreensão dos mecanismos biológicos envolvidos nas interações parasito-hospedeiro, na patogenia e epidemiologia da doença.

1.5 - Estudo da variabilidade genética do *S. mansoni*

Várias abordagens já foram utilizadas para se estudar a variabilidade genética do *S. mansoni*. Inicialmente esta variabilidade foi avaliada através da análise eletroforética de isoenzimas. Fletcher e cols., (1981b) realizaram estudos com cepas de *S. mansoni* que apresentavam diferenças em relação ao nível de infectividade em *B. glabrata.* Foi observado que 3 de 18 *loci* estudados mostravam-se polimórficos, sem, entretanto, ter sido determinada uma clara correlação entre infectividade e polimorfismo isoenzimático (Fletcher et al., 1981ac; LoVerde et al., 1985ab). Navarro e cols., (1992) utilizaram a mesma técnica para o estudo de cepas de diferentes regiões geográficas, observando diferenças entre as cepas. Porém, quando analisaram vermes individuais todos os *loci* mostraram-se monomórficos para os vermes de uma mesma cepa.

A técnica de Southern *blotting* foi a segunda técnica de análise molecular utilizada para o estudo do genoma do *S. mansoni*. A utilização de sondas de DNA ribosomal (rDNA), foi descrita para a distinção de espécies, cepas e sexo deste parasito (McCutchan et al., 1984; Simpson et al., 1984; van Keulen et al., 1985). Marcadores moleculares detectáveis como polimorfismos de tamanho de fragmentos de restrição RFLP (sigla em inglês para *Restriction Fragments Length Polymorphisms*) foram correlacionados com o surgimento de resistência ao hicantone em *S. mansoni* (Brindley et al., 1991), o que, entretanto, não foi encontrado por

Vieira et al., (1991), em relação à resistência à oxaminiquina, quando utilizando a mesma técnica.

Resultados obtidos utilizando RAPD (sigla em inglês para *Random Amplified Polymorphic DNA*), com diferentes cepas e populações de *S. mansoni*, isoladas de regiões geograficamente distantes como o Brasil e a África, ou mantidas por diferentes intervalos de tempo em laboratório, mostraram um alto grau de similaridade entre as cepas estudadas (Barral et al., 1993; Dias-Neto et al., 1993b; Pinto et al., 1997). Os primeiros autores encontraram apenas 5% de polimorfismo entre as mais de 300 bandas obtidas através do uso de 50 iniciadores, confirmando os achados obtidos por isoenzimas e sondas de rRNA.

A análise de polimorfismos de tamanho de regiões do DNA mitocondrial (mtDNA) também já foi usada no estudo da variabilidade em *S. mansoni*. Este marcador mostrou variabilidade dentro do gênero *Schistosoma* (Despres et al., 1991; 1992; Le et al., 2000), e entre diferentes clones de uma mesma cepa de *S. mansoni* (Pena et al., 1995; Jannotti-Passos et al., 1997). Recentemente, quatro regiões do mtDNA (COI, 16S-12S rDNA, cyt b-ND4L-ND4 e NDI) correspondendo a um fragmento de aproximadamente 2500pb foram utilizadas em estudos filogeográficos (Morgan et al., 2005). Foram comparados 143 espécimes de *S. mansoni* coletados de 53 localidades da África, Ásia, América do Sul e Caribe. Neste estudo foram encontrados 85 diferentes haplótipos possibilitando a separação dos espécimes estudados em 5 grupos, 3 exclusivamente do leste africano, sugerindo ser ali o local de origem da espécie. Este trabalho mostrou também a ocorrência de um pequeno número de haplótipos (7) no Novo Mundo, sugerindo a recente introdução deste parasito nas Américas (Morgan et al., 2005). O principal problema com o uso do mtDNA é que, devido à herança materna as análises refletirão apenas um lado dos processos genealógicos apresentados pela espécie ou população em estudo (Hartl & Clark, 1997).

Os marcadores do tipo polimorfismos de base única (SNP) representam uma fonte abundante de alelos de um determinado gene, podendo estar presentes em quase todos os *loci* gênicos, sendo, por esta abundância, mais informativos que microssatélites. Entretanto, poucos destes marcadores haviam sido identificados em *S. mansoni* (Oliveira & Johnston, 2001; Blanton et al., 2005). Recentemente nosso grupo identificou um grande número de SNPs após uma análise global do transcriptoma do parasito (Simões et al., 2007).

Devido à homogeneidade genotípica apresentada por *S. mansoni* frente aos marcadores citados, exceto SNPs, a busca de outros marcadores polimórficos que detectem um maior grau de variabilidade e, que sejam mais adequados aos estudos de genética de populações neste

parasito é imprescindível. Neste contexto, os microssatélites têm se mostrado como um dos mais promissores marcadores disponíveis atualmente.

1.6 - Os microssatélites.

Microssatélites, também conhecidos como: seqüências curtas repetidas em tandem, STR (*short tandem repeats sequences*), repetições de seqüências simples, SSR (*simple sequence repeats*) ou seqüências repetitivas simples, SRS (*simple repetitive sequences*), são regiões de DNA repetitivo, compostas de pequenos motivos de no máximo 6pb repetidos em tandem, e presentes em genomas de eucariotos e procariotos (Field & Wills, 1996; Toth et al., 2000). Amplamente utilizados como marcadores genéticos os microssatélites têm como especial atributo o fato de apresentarem maiores taxas de mutação que o restante do genoma (Jarne & Lagoda, 1996) o que gera polimorfismos nas populações. Estudos *in vitro* e *in vivo* indicaram que *loci* de microssatélites são altamente instáveis, possuindo algumas das maiores taxas de mutação já observadas. As taxas de mutação em humanos se encontram na ordem de 10^{-2} a 10^{-6} repetição por *locus* por geração (Ellegren, 2000; Sia et al., 2000).

Os microssatélites são classificados de acordo com tipo de repetição como perfeitos, imperfeitos, interrompidos e compostos (Weber, 1990). Nos perfeitos existe um motivo único de repetição, sem interrupções (TATATATATA), enquanto que nos imperfeitos, existe uma base diferente intercalada na seqüência que se repete (TATATACTATATA). No caso do interrompidos há uma pequena seqüência, diferente da principal, dentro da repetição (TATATACGTGTATATATA). Nos compostos, há uma repetição perfeita ou imperfeita associada a outra repetição com motivo diferente (TATATATAGTGTGTGTGT).

Até poucos anos atrás os microssatélites eram tidos como marcadores seletivamente neutros e não afetados por pressões seletivas. Entretanto, hoje existem evidências de que a expansão no número de repetições pode causar doenças em seres humanos, por exemplo, a doença de Huntington e causada pelo aumento do comprimento do bloco de repetição CAG presente no gene para proteína de huntington localizado no cromossomo 4 (Rubinsztein, 1999). Estas expansões podem estar ligadas à inativação do sistema de reparo de DNA, gerando instabilidade nos microssatélites. Instabilidades em repetições de microssatélites estão presentes em aproximadamente 15% dos casos de câncer de cólon intestinal (Grady, 2004) e é a única forma de mutação ligada a mais de 40 desordens neurológicas, neurodegenerativas e neuromusculares, especialmente na classe dos trinucleotídeos

(Rubinsztein, 1999; Cummings & Zoghbi, 2000; Everett & Wood, 2004; Pearson et al., 2005).

A estratégia utilizada para analisar o polimorfismo dos microssatélites é a amplificação pela PCR usando um par de iniciadores específicos flanqueando o segmento com as repetições e a verificação do tamanho do fragmento amplificado por eletroforese.

Embora extensivamente utilizados nas mais variadas áreas da genética a dinâmica mutacional destes marcadores não é ainda bem conhecida (Schlotterer, 2000). Vários mecanismos moleculares foram propostos como responsáveis pela geração da variabilidade dos *loci* de microssatélites, incluindo erros durante a recombinação, "crossing-over" desigual e o deslizamento das fitas do DNA durante a replicação – "*DNA polymerase slippage*" (Strand et al., 1993; Ellegren, 2000).

Com relação aos erros durante a recombinação, já foi mostrado que cepas de *Escherichia coli*, com ou sem sistema de recombinação funcional, apresentaram as mesmas taxas de mutação, sugerindo que a recombinação não era o mecanismo predominante na geração da variabilidade em microssatélites (Levinson & Gutman, 1987).

Quando ocorre *crossing-over* desigual, pode haver mudanças como a perda ou o ganho de grande número de repetições. Isto se deve à formação de alças nas regiões de microssatélites (Figura 2), significando que parte de diferentes tamanhos de cada cromossomo serão trocadas e que um dos homólogos receberá um grande número de repetições enquanto o outro receberá um pequeno número (Oliveira et al., 2006).

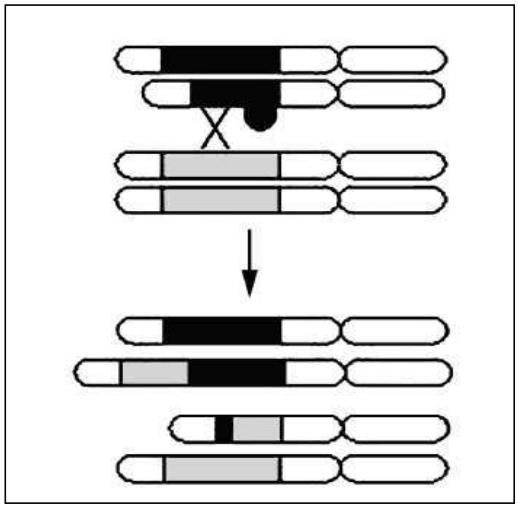


Figura 2 - Representação esquemática do crossing-over desigual entre 2 cromossomos homólogos. As regiões cinza e preta correspondem a sequências de microssatélites (Oliveira et al., 2006).

Durante a replicação as fitas de DNA podem se desanelar e se anelar novamente em locais errados formando alças. Em conseqüência, após o reanelamento, se a alça formada estiver na fita molde do DNA haverá uma contração do microssatélite com redução no número de repetições. Por outro lado, se esta alça estiver na fita que estiver sendo formada, haverá uma expansão do microssatélite com aumento no número de repetições (Figura 3).

Altas taxas de deslizamento foram demonstradas, mas parecem levar somente a pequenas alterações no número de repetições (Hentschel, 1982; Streisinger & Owen, 1985; Schlotterer & Tautz, 1992). Este mecanismo pode desestabilizar o microssatélite, seja por não haver sistema de reparo efetivo para as alças do DNA ou por alterações na DNA polimerase ou seus co-fatores, resultando em altas taxas de deslizamentos. Mutações nos genes do sistema de reparo de DNA aumentam em mais de 700 vezes a instabilidade de microssatélites em *E. coli* (Bichara et al., 2000), leveduras (Strand et al., 1993; Sia et al., 1997) e células de mamíferos (Kolodner & Marsischky, 1999), enquanto que mutações afetando os domínios DNA polimerase produzem efeitos menos drásticos (Sia et al., 1997).

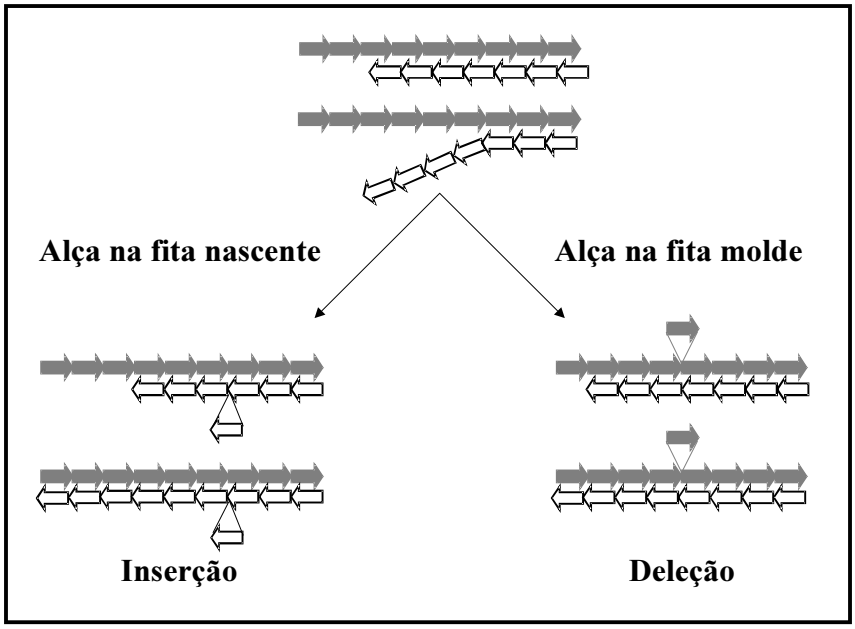


Figura 3 - Representação esquemática do deslizamento das fitas do DNA durante a replicação "*DNA polymerase slippage model*". Durante a replicação as fitas desanelam e ao reanelar podem formar alças. Conforme as alças estejam na fita nascente ou na fita molde irá ocorrer a formação de um novo alelo com a inserção ou deleção de uma unidade de repetição (Kokoska et al., 1998).

Devido ao fato de apresentarem herança mendeliana, ampla distribuição no genoma, alto nível de polimorfismo e serem facilmente analisáveis por PCR, os microssatélites têm se tornado uma fonte precisa de informações sobre genética e evolução de populações (Wilson & Balding, 1998).

Análises de microssatélites são empregadas com propósitos forenses (Pinheiro et al., 1997), genética de populações de humanos (Santos et al., 1993) e dos mais diferentes tipos de animais: mamíferos (Ishibashi et al., 1997), aves (Gibbs et al., 1996), peixes (Zheng et al., 1995), insetos (Solano et al., 1997), moluscos (Jarne et al., 1994); plantas (Lopes et al., 2006), fungos (Groppe et al., 1995), protozoários (Russell et al., 1999; Oliveira et al., 1999), e outros microrganismos (Field & Wills, 1996; King et al., 1997), além de aplicações em estudos de sócio-biologia, tais como, padrões de acasalamento, organização social de populações e outros (Schlotterer & Pemberton, 1994).

No campo da parasitologia, os microssatélites foram descritos no estudo de diversos parasitos, tanto humanos como animais. Na parasitologia humana podemos citar a identificação de microssatélites em diferentes espécies do gênero *Plasmodium* (van Belkum et al., 1992) com especial referência à espécie *Plasmodium falciparum,* parasito responsável pela malária em humanos (Su & Wellems, 1996). Citamos ainda a descrição de microssatélites em tripanosomatídeos como *Trypanosoma cruzi*, parasito causador da doença de Chagas (Oliveira et al., 1999) e *Leishmania* spp (Rossi et al., 1994; Rodriguez et al., 1997; Russell et al., 1999). No campo da parasitologia animal podemos citar, entre outros, a identificação e utilização de microssatélites na diferenciação de isolados de *Trichinella pseudospiralis*, parasito de aves e mamíferos, provenientes da América do Norte, Europa e Austrália (Zarlenga et al., 1996), e de populações do nematóide *Haemonchus contortus*, parasito gastrintestinal de ovinos e caprinos provenientes da Europa, África e Ásia (Hoekstra et al., 1997).

1.7 - Os microssatélites na análise de populações de *S. mansoni.*

Os primeiros trabalhos descrevendo a presença e variabilidade de microssatélites em *S. mansoni* datam de 2000 (Durand et al., 2000). Naquele trabalho os autores estudaram populações do parasito de ratos silvestres (*Rattus rattus*), naturalmente infectados, na ilha de Guadalupe e descreveram 33 *loci* obtidos de análises de bancos de dados e de bibliotecas enriquecidas, dos quais 11 mostraram-se polimórficos, com 2 a 8 alelos por *locus*. Em 2001, Blair e cols., (2001) isolaram e caracterizaram 10 outros *loci* polimórficos de bancos de dados e em uma biblioteca genômica do verme. Com estes marcadores os autores analisaram três populações de parasitos da África, encontrando entre 2 e 6 alelos por *locus* e uma heterozigosidade entre 0,33 e 1,0, demonstrando alta variabilidade tanto intra quanto interpopulações. Também em 2001, Curtis e cols., (2001) descreveram 5 outros *loci* polimórficos em isolados de *S. mansoni*, com 5 a 8 alelos por *locus* e heterozigosidade entre 0,58 e 0,70. Estes autores mostraram ainda resultados da amplificação destes marcadores em outras espécies do gênero. Em 2002, analisando os bancos de dados do Genbank, nosso grupo selecionou 6 *loci* que foram utilizados em estudos comparando uma cepa de *S. mansoni* mantida em laboratório (LE) a isolados de campo do parasito (Rodrigues et al., 2002a). Foram observados em média, 9,5 alelos por *locus* na cepa LE e de 14,5 nos isolados de campo, sugerindo maior variabilidade nos isolados de campo. Ainda em 2002 nosso grupo detectou outros 2700 *loci* em seqüências genômicas (GSS) e outros 100 *loci* em bibliotecas

enriquecidas para microssatélites (Rodrigues et al., 2002b). Em 2004, Stohler e cols., (2004) ao estudar cepas de laboratório e isolados de campo observaram, nas cepas de laboratório, um número médio de alelos variando entre 10 e 14% daquele encontrado para os isolados de campo, sugerindo também, maior variabilidade nestes que nas cepas de laboratório (Stohler et al., 2004). Ao estudar populações de ratos silvestres (*R. rattus*) naturalmente infectados na ilha de Guadalupe, foi observada a existência de maiores diferenças genética entre fêmeas do que entre machos de um mesmo hospedeiro. Isto implica em um padrão aleatório na escolha de parceiros sexuais (pangamia) entre os vermes, resultando em baixo nível de endogamia (Prugnolle et al., 2004ab). Ainda usando microssatélites, estes autores estudaram o padrão de dispersão deste parasito entre seus hospedeiros intermediários e definitivos, mas não observram correlações entre as distâncias genéticas das populações de parasitos e as das populações de caramujos ou ratos (Prugnolle et al., 2005).

Em 2006 Agola e cols., compararam a estrutura e a diversidade genética de 7 populações de parasitos de diferentes localidades do Kenia. Neste trabalho, utilizando 5 *loci* de microssatélites, os pesquisadores observaram um alto nível de diversidade genética e diferenças entre as populações. Estes dados indicam uma estruturação das populações, o que se deve ao limitado fluxo gênico entre elas e ao seu grande tamanho.

Trabalho publicado em 2007 (Gower et al., 2007) descreve o uso de miracídios e da PCR multiplex, na análise genética de 7 isolados de *S. mansoni* em Uganda. Neste trabalho é feita a comparação de amostras de miracídios diretamente isoladas de fezes humanas, com vermes adultos e miracídios oriundos de infecções de caramujos de laboratórios, também por miracídios destes mesmos isolados. As análises do material passado pelo laboratório mostram uma errônea estruturação geográfica destas populações indicando os possíveis problemas causados pela amostragem indireta em populações de *S. mansoni*. Este mesmo grupo já havia utilizado metodologia semelhante na genotipagem de miracídios de *S. japonicum* infectando animais silvestres na China e nas Filipinas (Shrivastava et al., 2005). Naquele trabalho os autores apresentam o uso de miracídios como uma forma de estimar de maneira acurada e ética, a estrutura genética de populações de parasitos que apresentam múltiplos hospedeiros definitivos.

Todos os trabalhos realizados até o momento com microssatélites ou outros marcadores, utilizaram como fonte de DNA principalmente vermes adultos, e em alguns casos, cercárias de *S. mansoni* e agora miracídios. O problema chave para o estudo com schistosomas até então era a não disponibilidade de parasitos adultos ou em estagio larval (cercarias), o que implicava na necessidade e custos da manutenção de pelo menos parte do ciclo do parasito em

laboratório além de significar a seleção artificial de genótipos (Gower et al., 2007). Outro problema com os microssatélites utilizados é que quase todos derivam de sequências de cDNA que estão sujeitas a pressões seletivas. Neste trabalho foi aumentado significativamente o número de marcadores não derivados de cDNA, com o uso de bibliotecas genômicas e diminuídos os custos do trabalho com a utilização de ovos como fonte de DNA do parasito.

II - JUSTIFICATIVA

2 - JUSTIFICATIVA

2.1 – Justificativa

Os estudos dos padrões genéticos de populações do parasito contribuem para o entendimento da dinâmica da doença, especialmente em áreas endêmicas sob a pressão de quimioterapia. Todos os trabalhos realizados até o momento, com os mais diversos marcadores, comparam populações de diferentes localidades geográficas ou obtidas em diferentes intervalos de tempo, e utilizam como fonte de material genético cercárias e/ou vermes adultos, e mais recentemente miracídios (Gower et al., 2007). Em nosso trabalho iremos aumentar significativamente o número de marcadores disponíveis e comparar infrapopulações de uma determinada área utilizando ovos individuais do parasito como material genético. Esperamos assim, descartar a influencia dos filtros genéticos, representados pelos moluscos e vertebrados de laboratório, e podermos determinar com maior precisão o "pool" gênico dos parasitos de cada paciente. Teremos desta forma um retrato bastante fiel do perfil genético destas populações de parasitos. Analisaremos a estruturação e a variabilidade genética das infrapopulações de parasitos infectando cada um dos moradores de uma área endêmica e o impacto ocorrido nesta variabilidade após o tratamento com a quimioterapia usual. Estas informações poderão ser úteis no desenho de futuras estratégias de controle da doença.

III - OBJETIVOS

3- OBJETIVOS

3.1 – Objetivo geral

Determinar a variabilidade e a estrutura inter e intrapopulacional do *Schistosoma mansoni* infectando pessoas residentes em área endêmica.

3.2 – Objetivos específicos

1 Identificar marcadores genômicos polimórficos do tipo microssatélites para *S. mansoni.*

2 Desenvolver metodologia para o uso de ovos de *S. mansoni* como fonte de material genético.

3 Determinar a diversidade genética dos parasitos infectando humanos em Virgem das Graças.

4 Avaliar a influência da quimioterapia na diversidade genética dos parasitos infectando humanos em Virgem das Graças.

IV - METODOLOGIA

4 – METODOLOGIA

4.1 – A identificação de microssatélites em seqüências de *S. mansoni*

Para o desenvolvimento de marcadores seguimos simultaneamente duas abordagens. Na primeira, 4 bibliotecas genômicas enriquecidas para seqüências repetitivas de AAT, CA, GA e TAGA (fragmentos entre 300 e 700bp) foram construídas e clonadas em pUC19 pela Genetics Information System, Chatsworth, CA. Clones recombinantes foram transformados em *Escherichia coli,* DH5α e selecionados em placas de LB/agar contendo X-gal 800µg/ml, IPTG 800µg/ml e ampicilina 0,1µg/µl (Invitrogen). As colônias brancas foram coletadas e crescidas a 37°C em 5ml LB com ampicilina (0,1µg/µl) sob agitação. Plasmídeos foram preparados com R.E.A.L. prep 96 plasmid Kit (Qiagen), seguindo instruções do fabricante e seqüenciados em seqüenciador automático de DNA (ALF – Pharmacia) usando Thermo Sequenase Fluorescent Primer Kit (Amersham Biosciences)

A segunda abordagem foi a busca por sequências apresentando microssatélites, dentre as seqüências de DNA transcritas - ESTs (*expressed sequence tags*) e seqüências genômicas (*BACs ends*) depositadas em bancos de dados públicos. As seqüências obtidas das bibliotecas e as depositadas nos bancos de dados foram agrupadas com o programa CAP3 (Huang & Madan, 1999) e analisadas utilizando-se o programa RepeatMasker (A.F.A. Smit, R. Hubley & P. Green RepeatMasker at http://repeatmasker.org) para a busca de regiões contendo microssatélites.

4.2 – Escolha de iniciadores para PCR de microssatélites

Somente as seqüências que apresentaram os microssatélites perfeitos localizados a mais de 50pb de suas extremidades 5'e 3' foram utilizadas para o desenho dos iniciadores. Para tal, usamos o programa Fast PCR Fast PCR© (Kalendar R, 2003), seguindo os seguintes parâmetros:

1) o tamanho esperado para cada fragmento gerado na PCR deveria estar situado entre 150pb e 500pb, possibilitando assim a análise simultânea de vários fragmentos gerados individualmente ou em PCR-multiplex. 2) O tamanho dos iniciadores deveria ser de 18 a 22 nucleotídeos, com as temperaturas de anelamento entre 54°C e 58°C. 3) Os iniciadores não poderiam apresentar complementaridade, principalmente nas suas extremidades 3', consigo mesmos ou com outros iniciadores. 4) Iniciadores que gerassem fragmentos com diferenças

de tamanho de aproximadamente 100pb não poderiam apresentar complementaridade entre si, possibilitando seu uso em PCR-multiplex.

Um dos iniciadores de cada par foi marcado em sua porção 5' com fluoresceína, o que possibilitaria seu uso na determinação exata do tamanho dos fragmentos amplificados quando analisados pelo seqüenciador automático de DNA, A.L.F-. *Automatic Laser Fluorescence* (Pharmacia-LKB), com o programa ALF Fragment Manager.

4.3 – Material Biológico

Usamos como fonte de DNA ovos de *S. mansoni* recolhidos de fezes de moradores de áreas endêmicas. Este trabalho foi realizado em colaboração com a Dra. Andréa Gazzineli que mantém atividades de campo nas áreas em estudo. Foram coletadas amostras de fezes de moradores de Virgem das Graças - VDG; (latitude -41,3º, longitude -16,9º), Município de Ponto dos Volantes (latitude -41,5º, longitude -16,7º), área endêmica no Vale do Jequitinhonha, região nordeste de Minas Gerais. Em março de 2001 foram feitos exames de fezes (Kato-Katz) de toda a população 589 pessoas, de acordo com o censo de 2001 (Gazzinelli et al., 2006) (exame 1), e em junho/2001 amostras de fezes foram coletadas de 79 moradores infectados por *S. mansoni*. Estes moradores foram tratados com praziquantel (PZQ1). Três meses após o tratamento (setembro/2001) foram realizados novos exames (2). Foram coletadas amostras de fezes de 26 moradores que ainda se apresentavam infectados e estes moradores passaram então por novo tratamento com praziquantel (PZQ2). Em outubro de 2002 foram realizados novos exames (3), novas amostras de fezes foram coletadas de 16 moradores infectados e realizamos novo tratamento (PZQ3). Em abril de 2003 foi realizado novo exame (4), e novo tratamento (PZQ4), desta vez não houve coleta de amostras. Outro exame (5) foi realizado em outubro de 2005 com coleta de amostras de 11 moradores infectados e novo tratamento foi realizado (PZQ5). No total foram coletadas amostras de fezes de 114 moradores em VDG. A Figura 4 apresenta o esquema com as datas dos exames de fezes, amostragens de material biológico e tratamentos de moradores realizados em VDG.

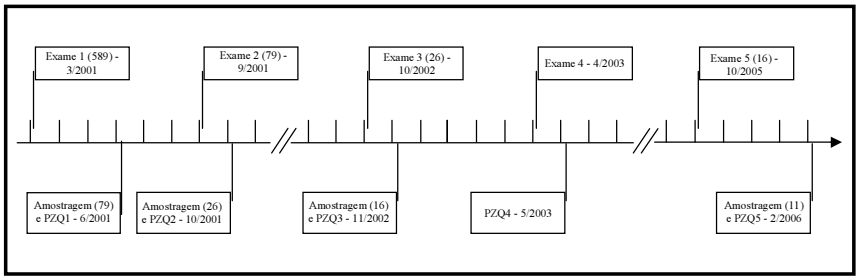


Figura 4 – Representação esquemática dos protocolos de exames de fezes, coleta de material biológico e tratamentos (PZQ1-5) de moradores em VDG. PZQ= Praziquantel. Entre parênteses, número de indivíduos examinados e/ou amostrados.

4.4 – Obtenção de ovos de *S. mansoni*.

Para a obtenção dos ovos de *S. mansoni* as amostras de fezes coletadas dos residentes de VDG passaram inicialmente pelo processo de sedimentação espontânea. As fezes foram diluídas em solução de NaCl a 1,7%, filtradas em gaze, sedimentadas em cálices de fundo cônico por 30min. O sobrenadante foi descartado e o sedimento ressuspendido em salina 1,7% e filtrado seqüencialmente em 3 telas com tamanho de malhas decrescentes (Tabela 1). Os ovos do parasito, que são elipsóides, e têm um diâmetro maior de 142µm e um menor de 60µm (Jourdane & Theron, 1987), ficaram retidos na terceira tela, juntamente com um mínimo de detritos. Este sedimento foi transferido para uma placa de petri com salina a 1,7%. Os ovos foram separados um a um com ajuda de uma lupa, transferidos para microtubos e armazenados a -20ºC em 10µl de salina a 1,7%.

Tabela 1: Tamanho de malhas das telas de filtração das fezes

Tela	Malha – *Mesh* (#)*	μm^2
1	90#	190
2	120#	120
3	500#	19

**Mesh* (#) = número de aberturas por polegada linear

Neste trabalho foram analisados ovos coletados de amostras de fezes de moradores na primeira coleta (33) e na segunda coleta (20). A região de Virgem das Graças foi subdividida

em 5 subáreas; Card1= Cardoso 1, Card2= Cardoso 2, Card3= Cardoso 3, Suss= Suçuarana, VdG= Virgem das Graças (Tabela 2 e Figura 5). Os moradores foram designados pela subárea de origem, número da casa e número de identificação. Os moradores da segunda coleta têm ainda uma letra 'a' acompanhando esta designação, assim temos, por exemplo, Card1-150/353 (1ª coleta) e Card1-150/353a (2ª coleta). As amostras de parasitos, coletadas de cada morador, passaram a compor uma infrapopulação.

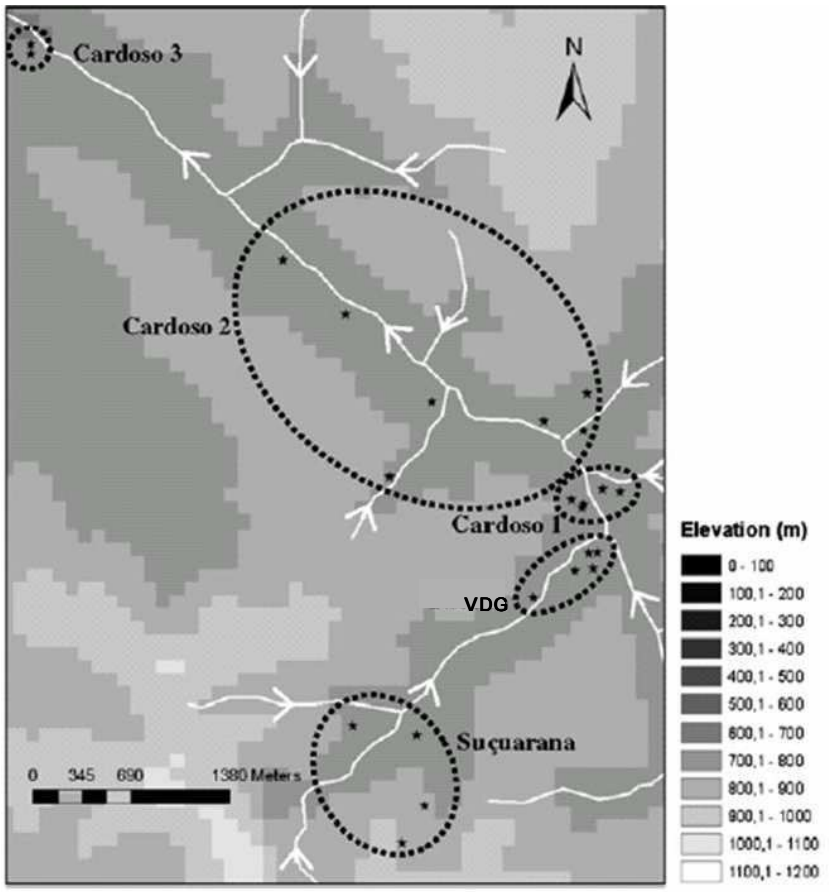


Figura 5 – Mapa da região de Virgem das Graças mostrando a localização aproximada das 5 subáreas em que esta foi subdividida: Cardoso 1 (Card1), Cardoso 2 (Card2), Cardoso 3 (Card3), Suçuarana (Suss) e Virgem das Graças (VdG), estrela representam casas com amostras coletadas, linhas brancas representam os rios com a direção do fluxo e a escala de cinza a topografia (Thiele et al., 2007)

4.5 – Obtenção de vermes adultos.

Alguns dos ovos de *S. mansoni* encontrados nas amostras de fezes coletados foram postos a eclodir em água fresca e os miracídios foram utilizados na infecção de caramujos *B. glabrata*. As cercárias liberadas por estes caramujos foram utilizadas na infecção de camundongos, dos quais, por perfusão de seus sistemas porta-mesentéricos, foram obtidos os vermes adultos (Pellegrino & Siqueira, 1956), o DNA extraído destes vermes foi utilizado na padronização das reações de PCR.

Tabela 2 - Lista de moradores de Virgem das Graças, dos quais foram analisados ovos coletados em amostras de fezes.

Área	Casa	Identificação	OPG 1ª Coleta	OPG 2ª Coleta
Card1	147	338	67	NC
Card1	149	348	NC	140
Card1	150	352	NC	720
Card1	150	353	80	380
Card1	151	359	232	NC
Card1	154	365	246	32
Card2	164	399	82	NC
Card2	165	403	152	NC
Card2	165	407	136	NC
Card2	165	408	224	48
Card2	165	409	448	NC
Card2	167	415	144	NC
Card2	169	542	NC	572
Card2	175	384	NC	1304
Card2	182	464	64	NC
Card3	198	495	744	NC
Card3	199	499	4184	NC
Card3	199	500	3616	NC
Card3	203	516	368	NC
Card3	204	523	2088	56
Card3	204	524	3032	NC
Card3	204	525	4312	84

Tabela 2 - Continuação

Área	Casa	Identificação	OPG 1ª Coleta	OPG 2ª Coleta
Card3	204	526	448	180
Card3	207	529	144	NC
Suss	114	259	56	NC
Suss	116	261	752	NC
Suss	122	275	272	NC
Suss	125	282	64	160
Suss	125	283	NC	36
Suss	125	284	NC	100
Suss	125	284	NC	100
Suss	133	305	76	NC
Suss	134	308	NC	48
Suss	137	315	64	108
Suss	145	330	152	NC
VdG	13	50	NC	32
VdG	26	70	88	NC
VdG	26	637	NC	320
VdG	46	91	NC	48
VdG	46	543	792	40
VdG	75	141	280	NC
VdG	75	145	88	NC
VdG	83	152	NC	184
VdG	97	205	1816	NC
VdG	99	210	252	NC

Card1= Cardoso 1, Card2= Cardoso 2, Card3= Cardoso 3, Suss= Suçuarana, VdG= Virgem das Graças, OPG=Ovos por grama de fazes, NC= não coletado.

4.6 – Extração de DNA de ovos e vermes adultos de *S. mansoni*

Para a extração de DNA dos ovos de *S. mansoni* foram testados 4 métodos:

1) "ROSE - Rapid One Step Extraction" (Steiner et al., 1995).

2) "Sorensen" (Sorensen et al., 1998).

3) Resina - Instagene matrix ® Bio-Rad

4) fenol/clorofórmio (Sambrook et al., 1989).

Em todos os 4 métodos, previamente à extração, os tubos contendo os ovos foram centrifugados por 5min a 13.000x g à temperatura ambiente, e o excesso de líquido foi retirado cuidadosamente para não retirar o ovo. Para o rompimento da casca do ovo, utilizamos choque térmico com nitrogênio líquido (-196ºC) por 30s e água fervendo por 30s. Este procedimento foi repetido três vezes.

Para a extração de DNA de vermes adultos de *S. mansoni,* utilizamos o método de fenol/clorofórmio.

4.6.1 - "ROSE - Rapid One Step Extraction"

A cada tubo contendo um ovo foram adicionados, 100µl de tampão de extração (50mM Tris-HCl, pH 8,0, 50mM EDTA, 100mM NaCl e 1% de SDS) e 50µg/ml de proteinase K (Invitrogen). A amostra foi incubada a 37°C por período entre 12h e 16h. Foram adicionados 200µl da solução tampão "Rose" (10mM Tris-HCl, pH 8,0; 0,3 M EDTA, pH 8,0; lauril sarcosil de sódio a 1% e polivinilpolipirrolidone-PVPP a 1%). O tubo foi incubado por 20min a 95°C em banho-maria, colocado em gelo por 5min, centrifugado a 13.000x g por 10min e a fase aquosa foi transferida para microtubos de 1,5ml. A esta foi adicionado acetato de sódio 3M, pH 5,2 na proporção de 1/10 do volume, seguida de centrifugação e lavagem com etanol 70% do sedimento de DNA. O etanol da amostra foi evaporado e o material foi ressuspendido em 25µl de TE (10mM Tris, 1mM EDTA, pH 8,0) e estocado a –20ºC.

4.6.2 - "Sorensen"

Ao tubo contendo o ovo foram adicionados 50µl de um tampão, constituído por 4 partes de um "tampão homogenizador" (0,1M NaCl; 0,2M Sucrose; 0,01M EDTA; 0,05M Tris-HCl pH 8,0) e uma parte de tampão de lise (0,25mM EDTA, 2,5% SDS; Tris 0,5mM pH 9,2). O tubo foi incubado a 65ºC por 30min. Posteriormente foram acrescentados 7µl de

acetato de potássio a 8M à mistura. O tubo foi incubado em gelo por 30min e centrifugado por 15min a 13.000x g à temperatura ambiente. O sobrenadante foi removido e transferido para outro tubo, ao qual foram acrescentados 150µl de etanol 100% a 4ºC. Este foi então, incubado a -20ºC por 12 a 16h, centrifugado por 5min a 13.000x g à temperatura ambiente e o sedimento lavado com etanol 70%. O etanol da amostra foi evaporado e o material foi ressuspendido em 25µl de TE e estocado a –20ºC.

4.6.3 - Resina

Ao tubo com o ovo foram adicionados 100µl de resina Instagene matrix ® Bio-Rad, e este foi incubado por 30min a 56ºC. O tubo foi agitado em vortex com alta velocidade por 10s, centrifugado a 13.000x g por 3min à temperatura ambiente, o sobrenadante foi retirado e estocado a –20ºC para posterior uso na PCR (a permanência do resíduo parece comprometer a estabilidade do DNA).

4.6.4 - Fenol/Clorofórmio

Ao tubo contendo o ovo ou verme, foram adicionados 100µl de tampão de extração (50mM Tris-HCl, pH 8,0, 50mM EDTA, 100mM NaCl, SDS 1%) e 50µg/ml de proteinase K. A amostra foi incubada a 37ºC por 12 a 16h. Foram adicionados 100µl de fenol, a amostra foi agitada em agitador de inversão, até formar uma emulsão e centrifugada a 13.000x g por 10min. A fase aquosa foi retirada, sendo transferida para um novo tubo, ao qual se acrescentaram 50l de fenol e 50l de clorofórmio-álcool isoamílico (24:1). A amostra foi agitada por 10min e centrifugada a 13.000x g por 10min. A fase aquosa foi retirada e transferida para novo tubo, ao qual foram adicionados 100µl de clorofórmio-álcool isoamílico. A amostra foi novamente agitada e centrifugada. A fase aquosa foi transferida para novo tubo ao qual foi acrescentado acetato de sódio 3M, pH 5,2, em volume igual a 1/10 do seu volume. Em seguida, para a precipitação do DNA, foram adicionadas 2,5 vezes o volume de etanol 100% a 4ºC. O tubo foi colocado a –70ºC por 1h. A amostra foi centrifugada a 13.000x g por 10 min a 4ºC. O sedimento foi lavado 2 vezes com etanol a 70% a 4ºC, centrifugando a 13.000x g por 10 min, a 4ºC. O sobrenadante etanólico foi eliminado e o tubo contendo o sedimento incubado, com a tampa aberta, a 37ºC para evaporar o etanol remanescente. O DNA foi ressuspendido em 25µl de tampão TE e estocado a –20ºC.

Devido às baixas quantidades de DNA obtidas, não foi possível uma quantificação, com o uso de nenhum dos métodos descritos, assim, testamos e utilizamos 4µl das preparações de DNA para 15µl de volume final na PCR.

4.7 – A PCR-multiplex

A PCR-multiplex é uma variação da PCR tradicional, na qual utilizamos simultaneamente vários pares de iniciadores, um par para cada um dos diferentes fragmentos que se queira amplificar. Inicialmente estabelecemos condições básicas de reação (concentração dos iniciadores, de *Taq* DNA polimerase, temperaturas e número de ciclos), para os pares de iniciadores isoladamente e depois em conjunto (multiplex). Os iniciadores que não geraram produtos foram retirados do multiplex. Nestas reações usamos, 0,15U de *Taq* DNA polimerase (Invitrogen), tampão da enzima, fornecido pelo fabricante (1,5mM de $MgCl_2$, Tris-HCl 10mM pH 8,0, KCl 50mM), 200µM de dNTP, 10pmoles de cada iniciador (Tabela 3), 4µl da preparação de DNA de ovo ou 1ng de DNA de verme adulto, e água deionizada para um volume total de reação de 15µl. Alternativamente, nos casos de não amplificação de algum *locus*, utilizamos a enzima AccuPrime® *Taq* DNA Polimerase (Invitrogen). Em todos os casos a PCR foi realizada em termociclador "PCR Express", Thermo Hybaid US.

Uma vez estabelecidas as condições de amplificação para cada par de iniciadores estes foram colocados nos devidos conjuntos e estabelecidas as novas condições de amplificação de acordo com Henegariu e cols., (1997). O protocolo final foi de 5min a 95ºC para desnaturação e 35 ciclos consistindo de 30s a 95ºC para desnaturação, 30s a 50ºC para anelamento dos iniciadores, 2 min a 65ºC para extensão, finalizando com uma extensão de 2min a 65ºC e mantido a 4ºC após o término da reação.

Tabela 3 - Lista dos *loci* analisados nas amostras de DNA de ovos coletados em fezes de moradores.

locus	Acesso	Repetição	Faixa (pb)	Seqüências (5'-3')	Referência
SmBr09	AF325694	(ATT)11	134/176	R: ATTGGCGTCAGTAGAAGAGATT	(Rodrigues et al., 2007)
				F: ATTCACCCATTGTCTTAAAACC	
SmBr15	AF325695	(GATA)10	431/483	R: TTGGATAAACTTAGTGACTTTTC	(Rodrigues et al., 2007)
				F: TATAGGACAAAACGCGGGTC	
SmBr16	L04480	(AT)10	319/345	R: GGCCTGATACAATTCTCCGA	(Rodrigues et al., 2007)
				F: TGTGACTTTGATGCCACTGA	
SmBr17	AQ841039	(TA)10	94/130	R: TGATCCTTTGTGCCAACA	(Rodrigues et al., 2007)
				F: CTGCAGGGGGAAATAGAAG	
SmBr19	C122522	(GATA)10	260/312	F: TAGATGATAGACAGATAGATCG	NP
				R: TAATTAAGCACACCAGCAAG	

Os números listados nas repetições são relativos aos encontrados nas primeiras seqüências obtidas destes *loci*, NP= não publicado

Para verificarmos o sucesso da PCR, 3µl dos produtos de amplificação foram submetidos à análise em géis de poliacrilamida a 8%. Os géis foram fixados em 150ml de solução de etanol a 10% com ácido acético 0,5% (v/v), e impregnados com nitrato de prata a 0,3%. Finalmente, os géis foram lavados em água deionizada e revelados em solução aquosa de hidróxido de sódio 3% (p/v) com 0,5% de formaldeído (v/v), até o aparecimento das bandas (Sanguinetti et al., 1994).

4.8 - Determinação do tamanho dos alelos de microssatélites

Para a determinação do tamanho dos alelos dos microssatélites amplificados, 1µl do produto da PCR foi desnaturado a 95°C por 3min com um tampão de amostra contendo 100% de formamida e 5mg/ml de Dextran Blue 2000, submetido à eletroforese em gel de poliacrilamida 6% desnaturante com 7M de uréia, utilizando-se tampão de corrida TBE 0,6X (1M de Tris, 0,83M de ácido bórico e 10mM de EDTA) em seqüenciador automático de DNA A.L.F - Automatic Laser Fluorescence (Pharmacia-LKB) por 3h. As condições de corrida foram: 1500V, 38mA, 34W, potência do laser 3, temperatura de 40ºC e amostragem a cada 0,84s. Também foram aplicados no gel 10fmoles de um padrão de tamanhos de fragmentos de 50pb a 500pb (Pharmacia) como referência de mobilidade dos fragmentos. Esta metodologia permitiu a análise simultânea de até 36 amostras em cada corrida. Os resultados obtidos foram

analisados com o programa AlleleLinks versão 1.00 (Pharmacia Biotech). Este programa calcula o tamanho de cada pico, tendo como referência o padrão de tamanho de fragmentos, e automaticamente nos dá o tamanho de cada alelo podendo compará-los com bancos de dados previamente montados.

4.9 – Análise dos dados obtidos com microssatélites: estimadores de genética de populações

Para as análises genéticas foi utilizado o programa Arlequin versão 3.0 (Excoffier et al., 2005). Para cada *locus* calculou-se o número de alelos, as freqüências alélicas observadas, as freqüências genotípicas observadas e os valores da heterozigose observada e esperada dentro do equilíbrio de Hardy-Weinberg.

No programa Arlequin a estrutura genética das populações é investigada pela análise de variância molecular (AMOVA), que é um método para estimar a proporção da variação devida a cada tipo de agrupamento das populações, e a partir daí permite testar hipóteses sobre estas diferenças (Excoffier et al., 1992).

Em nosso trabalho, para se verificar uma possível estruturação geográfica das populações estudadas, estas foram divididas em 5 grupos conforme sua localidade de origem (Card1, Card2, Card3, Suss e VdG). Para os testes de AMOVA foram feitos os seguintes agrupamentos: Grupos, as 5 subáreas (Card1, Card2, Card3, Suss, VdG), infrapopulações dentro dos grupos (cada paciente amostrado em cada subárea), indivíduos de cada população (cada ovo de uma infrapopulação), e todos os indivíduos (todos os ovos amostrados). Os componentes de variância em cada nível foram calculados e usados para computar uma série de estatísticas que resumem o grau de diferenciação entre as infrapopulações nos diversos níveis hierárquicos propostos (Figura 6). Estas estatísticas representam respectivamente os índices de fixação (FCT, FSC, FIS, e FIT) originariamente propostos por Wright (1965) em termos de coeficiente de endogamia e mais tarde por Slatkin (1991) em termos de tempo de coalescencia.

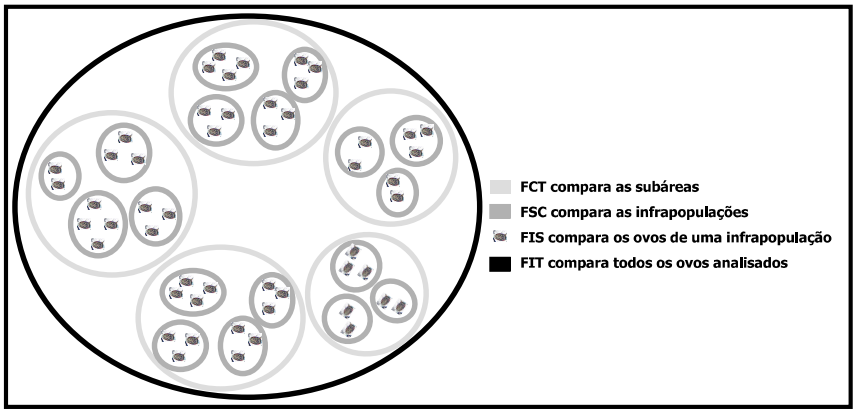


Figura 6 – Esquema representando os tipos de agrupamentos utilizados nos teste de AMOVA e os índices de fixação relativos a cada um.

Foi gerada ainda uma matriz de distâncias (FST) entre os pares de populações e estes dados foram agrupados via UPGMA – *Unweighet Pair Group Method Analysis* (Sneath & Sokal, 1962) e usados na construção de fenogramas utilizando-se o programa MEGA versão 3.1 (Kumar et al., 2004). Os fenogramas foram construídos com as 53 infrapopulações (número total de populações no estudo), com as 44 infrapopulações das quais tivemos apenas uma amostragem (pré ou pós-tratamento) e com os 9 pares de infrapopulações que foram analisados pré e pós-tratamento.

4.10 – Determinação do tamanho amostral

Para a determinação do tamanho amostral, ou o número mínimo de ovos que, com um nível de significância de 0,05, cobriria 80% do número provável de alelos por *locus* estudado, foram genotipados ovos coletados em fezes de 2 moradores e analisadas estatisticamente as possibilidades de ocorrência de 1, 2 ou nenhum alelo novo por ovo.

Foi analisado o tipo de distribuição dos eventos, distribuição Poisson, que é utilizada para descrever variáveis aleatórias que se expressam através de contagens como número de ocorrências (esta distribuição tem como característica a média e a variância serem iguais).

Logo: média = variância = μ

A fórmula é dada por: $$P(X) = \frac{e^{-\mu}\mu^{X}}{X!}$$

Onde: X = 0, 1, 2 alelos novos e μ = taxa média de ocorrências.

E a partir desta distribuição o tamanho amostral, cuja fórmula é dada por:

Tamanho amostral

$$n = \left(\frac{\sqrt{\mu}}{e}\right)^2 \text{ e } z = \left(b + 0{,}5 - \mu\right)\Big/\sqrt{\mu}$$

Onde: e = erro amostral, e b=0,1ou 2.

V – ARTIGOS & RESULTADOS

5 – ARTIGOS

5.1 – Artigo 1:

Microsatellite-enriched genomic libraries as a source of polymorphic *loci* for *Schistosoma mansoni*.

Publicado na revista Molecular Ecology Notes

Rodrigues NB, Silva SA, Pucci MM, Minchella DJ, Sorensen RE, LoVerde PT, Romanha AJ, Oliveira GC. Microsatellite-enriched genomic libraries as a source of polymorphic loci for *Schistosoma mansoni*. Molecular Ecology Notes. 2007. 7, 263-265

Este artigo apresenta a síntese dos resultados relativos ao objetivo específico 1:

1 Desenvolver marcadores genômicos polimórficos do tipo microssatélites para *S. mansoni.*

Neste trabalho demonstramos a utilização de bibliotecas enriquecidas como fonte de marcadores do tipo microssatélites para o estudo da estrutura genética de populações de *S. mansoni*. Adicionalmente, o trabalho apresenta outros 11 novos *loci* polimórficos que vêm se juntar aos cerca de 30 *loci* anteriormente descritos na literatura (Durand et al., 2000; Curtis et al., 2001; Rodrigues et al., 2002a; Rodrigues et al., 2002b; Silva et al., 2005).

Ainda com relação a estes objetivos, estamos em fase final de preparação de um manuscrito (Artigo 2) no qual descrevemos um *locus* de minissatélites (o primeiro não mitocondrial descrito em *Schistosoma mansoni*). Este *locus* apresenta uma porção interna variável composta por um microssatélite com o motivo repetitivo "CA". Neste trabalho descrevemos a utilização deste minissatélite, de seu microssatélite interno e de outros três *loci* de microssatélites na identificação de cepas de *S. mansoni* e na identificação de 9 diferentes espécies do gênero *Schistosoma*.

Molecular Ecology Notes (2007) 7, 263–265 doi: 10.1111/ j.1471-8286.2006.01575.x

PRIMER NOTE

Microsatellite-enriched genomic libraries as a source of polymorphic loci for Schistosoma mansoni

N. B. RODRIGUES,*M. R. SILVA,*M. M. PUCCI,*D. J. MINCHELLA,† R. SORENSEN,‡ P. T. LOVERDE,§ A. J. ROMANHA* and G. OLIVEIRA*¶

*Centro de Pesquisas René Rachou-FIOCRUZ, Av. Augusto de Lima 1715, 30190-002, Belo Horizonte, MG, Brazil, †Department of Biology, Purdue University, 915 W. State Street, West Lafayette, IN 47907, USA, ‡Department of Biology, Minnesota State University at Mankato, MN 56001, USA, §Southwest Foundation for Biomedical Research, PO Box 760549, San Antonio, TX 78245-0549, USA, ¶Programa de Pós-Graduação e Pesquisa, Santa Casa de Belo Horizonte, MG 30150-221, Brazil

Abstract

Microsatellite markers for Schistosoma mansoni were developed using four genomic microsatellite-enriched libraries. Microsatellites were observed in 65.4% of all sequences. Primer pairs were designed and tested for 23 loci. Eighteen loci produced amplification products, out of which 11 were polymorphic and were further characterized on 100 individuals of S. mansoni. Two to 19 alleles per locus were detected. The average values of expected and observed heterozygosities among the 11 loci were 0.79 and 0.59, respectively.

Keywords: computational biology, genome, genomic library, microsatellites, population genetics, Schistosoma mansoni

Received 10 July 2006; revision received 1 September 2006; accepted 1 September 2006

Schistosoma mansoni is the main causative agent of schistosomiasis, a human disease affecting over 200 million people (WHO 2002). The development of tools such as polymorphic microsatellite loci for the understanding of genetic variation and population structure of this parasite may improve our knowledge of disease transmission and allow the identification of genes of interest by linkage analysis. Approximately 30 polymorphic microsatellite markers, mostly derived from cDNAs, are available for S. mansoni (Durand et al. 2000; Blair et al. 2001; Curtis et al. 2001; Rodrigues et al. 2002). However, selective pressures on expressed genes may decrease microsatellite polymorphism (Stohler et al. 2004). We sequenced clones from genomic microsatellite-enriched libraries and identified new polymorphic microsatellite loci.

Four S. mansoni genomic libraries enriched for AAT, CA, GA and TAGA repetitive sequences were constructed and cloned into pUC19 (Genetics Information System). The insert sizes ranged from 300 to 700 bp. Recombinant clones were transformed into DH5α Escherichia coli, and selected on LB agar plates containing 800 μg X-gal, 800 μg IPTG, and 0.1 μg/ μL ampicillin (Invitrogen). Selected plasmids were purified from positive white colonies using the R.E.A.L. prep 96 plasmid Kit (QIAGEN). Three hundred and eighty-two clones were sequenced (automated ALF sequencer, Pharmacia) using the Thermo Sequenase Fluorescent Primer Kit (GE Healthcare).

The sequences were clustered with CAP3 (Huang & Madan 1999) and microsatellite repeats were identified using RepeatMasker (http:// repeatmasker.org). Microsatellites (di- to hexanucleotides) were observed in 65.4% of the sequences and 31.29% were perfect, 57.31% imperfect and 11.40% compound. Sequences were grouped into 24 clusters containing two to 35 sequences and 71 singlets. A complete description of sequences, BLAST hits, primers and clusters can be found at bioinfo.cpqrr.fiocruz.br/ microsat.

GenBank Accession numbers for the sequences are: AAT library, DQ137430 to DQ137590; CA library, DQ137591 to DQ137640; GA library, DQ137641 to DQ137722; and TAGA library, DQ137723 to DQ137796 and DQ514536. Sequences presenting the longest perfect repetitions and flanking regions were selected, and polymerase chain reaction (PCR) primers were designed using oligos version 3.6 (www.biocentre.helsinki.fi/ bi/ programs/ fastpcr.htm).

Correspondence: G. Oliveira. Fax: +55-31-3349-7785; E-mail: oliveira@cpqrr.fiocruz.br

Table 1 Characteristics of 11 Schistosoma mansoni microsatellite loci*. A total of 100 individuals were tested for each loci, except for loci SmBr12, SmBr13 and SmBr14, for which 386 individuals were tested

Locus	Accession no.	Repeat	Size range	Primer sequences (5′–3′)	AT (°C)	A	H_O	H_E
SmBr7	DQ137434	$(AAT)_9$	128–172	R: AATCACCAATGGCAACAATCTG F: CGTCATCACCTTAAACATGAAC	62	7	0.69**	0.82
SmBr8	DQ448292	$(ATT)_{12}$	246–255	R: ATGCCCACACACAAAGTAAA F: TAGGACAGGTTTTCCACCAA	56	4	0.53	0.60
SmBr9	DQ137431	$(TTA)_{11}$	149–164	R: ATTGGCGTCAGTAGAAGAGATT F: ATTCACCCATTGTCTTAAAACC	60	8	0.43	0.82
SmBr10	DQ448293	$(GATA)_{10}$	110–138	R: GTACATTTTATGTCAGTTAGCC F: CATGATCTTAGCTCAGAGAGC	60	7	0.69	0.85
SmBr11	AC112150.4	$(TCTA)_8$	206–214	R: TTCAGTCCCTGGAACACACA F: AAGAAGTGGAGGAGGCCTTT	60	2	0.23	0.58
SmBr12	DQ137724	$(CTAT)_{23}$	228–316	R: AGTAAAAACTATCCTATCCATTTCTATTG F: TATATAGCAAAAGTAGTCTATATTCGTAGC	60	19	0.94	0.87
SmBr13	DQ137790	$(CTAT)_{16}$	206–254	R: ACTCCCCAGCAATTTGTCC F: GTCACAGATACCTGACGAGCTG	60	13	0.86	0.88
SmBr14	DQ514536	$(TAGA)_9$ TGGA$(TAGA)_{19}$	268–308	R: TCTATGTATCTACCCCACCCTATC F: CTGCTCATCATAGAAGTGTGGC	60	11	0.84	0.83
SmBr15	AF325695	$(TAGA)_{10}$	450–474	R: TTGGATAAACTTAGTGACTTTTC F: TATAGGACAAAACGCGGGTC	60	8	0.51	0.82
SmBr16	L04480	$(TA)_{10}$	319–345	R: GGCCTGATACAATTCTCCGA F: TGTGACTTTGATGCCACTGA	60	9	0.39	0.77
SmBr17	AQ841039	$(TA)_{10}$	100–130	R: TGATCCTTTGTGCCAACA F: CTGCAGGGGGAAATAGAAG	60	16	0.41	0.90

AT, annealing temperature; A, number of alleles; observed (H_O) and expected (H_E) heterozygosities, *There is no indication of the presence of null alleles, except for SmBr7 that exhibited a significant deviation from Hardy–Weinberg equilibrium. **HWE significant for $P < 0.05$. Reverse (R) and Forward (F) primers. Reverse primers were labelled with fluorescein.

Primer pairs for 23 loci were tested on 13 adult S. mansoni worms of the LE strain. Each PCR was performed in a 10-μL volume containing 1.5 mm $MgCl_2$, 50 mm KCl and 10 mm Tris/ HCl, pH 8.3, 1 pmol of each primer, 200 μm of each dNTP, 0.75 U of Taq DNA polymerase (CENBIOT), and 1 ng of DNA. Thermal cycling was carried out in a PerkinElmer 9600 thermal cycler for 3 min at 95 °C; 35 cycles of 45 s at 95 °C, 1 min at primer-specific annealing temperature (Table 1), 30 s at 72 °C; and a final extension of 5 min at 72 °C. The amplicons were separated on 6% polyacrylamide 7 m urea gels using an ALF sequencer. Allele sizes were assessed using Allelinks (Amersham).

PCR products were observed in 18 of 23 loci tested and 11 were polymorphic. The polymorphic loci were further characterized on 100 S. mansoni worms from field isolates. Loci SmBr12–14 were tested on 386 individuals. DNA was extracted from single worms using the phenol/ chloroform extraction protocol as described in Rodrigues et al. (2002). Tests for deviations from Hardy–Weinberg equilibrium, observed (H_O) and expected (H_E) heterozygosity estimations, and linkage disequilibrium were performed using arlequin 3.0 (Schneider et al. 2000). The loci presented two to 19 alleles per locus, and H_E ranged from 0.58 to 0.90 (Table 1). Loci SmBr7 and SmBr8 were found to be linked. All loci, but SmBr12 and SmBr14, showed lower than expected H_O, albeit nonsignificant, except for locus SmBr7 ($P < 0.05$) (Table 1).

The use of microsatellite-enriched genomic libraries generated a much higher percentage of microsatellite-containing sequences (65.4%) compared to approximately 3% found by Rodrigues et al. (2002) in expressed sequence tags (ESTs) or BAC-end sequences and 1% found by Durand et al. (2000) by screening a partial genomic library using CA and GA oligonucleotide probes. In addition, 25% of the amplified loci were shown to be polymorphic. It is noteworthy that 80% of our sequences are novel in relation to the S. mansoni ESTs, and 91 (24%) were not detected in the TIGR and SANGER genomic databases (El Sayed et al. 2004). New sequences can aid in filling the gaps in the final genome assembly. The major limitation of using microsatellite-enriched genomic libraries is the elevated price of library construction. However, the approach used yielded over 60 times more microsatellite-containing sequences, 25% of which were shown to be polymorphic. The information presented here will be useful for the generation of additional markers for genetic studies, and mapping efforts and for genome gap filling.

Acknowledgements

This investigation received financial assistance from WHO/ TDR grant A20399, CNPq 521108/ 01-2, NIH-Fogarty (TW007012-01), NIH (AI042768-05A2) and FAPEMIG (407/ 02).

References

Blair L, Webster JP, Barker GC (2001) Isolation and characterization of polymorphic microsatellite markers in Schistosoma mansoni from Africa. Molecular Ecology Notes, 1, 93–95.

Curtis J, Sorensen RE, Page LK, Minchella DJ (2001) Microsatellite loci in the human blood fluke Schistosoma mansoni and their utility for other schistosome species. Molecular Ecology Notes, 1, 143–145.

Durand P, Sire C, Theron A (2000) Isolation of microsatellite markers in the digenetic trematode Schistosoma mansoni from Guadeloupe Island. Molecular Ecology, 9, 997–998.

El Sayed NM, Bartholomeu D, Ivens A, Johnston DA, LoVerde PT (2004) Advances in schistosome genomics. Trends in Parasitology, 20, 154–157.

Huang X, Madan A (1999) CAP3: A DNA sequence assembly program. Genome Research, 9, 868–877.

Rodrigues NB, Coura-Filho P, Souza CP et al. (2002) Populational structure of Schistosoma mansoni assessed by DNA microsatellites. International Journal for Parasitology, 32, 843–851.

Schneider S, Roessli D, Excoffier L (2000) ARLEQUIN: A Software for Population Genetics Data Analysis. Version 2.000. Genetics and Biometry Laboratory, Department of Anthropology, University of Geneva.

Stohler RA, Curtis J, Minchella DJ (2004) A comparison of microsatellite polymorphism and heterozygosity among field and laboratory populations of Schistosoma mansoni. International Journal for Parasitology, 34, 595–601.

WHO (2002) Prevention and control of schistosomiasis and soil-transmitted helminthiasis. World Health Organization Technical Report. Series 912, i–57.

5.2 – Artigo 2:

Identification of a first nuclear genome minisatellite, and 4 microsatellites *loci* in 9 different species of *Schistosoma.*

Diana Bahia, Nilton B. Rodrigues, Flávio M. G.Araújo, Alvaro Romanha, David Johnston & Guilherme Oliveira.

Artigo em fase final de preparação e que também apresenta resultados relativos ao objetivo específico 1:

Desenvolver marcadores genômicos polimórficos do tipo microssatélites para *S. mansoni.*

Neste trabalho descrevemos um *locus* de minissatélites, o primeiro não mitocondrial descrito em *Schistosoma mansoni*, o qual apresenta uma porção interna variável composta por um microssatélite com o motivo repetitivo "CA". Mostramos também a utilização deste minissatélite, de seu microssatélite interno e de outros três *loci* de microssatélites na identificação de cepas de *S. mansoni* e na identificação de 9 diferentes espécies do gênero *Schistosoma.*

Identification of a first nuclear genome minisatellite, and 4 microsatellites *loci* in 9 different species of *Schistosoma.*

Diana Bahia[1,2], Nilton B. Rodrigues[1], Flávio M. G.Araújo[1], Alvaro Romanha[1], David Johnston[3] & Guilherme Oliveira[1,4*]

[1]*Centro de Pesquisas René Rachou (CPqRR) – FIOCRUZ, Av. Augusto de Lima 1715, Belo Horizonte, MG, 30190-002, Brazil.*

[2] *Departamento de Microbiologia, Imunologia e Parasitologia, Escola Paulista de Medicina, UNIFESP, Rua Botucatu 862, 6º andar, 04023-062, São Paulo, SP, Brazil.*

[3] *The Natural History Museum, London, England*;

[4]*Programa de Pós-Graduação e Pesquisa, Santa Casa de Belo Horizonte, Av. Francisco Sales 1111, 9º andar, Ala C, Belo Horizonte, MG, 30150-221, Brazil*

*Corresponding author (E-mail:), Centro de Pesquisas René Rachou (CPqRR) – FIOCRUZ, Av. Augusto de Lima 1715, Belo Horizonte, MG, 30190-002, Brazil, Telephone: +55-31-3349-7785, FAX: +55-31-3295-3115

ABSTRACT

Here we describe the identification of the first non-mitochondrial minisatellite (so-called CA88 minisatellite) in the blood fluke parasite *Schistosoma*. It is present in a coding region of 9 *Schistosoma* species, three of which are of medical importance. The CA88 minisatellite was screened by BLAST searches against *S. mansoni* sequences from public data bases and by PCR in 10 individuals of each species analyzed. The minisatellite profiles obtained with the *Schistosoma* species using CA88 primers were classified into four genotypes, three of them are in accordance with the species' groups. Databases searches for similarity established links between CA88 minisatellites and a eggshell protein and other hypothetical *S. mansoni* protein. At present, CA88 has 903 nucleotides and at least three unity repeats of approximately 350 bp. In addition, a panel of 4 microsatellites *loci* has been characterized for all *Schistosoma* species. We have observed inter- and intra-specific differences among *Schistosoma* species. Each species has been identified and classified on the basis of a fingerprint panel provided by mini- and microsatellites analyses. Taken together, these results encourage the use of these markers to investigate the population dynamics of *Schistosoma* species of medical and non-medical importance in endemic areas as well as to answer a number of questions on the relationships among different populations of parasites.

Introduction

The presence of some types of repetitive DNA in many parasites has allowed the characterization of population structures as well as phylogenetic studies (1). As it has become clear that a substantial level of intraspecific variation exists within schistosome populations (2), interest has grown in studying the way in which that variation is partitioned in natural populations. A minisatellite region in the mitochondrial DNA of *Schistosoma mansoni*, that has high homology with transcribed nuclear sequences, was identified (3). That minisatellite repeat is composed of a large 558-pb subunit and a variable tandem array of the small 62-bp unit. Polymorphic DNA sequences, as the mtDNA, have been valuable for determining how schistosome genotypes are distributed among intermediate hosts (4-7) or even for examine *S. mansoni* population structure and subdivision in definitive hosts (8). Currently, microsatellites are the chosen markers used in population genetic studies of this parasite (9;10).

Micro- and minisatellites occur both within coding and non-coding regions of genomes (11). The polymorphism and wide genomic distribution make microsatellites one of the most useful genetic markers available for use in typing individuals (12) or population studies (13), and for constructing genetic maps to identify *loci* involved in genetic diseases (14). Micro- and minisatellites have become key elements to distinguish individuals in several eukaryotic species such as *Plasmodium* (15), *Trypanosoma brucei* (16), *T. cruzi* (17) and *Theileria parva* (18). Microsatellites are short (2–6 bp) tandenly repeated sequences, whereas minisatellites are longer (8–100 bp) tandenly repeated sequences (19;20). Both occur abundantly and randomly over eukaryotic genomes (1). The mutation mechanisms in micro- and minisatellites origins are nevertheless different. Microsatellite mutations tend to result from replication slippage and repair, whereas minisatellite mutations mechanisms include both meiotic crossing over and replication slippage (21). These mutations lead to high levels of polymorphism which make micro- and minisatellites ideal for high resolution molecular fingerprinting.

Several microsatellite markers were identified and developed for *S. mansoni* and these markers have proved to be highly useful for genetic and population studies of schistosome (10;22-26). So far, only one minisatellite, located on the mitochondrial DNA (3), has been identified in *S. mansoni.*

Here we describe the identification of the first nuclear genome schistosome minisatellite (so-called CA88 minisatellite) present in a genomic coding region of 9 different *Schistosoma* species, three of which of medical importance. The minisatellite presents inter- and intra-specific differences among *Schistosoma* species and the identification and classification of these species on the basis of a fingerprint panel provided by mini- and microsatellites analyses.

2. Material and Methods

2.1- Schistosoma species

Adult worms of 9 different *Schistosoma* species, kindly provided by Dr David Johnston from The Natural History Museum, London, England, were used in this study; *S. haematobium, S. margrebowiei*, *S. mattheei, S. bovis, S. intercalatum* and *S. curassoni,* from *S. haematobium* group, *S. mansoni and S. rodhaini* from *S. mansoni* group and *S. japonicum* from *S. japonicum* group.

2.2- DNA extraction

DNA was extracted from 10 individuals of each species. The worms were washed in phosphate buffered saline (PBS), placed, individually, in 1.5 ml microcentrifuge tubes and stored at 8^0C till DNA extraction. Before extraction the worms were pulverized on dry ice bath alternated with warm bath 5 times. The individual worms were resuspended in 50 ml of 50 mM Tris–HCl, pH=8.0, 100 mM ethylenediaminetetra-acetic acid (EDTA), 100 mM NaCl, and 0.5% sodium dodecyl sulphate (27) and incubated in the presence of 20 μg/ml proteinase K at 37^0C, overnight. Standard phenol/chloroform and ethanol precipitation (28) was used to pellet the DNA from the supernatant. DNA was resuspended in 30 μl of 10 mM Tris–HCl and 1 mM EDTA, pH 8.0 (TE).

2.3- Microsatellite amplification and detection

Microsatellite analyses were performed by PCR using 100 ng of template DNA, 0.75U of *Taq* DNA polymerase (Invitrogen), 1x PCR buffer (1.5mM $MgCl_2$, 50mM KCl,

10mM Tris/HCl pH 8.3), 10μM of each primer and 200μM of each dNTP. PCR amplifications were carried out in a Perkin Elmer model 9600 thermal cycler in a final reaction volume of 10μl. The cycling conditions were: one cycle of denaturation at 95°C for 3 min, followed by 35 cycles of denaturation at 95°C for 45 sec, 1 min for annealing and 30 sec at 72°C. Four microsatellite *loci* were used in this analyses SmBr5, SmBr6 (25), SmBr9 (10) e SmBr18, derived from CA88 minisatellite (ACC# DQ137431). For SmBr18 PCR primers were designed using the software Fast PCR (29).The annealing temperature varied from 55°C to 60 °C (Table 1). To check amplification, PCR products electrophoresis was carried out in an 8% polyacrylamide gel employing a Mini-Protean II apparatus (Bio-Rad, Hercules, CA) using 3 μl of the amplified products. The gel was then silver stained (30).One primer of each pair was labelled at the 5`end with fluorescein allowing alleles scoring using an ALF automated sequencer and the AlleleLinks software (Amersham), as previously described (10;25).

2.4- Minisatellite computational analyses

Bioinformatics tools were used to search for repetitive sequences in *Schistosoma* genome. DQ137431 sequence was analyzed in BLAST searches against NCBI, TIGR, *S. mansoni* geneDB version 4.0 (Sanger), and Sanger *S. mansoni* genome project databases.

3. Results

3.1. Microsatellite variability among 9 Schistosoma species

A panel of four microsatellites *loci*, all of them designed based on *S. mansoni* sequences (Table 1), was tested in DNA worms from 9 different species. All species presented amplification products for SmBr18 *loci* and *S. intercalatum* and *S. curassoni* showed amplification products only for this *locus*. In the *loci* that had been previously described for *S. mansoni* we observed from 1 to 7 alleles per amplified *locus* among the species (Table 2).

3.2. SmBr18 is a microsatellite within a minisatellite region in all 9 Schistosoma species

Locus SmBr18 presented a multi-band profile and BLAST search showed it was inserted in a minisatellite unity (CA88). It is present in a coding region of 9 studied *Schistosoma* species, three of which of medical importance (*S. japonicum, S. intercalatum* and *S. haematobium*) in addition to *S. mansoni*.

SmBr18 is a 364bp *S. mansoni* DNA sequence derived from a partial *S. mansoni* genomic library enriched for CA (Fig 1). BLAST searches against *S. mansoni* sequences from NCBI databases showed similarity with sequences DQ137590.1, DQ137585.1, DQ137520.1, DQ137504.1, DQ137567.1, DQ137605.1, DQ137539.1, DQ137537.1, DQ137526.1, DQ137489.1, DQ137525.1, DQ137461.1, DQ137466.1 (10). It also showed 84% similarity with TC34704 EST, a Eggshell protein precursor (Chorion protein) at TIGR database, 94% similarity, BlastN, 93% similarity in BlastX with Sm02551 (a putative hypothetical protein of *S. mansoni*) at Sanger *S. mansoni* geneDB and 90% similarity with shisto4743c05.p1k sequence, at Sanger databases. Those sequences were clustered using CAP3 (31) and the consensus sequence screened against the same databases. Those alignments showed that SmBr18 is inserted in a repetitive unity of a minisatellite.

This minisatellite has two different repetitive unities, one of about 70bp and other one of about 360bp. The consensus sequence has 9,030bp and presents 8 repetitions unities, three of the short type and five of the long type. The short type repetition presents a (CA)n microsatellite located at position 35 that varied from 7 to 14 repetitions. The long type repetition presents that same microsatellite and another one at position 205 that varied from 10 to 12 repetitions.

SmBr18 PCR primers were designed to amplify a fragment of about 145bp in the long microsatellite repetition. Those primers were applied in PCR using DNA samples from 10 individuals of each one of 9 different species. We observed a band pattern in PAGE gels consistent with the presence of more than one amplification site in all the species tested, indicating the presence of CA88 minisatellite in all of these species (Fig 2). *S. mansoni*, *S. rodhaini* and *S. japonicum* presented an extra band shorter than that expected, it could be explained by an amplification of CA88 short sub-unit (Fig 2).

3.3. CA88 minisatellite provides four different Schistosoma genotypes

The multiple band profiles of *Schistosoma* species on polyacrylamide gel (Fig 2A and B) obtained by PCR with SmBR18 primers were visually scored and analyzed for polymorphism based on the presence or absence of bands according to (32).

The species tested for the presence of CA88 minisatellite were classified into four genotypes, (Fig 2 A and B and Table 2). With the sole exception of *S. margrebowiei*, all the genotype profiles for the others were in complete accordance with the species group they are traditionally placed in (33).

In Figure 2B we estimate the apparent molecular weight of the CA88 minisatellite bands. Lanes 1 and 2 show the pattern obtained with *S. mansoni* and *S. rodhaini*, respectively. Both species present bands of around 141, 158, 210 and 240 kb, besides bands of 400 kb or more. Both profiles are identical, and were define as genotype I. Lanes 3, 4, 6, 7 and 8 correspondent to *S. haematobium, S. intercalatum, S. bovis, S. curassoni and S. mattheei,* respectively, show a profile of bands of 158, 280, 520 and 750 kb and were named as genotype II. All species are in *S. haematobium* group (Table 2). Lane 9 shows the profile of *S. japonicum* that corresponds to the *S. japonicum* group and the genotype IV, besides bands 141 and 158 also present in the *S. mansoni* group we could observe a banding pattern ranging from about 237 and 280 and a band of 500 kb, approximately. Curiously, only *S margrebowiei* presents a different profile from the others of the same group (*S. haematobium* group). Besides the 158 kb band, it clearly has a ladder of bands ranging from 158 to 280 kb.

As one primer of each pair was fluorescein labelled, allowing allele scoring, each species has been identified and classified on the basis of a fingerprint panel provided by mini- and microsatellite analyzes (Table 2). Figure 3 shows an example of SmBr18 microsatellite amplified from one individual of each different species and analyzed in the ALF automated sequencer. We observed peaks around 150bp, 300bp and 500bp indicating a minisatellite pattern and the extra band for *S. mansoni*, *S. rodhaini* and *S. japonicum* species.

4. Discussion

Here we describe the identification of the first non-mitochondrial minisatellite, CA88, in parasites of *Schistosoma* genus. It is present in a genomic coding region of 9 different *Schistosoma* species, three of which are of medical importance. The minisatellite profiles obtained with the *Schistosoma* species using SmBr18 primers were classified into four genotypes, three of them are in accordance with the traditional species' groups previously delineated by (33) Besides the alignment of CA88 with different contigs, in SangerDB, no link has yet been established between this minisatellites and a specific gene. At present, CA88 has 9,030 nucleotides and at least three repeat units of approximately 360 bp. In addition, a panel of 3 microsatellite *loci* has been characterized for all *Schistosoma* species. CA88 genotypes showed to be a useful tool to distinguish different *Schistosoma* groups. Together with the microsatellite *loci* here presented we have observed inter- and intra-specific differences among *Schistosoma* species and each species has been identified and classified on the basis of a fingerprint panel provided by mini- and microsatellites analyses, providing additional useful tool to assess population diversity within and between isolates.

In the *S. haematobium* species group are include *S. haematobium*, *S. intercalatum*, and the recently described *S. guineensis* (34), all human parasites, and *S. margrebowiei*, *S. leiperi*, *S. mattheei*, *S. bovis* and *S. curassoni*, predominantly parasitizing livestock (artiodactyls). These facts make this group, as a whole, a group of immense medical and veterinary importance. Undoubtedly, *S margrebowiei* might retain some characteristics that could separate this specie from the other members of the group. Webster et al (35), showed that the use of entire nuclear ssrDNA and lsrDNA and partial mitochondrial cox1 does not recognize *S. margrebowiei* and *S. mattheei* as sister taxa as was found by Kane et al. (36) using only cox1 sequences. Instead, *S. margrebowiei* is placed as a sister taxa to all other species within the group. Other interesting feature that put *S. margrebowiei* apart of the other species in this group is the fact that it is the unique that does not make hybrids within the group (37). In our mini satellite analysis, *S margrebowiei* exhibited a different profile when compared to other members of its group.

Taken together, these results encourage the use of these and the search for additional markers to investigate the population dynamics and phylogeny of *Schistosoma* species of medical and non-medical importance in endemic areas as well as to answer a number of questions on the relationships among different populations of parasites which can not be addressed with current methodologies.

Reference List

1. Oliveira,E.D., Pádua,J.G., Zucchi,M.I., Vencovsky,R., and Vieira,M.L.C. (2006) *Gen and Mol Biol* **29**, 294-307.

2. Curtis,J. and Minchella,D.J. (2000) *Parasitology Today* **16**, 68-71.

3. Pena,H.B., Souza,C.P., Simpson,A.J.G., and Pena,S.D.J. (1995) *Proc,Natl.Acad.Sci.USA* **92**, 915-919.

4. Minchella,D.J., Lewis,F.A., Sollenberger,K.M., and Williams,J.A. (1994) *Mol.Biochem.Parasitol.* **68**, 307-313.

5. Dabo,A., Durand,P., Morand,S., Diakite,M., Langand,J., Imbert-Establet,D., Doumbo,O., and Jourdane,J. (1997) *Acta Tropica* **66**, 15-26.

6. Eppert,A., Lewis,F.A., Grzywacz,C., Coura-Filho,P., Caldas,I., and Minchella,D.J. (2002) *J.Parasitol.* **88**, 232-236.

7. Sire,C., Durand,P., Pointier,J.P., and Theron,A. (2001) *Int.J.Parasitol.* **31**, 1609-1616.

8. Curtis,J., Fraga,L.A., Souza,C.P., Correa-Oliveira,R., and Minchella,D.J. (2001) *J.Heredity* **92**, 248-253.

9. Sorensen,R.E., Rodrigues,N.B., Oliveira,G., Romanha,A.J., and Minchella,D.J. (2006) *Parasitology* **133**, 443-451.

10. Rodrigues,N.B., Silva,S.A., Pucci,M.M., Minchella,D.J., Sorensen,R.E., LoVerde,P.T., Romanha,A.J., and Oliveira,G.C. (2007) *Mol Ecol Notes* **7**, 263-265.

11. Macleod,A. (2004) *Methods Mol.Biol.* **270**, 187-202.

12. Hagelberg,E., Gray,I.C., and Jeffreys,A.J. (1992) *Nature* **352**, 427-429.

13. Jarne,P. and Lagoda,J.L. (1996) *Trends in Ecology and Evolution* **11**, 424-429.

14. Dietrich,W.F., Miller,J., Steen,R., Merchant,M.A., Damron-Boles,D., Husain,Z., Dredge,R., Daly,M.J., Ingalls,K.A., O'Connor,T.J., and . (1996) *Nature* **380**, 149-152.

15. Collins,H.E., Li,H., Inda,S.E., Anderson,J., Laiho,K., Tuomilehto,J., and Seldin,M.F. (2000) *Hum.Genet.* **106**, 218-226.

16. Macleod,A., Tweedie,A., Welburn,S.C., Maudlin,I., Turner,C.M., and Tait,A. (2000) *Proc.Natl.Acad.Sci.U.S.A* **97**, 13442-13447.

17. Macedo,A.M., Martins,M.S., Chiari,E., and Pena,S.D.J. (1992) *Mol.Biochem.Parasitol.* **55**, 147-154.

18. Oura,C.A., Odongo,D.O., Lubega,G.W., Spooner,P.R., Tait,A., and Bishop,R.P. (2003) *Int.J.Parasitol.* **33**, 1641-1653.

19. Hamada,H., Petrino,M.G., and Kakunaga,T. (1982) *Proc.Natl.Acad.Sci.U.S.A* **79**, 6465-6469.

20. Tautz,D. and Renz,M. (1984) *Nucleic Acids Research* **12**, 4127-4139.

21. Armour,J.A., Brinkworth,M.H., and Kamischke,A. (1999) *Mutat.Res.* **445**, 73-80.

22. Blair,L., Webster,J.P., and Barker,G.C. (2001) *Molecular Ecology Notes* **1**, 93-95.

23. Durand,P., Sire,C., and Theron,A. (2000) *Mol.Ecol.* **9**, 997-998.

24. Prugnolle,F., de Meeus,T., Durand,P., Sire,C., and Theron,A. (2002) *Mol.Ecol.* **11**, 1231-1238.

25. Rodrigues,N.B., Coura-Filho,P., Souza,C.P., Jannoti Passos,L.K., Dias-Neto,E., and Romanha,A.J. (2002) *Inter.J.Parasitol.* **32**, 843-851.

26. Rodrigues,N.B., LoVerde,P.T., Romanha,A.J., and Oliveira,G. (2002) *Mem.Inst.Oswaldo Cruz* **97 Suppl 1**, 71-75.

27. Grossman,Z., Ram,D., Markovics,A., Tarrab-Hazdai,R., Lantner,F., Ziv,E., and Schechter,I. (1990) *Experimental Parasitology* **70**, 62-71.

28. Sambrook,J., Fritsch,E.F., and Maniatis,T. (1989) Molecular Cloning: a laboratory manual, Cold Spring Harbor Laboratory Press, New York, NY-USA.

29. Kalendar R. Fast PCR. (3.6). 2003. Finland., Institute of Biotechnology. University of Helsinki.
Ref Type: Computer Program

30. Sanguinetti,C.J., Dias-Neto,E., and Simpson,A.J.G. (1994) *BioTechniques* **17**, 915-918.

31. Huang,X. and Madan,A. (1999) *Genome Research* **9**, 868-877.

32. Zalloum,L., Gomes,M.L., Kinoshita,A.T., Toledo,M.J., Prioli,A.J., and de Araujo,S.M. (2005) *Exp.Parasitol.* **111**, 55-58.

33. Rollinson,D. and Southgate,V.R. (1987) *in* The biology of schistosomes: from genes to latrines. (Rollinson,D. and Simpson,A.J.G., Eds.), pp. 1-49 Academic Press, London.

34. Webster,B.L., Southgate,V.R., and Littlewood,D.T. (2006) *Int.J.Parasitol.* **36**, 947-955.

35. Webster,B.L., Southgate,V.R., and Littlewood,D.T. (2006) *Int.J.Parasitol.* **36**, 947-955.

36. Kane,R.A., Southgate,V.R., Rollinson,D., Littlewood,D.T., Lockyer,A.E., Pages,J.R., Tchuem Tchuente,L.A., and Jourdane,J. (2003) *Parasitology* **127**, 131-137.

37. Webster,B.L., Southgate,V.R., and Littlewood,D.T. (2006) *Int.J.Parasitol.* **36**, 947-955.

Legends to Figures

Figure 1- Global alignment of both, short and long sub-units of CA88 minisatellite. In bold the primers and underlined CA repetitions.

Figure 2- MW– Molecular Weight (50bp), *1– Schistosoma mansoni, 2– S. rodhaini, 3– S. haematobium, 4– S. intercalatum, 5– S. margrebowiei, 6– S. bovis, 7– S. curassoni, 8–S. mattheei, 9– S. japonicum*

Figure 3- Electropherogram showing analyses of the microsatellite *locus* SmBr18 in 9 individual worms from different species: S - a 50 to 500bp sizer. Lines 1 - *Schistosoma rodhaini*, 2 - *S. mansoni*, 3 - *S. haematobium*, 4 - *S. intercalatum*, 5 - *S. japonicum*, 6 - *S. margrebowiei*, 7 - S. *curassoni,* 8 – *S. bovis*, 9 - *S. mattheei.* Squares pick representing the amplification products of each minisatellite sub-unit.

Table 1: Characteristics of analyzed *S. mansoni* microsatellite *loci*.

Locus	Accession no.	Repeat	Primer sequences (5'-3')[a]	Amplicon	AT (°C)	Ref.
SmBr5	L25065	(ATT)7	GAATTACTGTCCCTTTATCTC *F*-AAACTATTCATTACTGTCGGG	328bp	58	(Rodrigues et al, 2002)
SmBr6	AF009659	(CTT)10	CTTAACAGACATACACGC *F*-GAATACAGGCTATAATCTACA	265bp	55	(Rodrigues et al, 2002)
SmBr9	DQ137431	(TTA)11	ATTGGCGTCAGTAGAAGAGATT *F*-ATTCACCCATTGTCTTAAAACC	161bp	60	(Rodrigues et al., 2007)
SmBr18	DQ448293	(CA)10	TTTTCTGTCTACATGTTGATGAAG *F*-TAACCATCATTCACCAAACATTC	141bp	60	-

a: The reverse primer of each pair is 5' Fluorescein labelled (*F-)*.

Table 2: Schistosoma specie's profile for three different microsatellite *loci*

Species	SmBr5 (ATT)n Alleles	SmBr6 (CTT)n Alleles	SmBr9 (TTA)n Alleles	Genotypes
**S. haematobium*	448	257	149, 152	II
S. margrebowiei	445,448,451	254, 257, 260	146, 149, 152	III
S. mattheei	448,451,460, 463, 466	NA	158	II
S. bovis	445, 448, 457	NA	146	II
S. intercalatum	NA	NA	NA	II
S. curassoni	NA	NA	NA	II
**S. mansoni*	451, 454, 457, 463, 475, 478, 481	254, 257, 260, 263, 266, 272	146, 149, 152, 158	I
S. rodhaini	445	263	146, 152	I
**S. japonicum*	NA	254, 260	NA	IV

*specie's groups
NA= not amplified

Figure 1

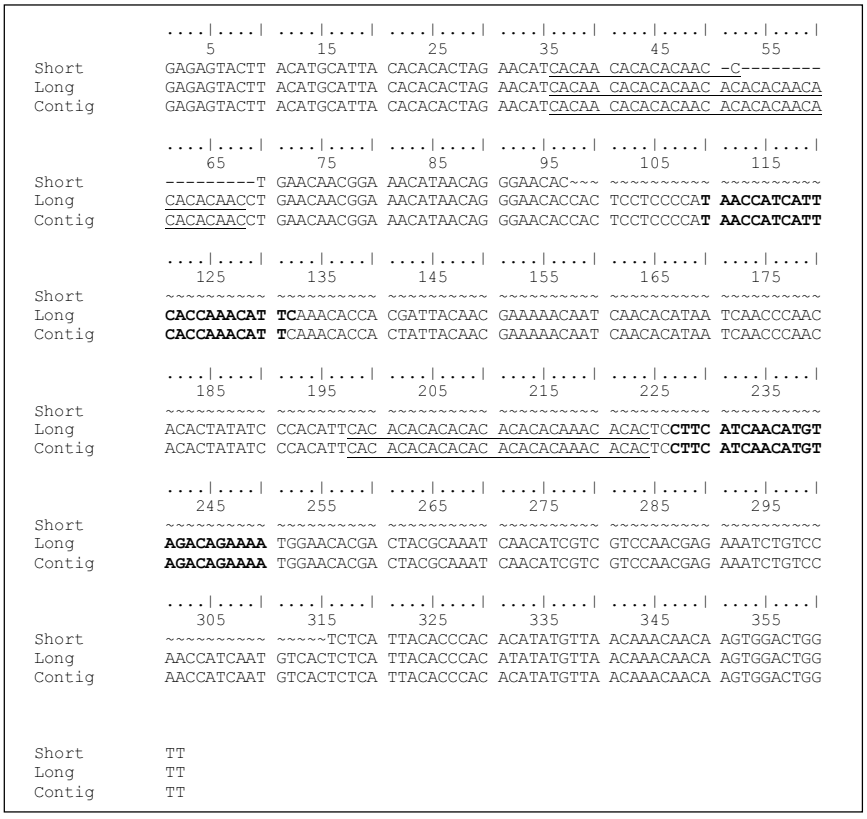
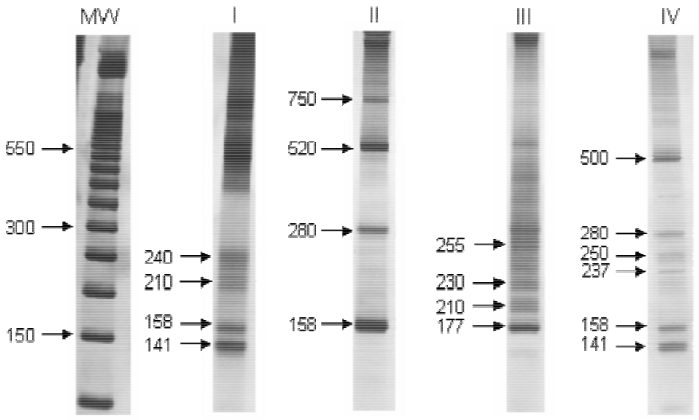

....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
5 15 25 35 45 55
Short GAGAGTACTT ACATGCATTA CACACACTAG AACATCACAA CACACACAAC -C--------
Long GAGAGTACTT ACATGCATTA CACACACTAG AACATCACAA CACACACAAC ACACACAACA
Contig GAGAGTACTT ACATGCATTA CACACACTAG AACATCACAA CACACACAAC ACACACAACA
....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
65 75 85 95 105 115
Short ---------T GAACAACGGA AACATAACAG GGAACAC~~~ ~~~~~~~~~~ ~~~~~~~~~~
Long CACACAACCT GAACAACGGA AACATAACAG GGAACACCAC TCCTCCCCAT AACCATCATT
Contig CACACAACCT GAACAACGGA AACATAACAG GGAACACCAC TCCTCCCCAT AACCATCATT
....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
125 135 145 155 165 175
Short ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
Long CACCAAACAT TCAAACACCA CGATTACAAC GAAAAACAAT CAACACATAA TCAACCCAAC
Contig CACCAAACAT TCAAACACCA CTATTACAAC GAAAAACAAT CAACACATAA TCAACCCAAC
....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
185 195 205 215 225 235
Short ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
Long ACACTATATC CCACATTCAC ACACACACAC ACACACAAAC ACACTCCTTC ATCAACATGT
Contig ACACTATATC CCACATTCAC ACACACACAC ACACACAAAC ACACTCCTTC ATCAACATGT
....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
245 255 265 275 285 295
Short ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
Long AGACAGAAAA TGGAACACGA CTACGCAAAT CAACATCGTC GTCCAACGAG AAATCTGTCC
Contig AGACAGAAAA TGGAACACGA CTACGCAAAT CAACATCGTC GTCCAACGAG AAATCTGTCC
....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
305 315 325 335 345 355
Short ~~~~~~~~~~ ~~~~~TCTCA TTACACCCAC ACATATGTTA ACAAACAACA AGTGGACTGG
Long AACCATCAAT GTCACTCTCA TTACACCCAC ATATATGTTA ACAAACAACA AGTGGACTGG
Contig AACCATCAAT GTCACTCTCA TTACACCCAC ACATATGTTA ACAAACAACA AGTGGACTGG
Short TT
Long TT
Contig TT

Figure 2

A


MW
I
II
III
IV
550
300
150
240
210
158
141
750
620
280
158
255
230
210
177
500
280
250
237
158
141

B

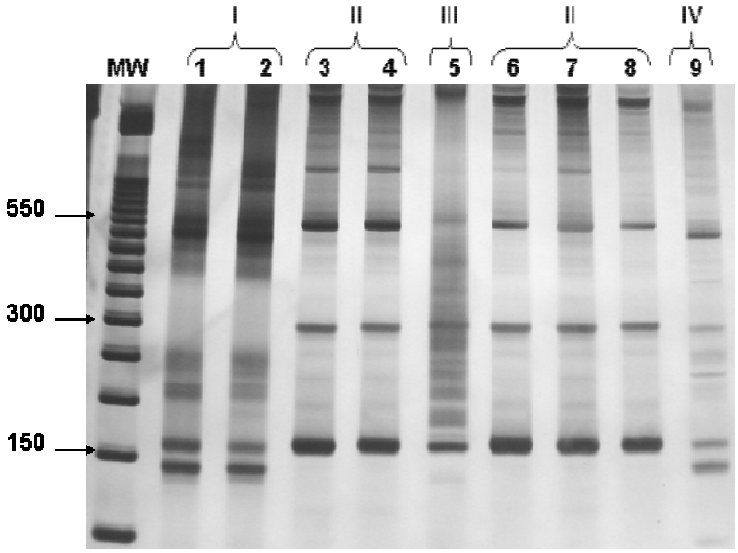
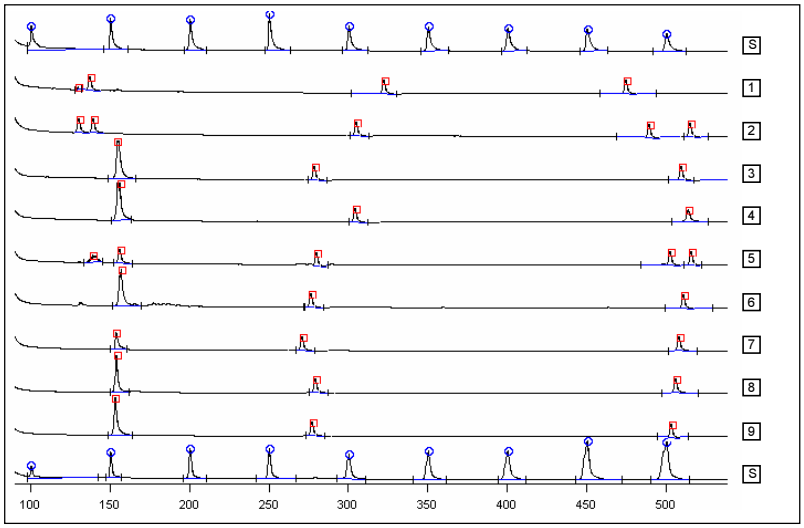

I
II
III
II
IV
MW
1
2
3
4
5
6
7
8
9
550
300
150

Figure 3



5.3 – Comparação dos diferentes métodos de extração de DNA.

Testamos quatro diferentes métodos de extração de DNA de ovos com o objetivo de avaliar qual seria o mais eficaz. Os DNAs extraídos destes ovos foram utilizados em rações de PCR com iniciadores específicos para *loci* de microssatélites.

De nove amostras de DNA extraídas com fenol/clorofórmio, obtivemos produtos de amplificação para apenas uma amostra (Figura 7A). De 10 amostras de DNA extraídas pelo método ROSE, três apresentaram produtos de amplificação (Figura 7B). Das nove amostras de DNA extraídas com o método "Sorensen", obtivemos produtos de amplificação para apenas uma (Figura 7C). Com o método que utiliza a resina "Instagene matrix", obtivemos produtos de amplificação para oito das 10 amostras de DNA extraídas (Figura 7D). É importante ressaltar que nestas reações utilizamos o mesmo para de iniciadores (SmBr09) que gera um fragmento de aproximadamente 161pb. Devemos ressaltara também que as duas amostras não amplificadas nesta reação (canaletas 1 e 9 da Figura 7D), apresentaram produtos em outras reações com estes e com outros iniciadores.

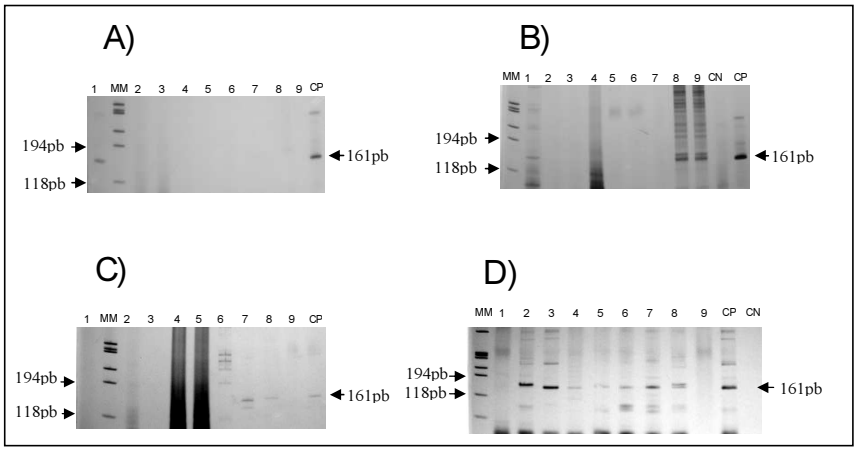


Figura 7 – Produtos de amplificação de DNA de ovos individuais de *S. mansoni* utilizando o par de iniciadores SmBr09. Preparações de DNA obtidas por diferentes métodos: A) Fenol/clorofórmio, B) "ROSE", C) "Sorensen" e D) Resina Instagene Matrix®. CP= controle positivo da reação, DNA de verme adulto. MM= marcador molecular. CN= controle negativo, reação sem adição de DNA.

5.4 – Artigo 3:

Genetic filtering and optimal sampling of *Schistosoma mansoni* populations

Publicado na revista Parasitology (Sorensen et al., 2006)

Sorensen RE, Rodrigues NB, Oliveira G, Romanha AJ, Minchella DJ. Genetic filtering and optimal sampling of *Schistosoma mansoni* populations. Parasitology. 2006. 133: Pt 4, 443-451

Neste trabalho estão sintetizados os resultados relativos ao objetivo específico 2:

2 Desenvolver metodologia para o uso de ovos de *S. mansoni* como fonte de material genético para genotipagem

Neste trabalho mostramos, pela primeira vez, a utilização de ovos individualizados de *S. mansoni* como fonte de DNA para PCR, bem como um novo protocolo para extração deste DNA. O protocolo apresentado é o que utiliza resina na extração e foi o que apresentou os melhores resultados dentre os 4 descritos na metodologia.

Mostramos também, pela primeira vez, a utilização de PCR-multiplex na genotipagem de miracídios e vermes adultos de *S. mansoni.* Tanto os miracídios quanto os vermes utilizados são provenientes das amostras de fezes coletadas de moradores de Virgem das Graças – Minas Gerais.

Genetic filtering and optimal sampling of Schistosoma mansoni populations

R. E. SORENSEN[1]*, N. B. RODRIGUES[2], G. OLIVEIRA[2], A. J. ROMANHA[2] and D. J. MINCHELLA[3]

[1] Department of Biological Sciences, Minnesota State University Mankato, Mankato, MN 56001, USA
[2] Centro de Pesquisas René Rachou – FIOCRUZ, Belo Horizonte, Minas Gerais, Brazil
[3] Department of Biological Sciences, Purdue University, West Lafayette, IN 47907, USA

(Received 28 March 2006; revised 29 April 2006; accepted 1 May 2006; first published online 4 July 2006)

SUMMARY

Allelic variation in 6 microsatellite markers was compared between frozen Schistosoma mansoni eggs and laboratory-passaged worms originating from the same 5 fecal samples obtained from Brazilian residents. Based on allelic richness values, the number of alleles detected per locus did not differ between egg and worm DNA templates. However, our ability to score loci differed between these DNA templates, with worms providing more scored loci per individual than eggs. Differences also existed between the worms and eggs in the identity of the specific alleles that were detected. Additionally, we observed a reduction in homozygous genotypes among laboratory-passaged worms relative to the eggs. Allelic diversity curves were calculated by genotyping all worms from a representative host sample to determine the relationship between the number of alleles detected at a locus and the number of worms genotyped. Curves for the 5 residents' worm infrapopulations for each of the loci were very similar. The equation $y = 19{\cdot}55r\ \ln(x) + 9{\cdot}992$ explained the association between sampling effort (x) and number of alleles detected (y) with an R^2 of 0·775. In conclusion, egg DNA templates and allelic diversity curves can benefit efforts to discern the sociological, ecological and evolutionary forces impacting the genetic diversity and disease epidemiology of human schistosomes.

Key words: Schistosoma mansoni, microsatellites, allelic diversity, miracidia, population genetics, host-induced selection, sampling approaches.

INTRODUCTION

Schistosomiasis remains a major public health concern affecting over 200 million people worldwide. As such, considerable financial, medical and scientific effort has been devoted to understanding the epidemiological factors influencing the transmission and spread of the parasite. Population-genetics theories and methodologies have enhanced our knowledge of Schistosoma mansoni population dynamics in light of the sociological, ecological and evolutionary forces that impact disease epidemiology. As studies investigating Schistosoma population genetics increase, our ability to discern the relative strength of these forces over space and time improves. However, accurate interpretation of natural genetic patterns of schistosomes requires both direct collection of genotypic data and assurance that the portion of sampled parasites constitutes an unbiased estimate of genetic diversity.

* Corresponding author: Department of Biological Sciences, Minnesota State University Mankato, Mankato, MN 56001, USA. Tel: + 507 389 1280. Fax: + 507 389 2788. E-mail: robert.sorensen@mnsu.edu

Schistosoma mansoni genetic studies involving both sylvatic and human transmission cycles have been performed to discern the forces shaping parasite transmission and maintenance in endemic areas (Minchella et al. 1995; Barral et al. 1996; Sire et al. 1999; Curtis et al. 2001, 2002; Sire et al. 2001a, b; Eppert et al. 2002; Prugnolle et al. 2002, 2004, 2005; Theron et al. 2004). While the intramolluscan parasite stages can be studied equitably in either cycle, differences exist in how parasite tissue is recovered from the vertebrate host. In sylvatic cycles of S. mansoni, adult worms are harvested directly from sacrificed murine hosts (Barral et al. 1993, 1996; Sire et al. 2001a, b; Prugnolle et al. 2002, 2004, 2005; Theron et al. 2004). Alternatively, among human hosts, S. mansoni eggs are collected from patient feces and subsequently passaged through laboratory populations of snails and rodents to obtain adult worms for genetic analysis (Minchella et al. 1994; Curtis et al. 2001, 2002; Rodrigues et al. 2002; Stohler et al. 2004). An important problem associated with this methodology is the potential for artificially selecting genotypes of the parasite that are better adapted for laboratory hosts, while filtering out genotypes adapted to natural host populations (LoVerde et al. 1985). The ability to genotype multiple loci from

Parasitology (2006), 133, 443–451. £ 2006 Cambridge University Press
doi:10.1017/S0031182006000552 Printed in the United Kingdom

individual miracidia collected from eggs in human stool samples would circumvent the need for laboratory passage of the schistosomes, allowing a more direct assessment of genotypes associated with human schistosomiasis.

Determining the appropriate number of worms to sample from infected hosts is also an important problem facing researchers investigating S. mansoni population genetics. Currently, there are 2 commonly employed methodologies for collecting genotypic information from vertebrate hosts: genotyping all worms within a single host (Sire et al. 2001a; Theron et al. 2004) or genotyping an arbitrarily-sized subsample of worms (Curtis et al. 2001, 2002; Sire et al. 2001a; Rodrigues et al. 2002). Both of these methods have shortcomings. Genotyping all worms can lead to sampling redundancy (and wasted effort), as hosts often carry redundant clonal genotypes due to transmission from snail hosts. However, choosing an arbitrary number of worms from each host presents the risk of under-sampling the available parasite genotypes. A more efficient method would involve genotyping all worms from a subset of hosts and determining the relationship between the number of unique alleles and the number of genotyped worms. This relationship could help to determine the proportion of an infrapopulation that should be genotyped in the remainder of the hosts.

This paper presents our approaches to optimize S. mansoni sampling efforts in endemic regions. Herein, we describe a protocol for extracting DNA from single S. mansoni eggs (from human feces) that allows subsequent multilocus, microsatellite genotyping. We also assess the importance of genetic filtering during laboratory passage of S. mansoni, by genotyping parasites prior to laboratory infections (from eggs) and after passage through both snails and rodent hosts (from adult worms). Lastly, we present a methodology to empirically optimize the trade-off between capturing sufficient parasite genetic variability in adult worms and reducing the sampling effort necessary to gather that information.

MATERIALS AND METHODS

Collection of allelic information

Schistosoma mansoni eggs (in feces) were collected from 5 residents (hereafter, individually referred to as PID142, PID259, PID404, PID447, and PID008) of the Virgem das Graças (VdG) study area in the Jequitinhonha Valley of northern Minas Gerais, Brazil. These residents were expected to comprise a good sample of S. mansoni genetic variability within the larger VdG study population based on their high parasite intensities and the dispersed nature of their residences across the study area. The larger study population consisted of 47 individuals residing in a 60 km^2 area surrounding Corrego do Cardoso and Corrego do Suçuarana. The village of VdG is near the centre of this study area. Participants of this study were divided into 5 geographical groups based on their proximity to the village. The 5 residents considered in this paper represent all 5 of the geographical groups. Evaluation of which residents to select as representatives of a geographical group was based on the number of S. mansoni adults that were collected following laboratory passage of miracidia through Biomphalaria glabrata snails and BALB/c mice. Within each geographical group, the resident providing the most worms from a mouse infection involving o 5 cercariae-releasing snails was chosen for this study. This selection criteria offered the best a priori opportunity of detecting maximal levels of genetic variability within each of the study areas, while minimizing the number of parasite infrapopulations assessed since the greatest loss of schistosome genetic diversity is assumed to occur when the snails and the mice are exposed to their respective infective stages (mean number of snails per infected mouse for all 47 VdG residents ¡ 1 standard deviation= 3·54¡ 1·75 based on 130 mice infected in the larger study).

Eggs were removed from feces using 2 washes and sedimentation in a 1·85% NaCl solution. These eggs were either (1) transferred individually to 0·5 ml microfuge tubes containing 50 ml of a 1·85% NaCl solution for storage at x 20 xC (14–17 eggs per resident) or (2) transferred to dechlorinated tap water under a 100 W light source to induce miracidial hatching. Five groups of 15 laboratory-reared Biomphalaria glabrata snails were individually exposed to 10 miracidia per snail to represent the S. mansoni infections of the 5 VdG residents. Unfortunately, logistical constraints limited our ability to freeze enough eggs to provide a more equitable number of eggs and miracidia. Cercariae from each group of snails were used to infect 2–5 BALB/c mice (100–120 cercariae/mouse). Mouse infections and the subsequent collection of adult worms followed procedures described by Curtis et al. (2001). For each of the 5 VdG residents, the mouse that offered the greatest potential recovery of schistosome diversity, based on the number of snails providing cercariae and the number of worms recovered, was selected to represent that individual's schistosome infrapopulation. These experimentally infected mice yielded 414 adult schistosomes (Table 1). DNA from these worms was extracted using mechanical pulverization, chemical degradation, and ethanol precipitation (Sorensen et al. 1998).

In order to extract DNA from the 74 individual S. mansoni eggs, the eggs were first isolated from the saline storage buffer by centrifuging each tube at 13400 g for 5 min and eluting the supernatant. Tubes containing the single egg were then submerged in liquid nitrogen for 30 sec to rupture the surface of the eggshell and then 100 ml of InstaGene Matrix (Bio-Rad) was added to the disrupted egg,

Table 1. Summary table showing allelic data for 6 microsatellite loci detected among adult Schistosoma mansoni worms (W) and miracidia within eggs (E). Worms were harvested from BALB/c mice following laboratory passage of miracidia obtained from fecal samples from 5 Virgem das Graças residents (PID142–PID008). Eggs were collected when miracidia for laboratory passage were harvested.

	PID142		PID259		PID404		PID447		PID008	
Parameter	W	E	W	E	W	E	W	E	W	E
Total Collected	54	15	86	14	90	17	74	14	110	14
Total Genotyped	48	14	78	13	84	15	69	14	106	14
DA03 Locus										
Number Genotyped	47	6	78	11	84	11	69	12	106	10
Number of Alleles	8	4	12	8	12	8	16	8	8	8
Frequency f 0·05	2	0	6	4	8	5	9	4	3	0
Frequency o 0·20	2	2	2	2	2	2	2	1	2	1
H_O[a]	1·000	0·333	0·885	0·545	0·857	0·364	0·957	0·583	1·000	0·000
H_E[b]	0·809	0·625	0·835	0·781	0·834	0·781	0·865	0·757	0·717	0·840
Allelic Richness	3·93	2·97	4·14	4·00	4·16	3·94	4·46	3·85	3·26	4·57
DO03 Locus										
Number Genotyped	43	11	78	11	83	9	69	9	106	4
Number of Alleles	7	9	10	8	11	8	10	7	5	6
Frequency f 0·05	3	3	4	3	5	0	4	0	0	0
Frequency o 0·20	2	1	2	1	1	1	2	1	3	2
H_O[a]	0·744	0·545	0·974	0·545	0·590	0·444	0·957	0·556	0·991	0·500
H_E[b]	0·777	0·855	0·789	0·831	0·773	0·815	0·789	0·796	0·738	0·813
Allelic Richness	3·69	4·64	3·81	4·36	3·75	4·13	3·83	4·39	3·32	4·93
DO23 Locus										
Number Genotyped	47	7	71	12	84	10	64	11	103	5
Number of Alleles	9	5	10	9	8	7	11	6	8	5
Frequency f 0·05	2	0	5	1	2	0	5	1	5	0
Frequency o 0·20	2	1	1	1	2	3	1	1	3	5
H_O[a]	0·872	0·143	0·465	0·500	0·690	0·300	0·797	0·273	0·903	0·000
H_E[b]	0·835	0·622	0·731	0·847	0·824	0·835	0·858	0·806	0·714	0·800
Allelic Richness	4·13	3·24	3·58	4·58	4·03	4·40	4·38	4·06	3·17	4·33
ED28 Locus										
Number Genotyped	28	9	78	11	83	12	66	10	106	14
Number of Alleles	2	2	3	2	3	2	5	2	2	2
Frequency f 0·05	0	0	1	0	1	0	2	0	0	0
Frequency o 0·20	2	2	2	2	2	2	2	2	2	2
H_O[a]	0·143	0·111	0·128	0·455	0·277	0·083	0·424	0·100	0·321	0·000
H_E[b]	0·436	0·500	0·165	0·500	0·386	0·413	0·395	0·495	0·429	0·490
Allelic Richness	1·91	1·99	1·46	1·99	1·91	1·93	2·02	1·99	1·90	1·98
ED57 Locus										
Number Genotyped	46	9	78	12	82	12	67	10	105	3
Number of Alleles	6	10	7	9	14	9	12	9	6	3
Frequency f 0·05	2	0	2	4	7	4	5	4	2	0
Frequency o 0·20	2	0	3	1	2	1	1	0	2	3
H_O[a]	0·522	0·444	0·667	0·583	0·854	0·583	0·716	0·300	0·533	0·000
H_E[b]	0·704	0·883	0·735	0·771	0·820	0·771	0·860	0·845	0·707	0·667
Allelic Richness	3·19	5·07	3·36	4·03	4·10	4·80	4·38	4·59	3·16	3·00
G5 Locus										
Number Genotyped	47	8	71	10	83	10	63	8	104	6
Number of Alleles	3	3	6	3	3	4	4	3	4	3
Frequency f 0·05	0	0	3	1	0	0	1	0	1	0
Frequency o 0·20	2	2	2	2	2	2	2	2	3	2
H_O[a]	0·681	0·125	0·620	0·300	0·602	0·100	0·635	0·125	0·633	0·167
H_E[b]	0·567	0·617	0·560	0·485	0·559	0·675	0·566	0·570	0·661	0·569
Allelic Richness	2·42	2·73	2·39	2·22	2·35	3·22	2·38	2·48	2·75	2·68

[a] Observed heterozygosity value.
[b] Expected heterozygosity value.

followed by a 3-min. incubation at 56 xC. This was followed by 30 sec of high-speed vortexing, a second incubation for 8 min at 100 xC, a final 10-sec high-speed vortexing, and 3 min of centrifugation. An aliquot of this supernatant was used as template DNA in 10 ml vol. PCRs.

Six schistosome microsatellite loci (ED28, ED57 (Durand et al. 2000), DA03, DO03, DO23, and G5 (Rodrigues et al. manuscript in preparation)) were amplified in 15 ml vol.multiplexed PCR reactions for the worms and in 10 ml vol.single-locus PCR reactions for the eggs; multiplexed reactions using egg template failed to produce detectable alleles. Whether for eggs or worms, each amplification reaction contained 10 mM Tris-HCl, pH 9·0, 50 mM KCl, 0·1% Triton[1] X-100, 200 mM dNTP, 2 mM $MgCl_2$, 167 nM of each primer (3–4 primer pairs/reaction for multiplex reactions), 0·6 units of Taq DNA Polymerase (Promega), 0·8 mg BSA and 50 ng of worm DNA or a 4 ml aliquot of the InstaGene Matrix supernatant. Thermal cycling was performed under the following conditions: 4 min at 94 xC; 32 cycles of 30 s at 94 xC, 30 s at 55 xC, and 1 min at 65 xC; and a final 3 min extension at 65 xC, in either a MJ Research PTC-100 or an Eppendorf Mastercycler gradient thermocycler. The fluorescently labelled PCR products were electrophoresed on an ABI 377 sequencer using GENESCAN[1]-500[TAMRA] as an internal size standard for each lane. Allele sizes were determined using GENESCAN v3.1 and GENOTYPER v2.5 software (PE Applied Biosystems).

Assessing equality of egg and worm DNA sources

Allelic diversity (number of alleles and allele frequency), heterozygosity measures (observed and expected values), and Hardy-Weinberg equilibrium statistics of egg and worm DNA templates were determined using GenAlEx 6 (Peakall and Smouse, 2005). Since exhaustive collection of the parasites from experimentally infected mice yielded over 5 times more worms than the number of eggs that were initially reserved for genetic analyses, comparison of allelic diversity from the two sources was biased in favour of worms. This bias was addressed by calculating the allelic richness of each locus (El Mousadik and Petit, 1996; Petit et al. 1998) for the egg and worm templates using FSTAT v.2.9.3 (Goudet, 2001). All statistical tests were performed using SPSS 11 for Mac OS X. One-way ANOVA was used to compare the number of alleles detected per locus and overall allelic richness values between eggs and worms, while Mann-Whitney U or Kruskal-Wallis tests were used to make statistical comparisons when assumptions of normality or homoscedasticity were not met. Frequency differences were tested using chi-square analysis. An a of 0·05 was used to evaluate statistical significance in all cases.

Construction of an allelic diversity curve

Genotypic data from the most diverse microsatellite locus for each of the 5 VdG worm infrapopulations was used to determine how the cumulative number of unique alleles at these loci increased as a function of the number of worms assayed. By considering the most diverse locus in worms obtained from each of the VdG residents, we necessarily encompass the diversity of other less-diverse loci within their worm infrapopulations. Because the relationship between the number of unique alleles detected and the number of worms genotyped depends upon the order in which individual worms in a population are considered, an algorithm was written in Python 2.3.3 to randomly reorder and resample the original data from each of the 5 VdG worm infrapopulations. This program allowed us to consider the relationship between number of alleles detected and number of worms genotyped by iteratively shuffling the order in which individuals were genotyped (program available upon request from RES via email). Python's random.shuffle function was used to randomly change the order in which alleles from these worms were considered as unique or redundant.

Regression analysis was used to determine the most biologically relevant, mathematical relationship (based on R^2 values) between the percentage of unique alleles detected and the percentage of the infrapopulation assayed using re-sampled data and SPSS 11 for Mac OS X. The optimal number of times to shuffle allelic data from each of the 5 VdG resident infrapopulations was determined iteratively using the single most diverse locus among the 5 worm infrapopulations. In this case, genotypic data from locus DA03 in PID447 was iteratively shuffled and reassembled 5, 10, 15, 20, 30, or 60 times. The best-fit curve was calculated for each of the resulting 6 datasets. The coefficients of determination (R^2) from these curves were retained, and the process was repeated 4 additional times to determine the mean R^2 and its standard deviation based on 5 trials for each of the 6 reassembly alternatives. This allowed to us to find the number of re-sampling iterations that minimized variation in the R^2 value. Outputs generated using the optimal number of re-sampling iterations were combined into a single dataset to determine an overall best-fit regression curve.

RESULTS

Equality of egg and worm DNA sources

The 5 groups of experimentally infected BALB/c mice considered in this study yielded 414 adult schistosomes (Table 1) representing offspring of the infections in 5 Virgem das Graças residents. Of these, 385 (93·0%) produced allelic information for at least 1 of the 6 microsatellite loci under consideration (Table 1). PID142 yielded the fewest scored worms (88·9%), and PID008 offered the most (96·4%). Seventy (94·6%) of the collected eggs provided allelic information for at least 1 of the loci being studied,

and the number of genotyped eggs appeared evenly distributed across the 5 residents (Table 1).

Our ability to detect microsatellite alleles differed somewhat among egg and worm DNA templates. This difference was observed in the mean number of loci scored per individual and in the proportion of genotypes detected per locus. Worms provided significantly more scored loci per individual than did eggs (mean= 5·8 vs 4; Z= x 9·985, U= 6164, P< 0·001). In the same manner, a significantly smaller proportion of eggs provided genotypic information at each locus than did worms, with 60·0–80·0% of the eggs and 93·8–99·7% of the worms offering genotypic information across the 6 loci (mean= 67·2% vs 96·9%; x^2= 445·42, D.F.= 1, P< 0·001). However, at a given locus, the number of alleles that were detected did not differ between the egg and worm populations assayed ($F_{1, 10}$= 0·586, P= 0·462).

Allelic variability was similar for both DNA sources in terms of the number of alleles per locus. Four of the loci, DA03, ED57, DO03, and DO23, all possessed > 10 alleles, while G5 and ED28 presented < 10 alleles regardless of the DNA source (Table 1). The frequency of more common alleles (frequency o 0·20) was also similar in both the worm and egg portion of the population (Z= x 0·589, U= 14·500, P= 0·589). However, rare alleles (frequency f 0·05) were detected more frequently in the worm population compared to the egg population (Z= x 2·166, U= 4·500, P= 0·026). The occurrence of specific alleles differed between worms and eggs as well. Twenty alleles were found among the worms that were absent from the eggs, while 5 alleles were detected among the eggs that were lacking in the sampled worms (data not shown).

Allelic richness per locus did not differ between the 5 VdG residents when the egg samples and the worm samples were considered separately (worm: $F_{4, 25}$= 0·434, P= 0·783; egg: $F_{4, 25}$= 0·056, P= 0·994; Table 1). Furthermore, when the egg and worm samples were combined, allelic richness per locus did not differ between the residents or between the sources (egg or worm) of the DNA (resident: $F_{4, 55}$= 0·260, P= 0·903; source: $F_{1, 58}$= 1·714, P= 0·196). Although, allelic richness values were similar in eggs and worms, whether considered on a per locus or a per resident basis, the number of microsatellite alleles present differed among the loci, both when the egg and worm samples were considered separately and when they were grouped (egg: H= 21·856, D.F.= 5, P= 0·001; worm: H= 21·388, D.F.= 5, P= 0·001; grouped: H= 40·295, D.F.= 5, P< 0·001).

Measures of observed and expected heterozygosity differed between the worm and egg portions of the S. mansoni population (Table 1). Observed levels of heterozygosity (H_o) were significantly larger among worms compared to eggs across all loci (U= 4·000, Z= x 2·246, P= 0·026), while expected heterozygosity (H_e) values were comparable for the eggs and the worms (U= 13·5, Z= x 0·727, P= 0·485). When eggs were used as the DNA source, H_o per locus and over all loci in each patient were lower than H_e. Nevertheless, some of the loci (PID142: loci DA03 and DO03; PID259: loci DO03, ED28, and G5; PID404: locus DO03; PID447: loci DA03 and DO03; PID008 loci DO03 ED57, and G5; P> 0·05) did not deviate significantly from Hardy-Weinberg equilibrium (HWE), which may be due to the paucity of eggs genotyped for these loci among the individual residents. Unexpectedly, however, when adult worms were used as the DNA source, 16 loci showed H_o higher than the H_e, while 14 loci showed the opposite outcomes (Table 1). All but 1 (PID259: G5; P> 0·05) loci in the worm populations deviated significantly from HWE (P< 0·05).

Estimation of an allelic diversity curve

Using reshuffled data from the most diverse locus in each of the 5 VdG resident's worms, a logarithmic relationship, of the form y= mr ln(x)+ b, best fit the association between the percentage of the unique alleles detected (x) and the percentage of the worm population sampled (y) better than alternative relationships (data not shown). The coefficient of determination's (R^2) mean standard deviation decreased dramatically as the number of shuffling iterations increased from 5 to 15 and remained relatively constant as iterations increased beyond 15. As such, 30 re-sampling iterations of the most diverse microsatellite locus for each worm infrapopulation was used to determine the logarithmic relationship between the number of worms analysed and the number of unique alleles detected. The allelic diversity curves for each of the 5 VdG resident's worm populations were very similar (Fig. 1). Furthermore, for all resident-locus combinations, other than PID008, R^2 for the regression exceeded 0·780 (range= 0·544–0·916; mean= 0·806) indicating a reasonable fit of the calculated relationship to the data. When the data used to produce these 5 curves were combined to determine an overall relationship, we found that the equation y= 19·55r ln(x)+ 9·992 explained the association between worm number and unique allele number with an R^2 of 0·775.

DISCUSSION

Our results show that sufficient high-quality DNA for microsatellite analyses can be obtained from individual Schistosoma mansoni eggs (collected from human feces) when an extraction protocol containing InstaGene Matrix is used. This technique greatly reduces the time and expense necessary to compile schistosome allelic data from human hosts by making laboratory passage of the parasites prior to genetic analysis unnecessary. PCR amplification of schistosome DNA has been utilized for microsatellite loci

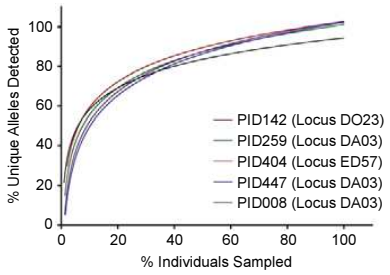


Fig. 1. Trendlines representing logarithmic relationships for the percentage of unique alleles detected for each of 5 microsatellite loci (DA03–G5) as a function of the percentage of worms in the infrapopulation (PID142–PID008) that were sampled. Each trendline is based on re-sampling allelic data from the single most diverse locus detected in a single patient infrapopulation of Schistosoma mansoni 30 times.

from pooled miracidia (Silva et al. 2006) and multicopy or mitochondrial loci from eggs within fecal samples (Pontes et al. 2002; Gobert et al. 2005; Shrivastava et al. 2005), but to our knowledge this is the first demonstration of multilocus microsatellite genotyping of single copy loci from individual miracidia within schistosome eggs. Collection of multilocus genotypic information from individual eggs, as shown herein, enables the most comprehensive testing of hypotheses about the evolution of human schistosome populations. Although, allele frequency data can be obtained by genotyping a single locus from individual miracidia (Shrivastava et al. 2005), single locus genotypes are limited in their ability to identify alleles that are identical by descent compared to multilocus genotypes. Likewise, estimates of allele frequency that are calculated from pools of schistosome eggs (Silva et al. 2006) cannot provide measures of heterozygosity, limiting their usefulness in explaining modes of evolution within and among populations.

The use of individual S. mansoni eggs, rather than worms, as a DNA source may warrant reservations when successful detection of all microsatellite alleles under investigation is essential or when multiplexed PCR reactions are preferred. The success rate of microsatellite PCR reactions was greater using a worm DNA template in terms of the number of loci detected per individual and the proportion of genotypes detected per locus. Furthermore, the fact that worm DNA allowed multiplexed PCR reactions and egg DNA did not, may influence the economics of some studies. These negative aspects of using schistosome eggs may be countered by our finding that a comparable number of alleles was detected per locus whether egg or worm DNA was used and that the same common alleles were detected with both template sources. Similarly, further experimentation with single locus and multiplex PCR conditions should increase the efficiency of DNA amplification from egg templates. A somewhat similar DNA extraction protocol for nematode eggs that does not require freezing the eggs in liquid nitrogen has been described by Floyd et al. (2002), and may be worth considering in future studies.

In this study, we detected rare alleles more frequently among sampled worms than among genotyped eggs. However, this finding is confounded by the fact that 5 times more worms were genotyped than eggs in this study. It is worth reiterating that the lack of eggs was not related to the availability of eggs for collection, rather it was limited by our ability to store and freeze more eggs when the fecal samples were being processed. Because this method of collecting schistosome DNA from an individual egg was unproven at the time the eggs were harvested, greater emphasis was placed on the proven technique for collecting schistosome genetic information, which meant that infecting snails was the primary focus. Although more rare alleles were detected among the worms, 5 alleles were detected among the eggs that were not observed among the worms, which is noteworthy given that many times fewer eggs were analysed. This finding demonstrates that S. mansoni allelic diversity was lost from the original egg population as evidenced by their absence following passage through laboratory hosts. This supports previous evidence (Stohler et al. 2004) that genetic diversity is lost while maintaining laboratory populations of this parasite. However, this does not require that these alleles were lost due to selection pressures imposed by the laboratory hosts since rare alleles are strongly affected by chance events.

To more equitably compare the genetic diversity of our egg populations to the larger worm populations, we utilized allelic richness values to compare the number of alleles detected among the egg and worm samples. Allelic richness per resident did not differ among the 5 VdG residents suggesting that S. mansoni expresses similar intra-host genetic diversity in our study site. In other words, residents sample similar levels of the S. mansoni genetic diversity (richness) within the study area. Differences were detected in allelic richness between loci when eggs and worms were analysed individually. However, no such differences were found between the egg and worm subpopulations when data from both subpopulations were pooled. These results show that harvesting S. mansoni eggs from stool samples does not limit our ability to detect microsatellite alleles.

This study also considered the genetic consequences of laboratory passage of natural S. mansoni isolates since the worms we analysed were derived from the same stool samples as the genotyped eggs. We anticipated no difference in the expected and observed heterozygosity values for the two DNA

sources because the genotyped eggs and the miracidia, which produced the worms, were obtained through replicate sampling of a common set of eggs. Our results do not support this hypothesis. When eggs were used as a DNA source, deficits in heterozygosity, relative to HWE, were detected in all loci across the 5 VdG residents. This finding does not seem to be related to technical limitations associated with our DNA extraction technique since heterozygotes were detected among 30·5% of the loci across all individual eggs. Furthermore, 74·7% of the sampled eggs were heterozygous at one or more loci and 63·1% of the eggs that yielded wholly homozygous multilocus genotypes came from PID008 (data not shown). Worms from PID008 also showed high levels of homozygosity. Therefore, the observed lack of heterozygous genotypes among eggs is likely due to natural causes including: inbreeding of the worms within a resident, population substructuring (Wahlund effect) due to residents sampling cercariae from several different S. mansoni subpopulations, or from positive selection pressures. Interestingly, when passaged adult worms were used as the DNA source, an excess of heterozygotes relative to HWE was observed in 3–4 of the loci among all the studied VdG residents. Selection pressures that favour heterozygous individuals would favour such an outcome. Comparison of H_o and H_e values in eggs and worms also suggests that heterozygote advantage may have affected the genetic diversity of the worm subsample since H_e values were equivalent for the two DNA sources but H_o for worms was significantly larger than that for eggs. Since the worms and eggs used in this study were derived from a common stool sample, this result suggests that passage through laboratory strains of snails and mice favoured heterozygous S. mansoni genotypes. This is an intriguing result as long-term passage of schistosomes through laboratory hosts results in reduced heterozygosity (Stohler et al. 2004), whereas heterozygote advantage should promote the maintenance of allelic diversity. In light of our current results, it appears that the reduction in diversity of laboratory populations, due to founder effects and potential bottlenecks, is sufficient to counter any initial heterozygote advantage that may result from passage through laboratory hosts.

To the extent that the 5 individuals from VdG used as the source of worms for this study represent all VdG residents, $y = 19{\cdot}55r\ \ln(x) + 9{\cdot}992$, allows us to determine the proportion of worms required to obtain high allelic diversity among parasites from other VdG residents. Based on our overall equation, we can predict that when 60% of the S. mansoni tissue (laboratory passaged worms, most specifically) from other VdG resident infrapopulations is genotyped 90% of the distinct alleles present among that worm population should be detected. This is a powerful approach because it relies on allelic data from a subset of the original study population to define sampling effort for the remainder of the study population rather than arbitrarily choosing a number of individuals to genotype. It is likely that population genetic patterns discerned using this method are more indicative of the actual population structure than patterns suggested by less stringent sampling methods. Since this approach is based on previously obtained evidence of genotypic redundancy for a defined study area, its use can substantially reduce the overall cost of genetic analyses for parasite populations where clonality is an important determinant of population structure. Use of a cost-benefit curve can also limit over-sampling, which would be accompanied by exclusion of redundant genotypes.

Selecting appropriate hosts for generating parasite allelic diversity curves is critical for the applicability of this methodology. Variables that we consider most important for choosing host candidates include: the number of infection foci in the study area, the spatial distribution of these foci, aggregation of hosts relative to the foci, and the host's parasite intensity. As a general rule, the number of sampled hosts should be increased relative to the number of infection foci since a single host is unlikely to adequately represent the genetic diversity of multiple foci unless the foci are close together. As a result, aggregated host populations and/or foci may require fewer sampled hosts relative to populations and/or foci that are more diffuse. Lastly, hosts with similarly high parasite intensities (based upon initial intensity screening) should be given priority, as these individuals likely posses the richest parasite diversity. Each of these variables (and their interactions) should be considered within a particular study area to avoid biasing estimates of parasite genetic diversity.

In summary, the schistosome sampling approaches presented here provide worthwhile methods to increase the efficiency of our efforts to discern the sociological, ecological and evolutionary forces that impact genetic diversity and disease epidemiology. The use of template DNA from individual schistosome eggs can dramatically reduce the time and costs associated with obtaining multilocus genotypes by negating the need to artificially passage life-cycle stages in the laboratory. Likewise, sampling schemes that are defined by the underlying allelic diversity provide a realistic view of the genetic diversity in that population while limiting both the time and resources necessary to collect that view.

We are indebted to Laura Warlow, Kathy Mounts, Esther Dubrovsky, Marcilene Rezende Silva and Maira Mazzoni Pucci for their technical support. We also wish to thank Liana Konovaloff Jannotti Passos for help with purifying the eggs and infecting snails as well as recognizing Andrea Gazzinelli and Rodrigo Correa-Oliveira for assistance with the field portion of this work. Manuscript comments, by Monika Zavodna, Greg Sandland and Jason Curtis were

especially valuable. This work was funded by a National Institutes of Health (USA) Research Grant (R01-AI-42768) to D. J. M. and in part by a Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG) Research Grant (407/02) and a World Health Organization/Tropical Disease Research (WHO/TDR) Grant (A20399) to G. O.

REFERENCES

Barral, V., This, P., Imbert-Establet, D., Combes, C. and Delseny, M. (1993). Genetic variability and evolution of the Schistosoma genome analysed by using random amplified polymorphic DNA markers. Molecular and Biochemical Parasitology 59, 211–222.

Barral, V., Morand, S., Pointier, J. P. and Theron, A. (1996). Distribution of schistosome genetic diversity within naturally infected Rattus rattus detected by RAPD markers. Parasitology 113, 511–517.

Curtis, J., Fraga, L. A., de Souza, C. P., Correa-Oliveira, R. and Minchella, D. J. (2001). Widespread heteroplasmy in schistosomes makes an mtVNTR marker ''nearsighted''. Journal of Heredity 92, 248–253.

Curtis, J., Sorensen, R. E. and Minchella, D. J. (2002). Schistosome genetic diversity: the implications of population structure as detected with microsatellite markers. Parasitology 125 (Suppl.), S51–S59.

Durand, P., Sire, C. and Theron, A. (2000). Isolation of microsatellite markers in the digenetic trematode Schistosoma mansoni from Guadeloupe Island. Molecular Ecology 9, 997–998.

El Mousadik, A. and Petit, R. J. (1996). High level of genetic differentiation for allelic richness among populations of the argan tree [Argania spinosa (L.) Skeels] endemic to Morocco. Theoretical and Applied Genetics 92, 832–839.

Eppert, A., Lewis, F. A., Grzywacz, C., Coura Filho, P., Caldas, I. and Minchella, D. J. (2002). Distribution of schistosome infections in molluscan hosts at different levels of parasite prevalence. Journal of Parasitology 88, 232–236.

Floyd, R., Abebe, E., Papert, A. and Blaxter, M. (2002). Molecular barcodes for soil nematode identification. Molecular Ecology 11, 839–850.

Gobert, G. N., Chai, M., Duke, M. and McManus, D. P. (2005). Copro-PCR based detection of Schistosoma eggs using mitochondrial DNA markers. Molecular and Cellular Probes 19, 250–254.

Goudet, J. (2001). FSTAT, a program to estimate and test gene diversities and fixation indices (version 2.9.3). Available from http://www.unil.ch/izea/softwares/fstat.html. Updated from Goudet (1995).

LoVerde, P. T., DeWald, J., Minchella, D. J., Bosshardt, S. C. and Damian, R. T. (1985). Evidence for host-induced selection in Schistosoma mansoni. Journal of Parasitology 71, 732–734.

Minchella, D. J., Lewis, F. A., Sollenberger, K. M. and Williams, J. A. (1994). Genetic diversity of Schistosoma mansoni: quantifying strain heterogeneity using a polymorphic DNA element. Molecular and Biochemical Parasitology 68, 307–313.

Minchella, D. J., Sollenberger, K. M. and Pereira de Souza, C. (1995). Distribution of schistosome genetic diversity within molluscan intermediate hosts. Parasitology 111, 217–220.

Peakall, R. and Smouse, P. E. (2005). GenAlEx 6: Genetic Analysis in Excel. http://www.anu.edu.au/BoZo/GenAlEx/. Population genetic software for teaching and research. The Australian National University.

Petit, R. J., El Mousadik, A. and Pons, O. (1998). Identifying populations for conservation on the basis of genetic markers. Conservation Biology 12, 844–855.

Pontes, L. A., Dias Neto, E. and Rabello, A. (2002). Detection by polymerase chain reaction of Schistosoma mansoni DNA in human serum and feces. American Journal of Tropical Medicine and Hygiene 66, 157–162.

Prugnolle, F., De Meeus, T., Durand, P., Sire, C. and Theron, A. (2002). Sex-specific genetic structure in Schistosoma mansoni: evolutionary and epidemiological implications. Molecular Ecology 11, 1231–1238.

Prugnolle, F., Théron, A., Durand, P. and De Meeus, T. (2004). Test of pangamy by genetic analysis of Schistosoma mansoni pairs with its natural murine host in Guadeloupe. Journal of Parasitology 90, 507–509.

Prugnolle, F., Théron, A., Pointier, J.-P., Jabbour-Zahab, R. and Jarne, P. (2005). Dispersal in a parasitic worm and its two hosts: consequence for local adaptation. Evolution 59, 296–303.

Rodrigues, N. B., Loverde, P. T., Romanha, A. J. and Oliveira, G. (2002). Characterization of new Schistosoma mansoni microsatellite loci in sequences obtained from public DNA databases and microsatellite enriched genomic libraries. Mem Inst Oswaldo Cruz 97 (Suppl. 1), 71–75.

Shrivastava, J., Gower, C. M., Balolong, E. J., Wang, T. P., Quian, B. Z. and Webster, J. P. (2005). Population genetics of multi-host parasites—the case for molecular epidemiological studies of Schistosoma japonicum using larval stages from naturally infected hosts. Parasitology 131, 1–10.

Silva, L. K., Liu, S. and Blanton, R. E. (2006). Microsatellite analysis of pooled Schistosoma mansoni DNA: an approach for studies of parasite populations. Parasitology 132, 331–338.

Sire, C., Durand, P., Pointier, J. P. and Theron, A. (1999). Genetic diversity and recruitment pattern of Schistosoma mansoni in a Biomphalaria glabrata snail population: a field study using random-amplified polymorphic DNA markers. Journal of Parasitology 85, 436–441.

Sire, C., Durand, P., Pointier, J. P. and Theron, A. (2001a). Genetic diversity of Schistosoma mansoni within and among individual hosts (Rattus rattus): infrapopulation differentiation at microspatial scale. International Journal for Parasitology 31, 1609–1616.

Sire, C., Langand, J., Barral, V. and Theron, A. (2001b). Parasite (Schistosoma mansoni) and host (Biomphalaria glabrata) genetic diversity: population structure in a fragmented landscape. Parasitology 122, 545–554.

Sorensen, R. E., Curtis, J. and Minchella, D. J. (1998). Intraspecific variation in the rDNA ITS loci of 37-collar-spined echinostomes from North America: implications for sequence-based diagnoses and phylogenetics. Journal of Parasitology 84, 992–997.

Stohler, R. A., Curtis, J. and Minchella, D. J. (2004). A comparison of microsatellite polymorphism and heterozygosity among field and laboratory populations of Schistosoma mansoni. International Journal of Parasitology 34, 595–601.

Theron, A., Sire, C., Rognon, A., Prugnolle, F. and Durand, P. (2004). Molecular ecology of Schistosoma mansoni transmission inferred from the genetic composition of larval and adult infrapopulations within intermediate and definitive hosts. Parasitology 129, 571–585.

5.5 – Determinação da variabilidade genética e estudo da influência da quimioterapia sobre esta variabilidade.

Para a determinação da variabilidade genéticas dos parasitos (objetivo 3), foram analisadas 53 infrapopulações de parasitos de 33 de amostras de fezes da primeira coleta e de 20 da segunda coleta (Tabela 2). Para o estudo da influência da quimioterapia sobre esta variabilidade (objetivo 4), parasitos coletados em amostras de fezes após o tratamento e em maior detalhe, de 9 moradores, amostrados nas duas coletas, foram analisados; Card1-150/353, Card1-154/365, Card2-165/408, Card3-204/523, Card3-204/525, Card3-204/526, Suss-125/282, Suss-137/315 e VdG-46/543.

Foram selecionados 5 pares de iniciadores para esta fase do trabalho por apresentarem melhores resultados com DNA de ovos e em PCR-multiplex (Tabela 3).

A determinação do tamanho amostral indicou que um 'N'=21 ovos seria o número ideal a ser analisado de cada amostra de fezes. Entretanto, este número não foi possível em todas as amostras, desta forma o tamanho de nossa amostra variou de 10 a 22 ovos.

Em uma análise global, com as 53 infrapopulações observamos que o número de alelos por *locus* variou de 2 a 13, não houve correlação entre os alelos e o tratamento nas 9 infrapopulações estudadas (Tabela 4, sombreado). A heterozigozidade observada variou de 0,00 a 1,00. Com exceção de 13 infrapopulações (Tabela 4 índice b), em todas as outras e para todos os *loci* a heterozigosidade observada (Ho) foi sempre menor que a esperada (He). Em 39 das 53 populações, em pelo menos um *locus*, Ho foi não significativamente diferente de He ($p<0,05$) representando desvios para equilíbrio de Hardy-Weinberg. (Tabela 4*).

Tabela 4: Número de alelos por *locus,* heterozigosidade observada (Ho) e esperada (He), encontradas nas populações de parasitos de cada paciente estudado. As linhas sombreadas indicam indivíduos que foram amostrados antes e após o tratamento

Locus	SmBr19			SmBr16			SmBr17			SmBr9			SmBr15		
Moradores (N)	Alelos	Ho	He	Alelos	Ho	He	Alelos	Ho	He	Alelos	Ho	He	Alelos	Ho	He
Card1-147/338 (14)	6	0.29	0.73	8	0.43	0.83	7	0.14	0.71	4	0.43	0.60*	6	0.64	0.81*
Card1-149/348a (12)	8	0.25	0.92	8	0.42	0.88	9	0.50	0.85	5	0.83b	0.67*	6	0.67	0.84*
Card1-150/352a (12)	7	0.42	0.87	6	0.17	0.85	11	0.75	0.90*	4	0.08	0.78	5	0.50	0.80*
Card1-150/353 (13)	6	0.08	0.67	7	0.23	0.88	13	0.85	0.95*	4	0.38	0.80*	5	0.63	0.84*
Card1-150/353a (14)	8	0.64	0.79	8	0.14	0.81	8	0.29	0.90	5	0.64	0.69	5	0.50	0.81*
Card1-151/359b (18)	4	0.06	0.54	9	0.17	0.90	8	0.11	0.75	4	0.44	0.75	6	0.89b	0.78*
Card1-154/365 (10)	3	0.60b	0.55*	3	0.60b	0.57*	9	0.30	0.92	5	0.30	0.85	5	0.50	0.79*
Card1-154/365a (13)	6	0.23	0.82	7	0.23	0.88	7	0.15	0.88	4	0.77	0.75*	6	0.77	0.83*
Card2-164/399 (16)	6	0.06	0.84	8	0.00	0.88	3	0.25	0.61	5	0.44	0.68	6	0.57	0.80
Card2-165/403 (11)	6	0.18	0.88	7	0.00	0.95	6	0.09	0.66	3	0.18	0.26*	6	0.64	0.85*
Card2-165/407 (21)	5	0.00	0.81	8	0.57	0.84	10	0.71	0.78	6	0.19	0.84	6	0.48	0.82
Card2-165/408b (11)	7	0.09	0.89	10	0.27	0.94	6	0.36	0.83	4	0.64	0.68*	5	0.82b	0.67*
Card2-165/408a (10)	8	0.50	0.92	5	0.20	0.86	10	0.60	0.93	3	0.40	0.44*	5	0.60	0.80
Card2-165/409 (13)	5	0.00	0.73	7	0.15	0.86	11	0.31	0.93	4	0.08	0.77	7	0.54	0.87
Card2-167/415 (18)	6	0.06	0.78	6	0.06	0.59	8	0.28	0.83	4	0.56	0.75	6	0.50	0.82
Card2-169/542ab (12)	10	0.83	0.89	6	0.17	0.86	10	0.67	0.89	5	0.92b	0.82*	6	0.67	0.84
Card2-175/384a (18)	10	0.44	0.90	11	0.50	0.91	12	0.61	0.87	4	0.72	0.77	6	0.84	0.77*
Card2-182/464 (11)	5	0.09	0.84	7	0.00	0.93	3	0.18	0.61	2	0.09	0.09*	5	0.82	0.77*
Card3-198/495 (12)	7	0.00	0.86	8	0.17	0.93	5	0.33	0.70	4	0.42	0.70*	5	0.50	0.78
Card3-199/499 (12)	5	0.17	0.84	7	0.00	0.90	6	0.42	0.63	5	0.17	0.68	6	0.75	0.83*
Card3-199/500 (12)	7	0.42	0.84	5	0.00	0.85	7	0.17	0.80	4	0.50	0.71*	5	0.50	0.84

Tabela 4: Continuação.

Locus	SmBr19			SmBr16			SmBr17			SmBr9			SmBr15		
Moradores (N)	Alelos	Ho	He	Alelos	Ho	He	Alelos	Ho	He	Alelos	Ho	He	Alelos	Ho	He
Card3-203/516 (10)	3	0.00	0.66	7	0.20	0.92	10	0.60	0.95	3	0.40	0.44*	5	0.40	0.76
Card3-204/523 (15)	9	0.07	0.92	13	0.00	0.99	9	0.33	0.90	5	0.54	0.74*	6	0.80	0.82*
Card3-204/523a (14)	7	0.00	0.87	6	0.07	0.84	5	0.07	0.54	5	0.43	0.78	6	0.50	0.83
Card3-204/524 (10)	6	0.00	0.89	9	0.00	0.98	5	0.30	0.84	5	0.60	0.71*	5	0.70	0.81*
Card3-204/525 (20)	11	0.10	0.85	11	0.00	0.94	16	0.45	0.92	5	0.20	0.77	6	0.55	0.77
Card3-204/525a (13)	6	0.31	0.82	7	0.15	0.85	8	0.08	0.93	3	0.38	0.38*	6	0.77	0.83*
Card3-204/526 (20)	9	0.10	0.82	7	0.15	0.86	12	0.45	0.88	6	0.40	0.75	6	0.50	0.81
Card3-204/526ab (14)	8	0.14	0.86	8	0.21	0.90	7	0.14	0.74	6	0.86b	0.84	6	0.86b	0.72*
Card3-207/529 (11)	8	0.33	0.76	10	0.08	0.95	7	0.08	0.79	5	0.42	0.77	7	0.67	0.83*
Suss-114/259 (18)	3	0.06	0.39	7	0.28	0.83	12	0.39	0.89	5	0.61	0.70	7	0.61	0.80
Suss-116/261 (18)	4	0.11	0.57	5	0.17	0.78	13	0.67	0.92	4	0.17	0.67	5	0.50	0.80
Suss-122/275b (19)	6	0.26	0.79	5	0.42	0.83	9	0.15	0.89	5	0.64	0.80	6	0.84b	0.79*
Suss-125/282 (15)	13	0.27	0.95	8	0.07	0.91	14	0.33	0.97	5	0.27	0.84	5	0.53	0.80*
Suss-125/282a (20)	9	0.55	0.89	10	0.25	0.91	15	0.50	0.96	5	0.15	0.74	6	0.50	0.80
Suss-125/283ab (14)	5	0.43	0.77	10	0.64	0.90	16	0.64	0.98	4	0.64	0.79*	5	0.79b	0.73*
Suss-125/284a (13)	10	0.38	0.87	6	0.08	0.89	13	0.54	0.94	4	0.46	0.77	6	0.62	0.80*
Suss-133/305b (12)	5	0.17	0.82	5	0.17	0.66	9	0.42	0.82	4	0.58	0.73*	5	0.83b	0.77*
Suss-134/308a (10)	5	0.00	0.84	5	0.00	0.89	6	0.11	0.93	3	0.00	0.78	5	0.56	0.78
Suss-137/315 (14)	7	0.21	0.77	8	0.36	0.86	7	0.29	0.73	4	0.57	0.75	7	0.64	0.86
Suss-137/315a (12)	5	0.17	0.75	9	0.58	0.92	13	0.50	0.94	4	0.42	0.49*	5	0.58	0.80*
Suss-145/330b (22)	7	0.23	0.83	6	0.27	0.82	9	0.32	0.82	5	0.41	0.75	6	0.83b	0.79*
VdG-13/50ab (10)	2	0.33	0.48*	4	0.33	0.76	5	1,00b	0.79	3	0.00	0.68	5	0.50	0.83*

Tabela 4: Continuação.

Locus	SmBr19			SmBr16			SmBr17			SmBr9			SmBr15		
Moradores (N)	Alelos	Ho	He	Alelos	Ho	He	Alelos	Ho	He	Alelos	Ho	He	Alelos	Ho	He
VdG-26/637a (17)	7	0.35	0.71	7	0.12	0.73	8	0.59	0.73	4	0.41	0.49*	7	0.59	0.80*
VdG-26/70 (15)	7	0.20	0.81	7	0.33	0.87	7	0.40	0.75	4	0.27	0.68	6	0.60	0.80*
VdG-46/543 (16)	7	0.06	0.89	8	0.31	0.85	9	0.31	0.89	5	0.44	0.76	5	0.63	0.83*
VdG-46/543a (14)	7	0.14	0.90	9	0.07	0.93	3	0.07	0.21	3	0.21	0.36*	5	0.57	0.78*
VdG-46/91ab (13)	3	0.15	0.61	5	0.31	0.76	12	0.38	0.96	5	0.77	0.81*	5	1,00b	0.78*
VdG-75/141 (10)	5	0.00	0.79	7	0.60	0.90*	7	0.20	0.71	3	0.30	0.54*	5	0.60	0.84
VdG-75/145 (11)	4	0.45	0.82	4	0.36	0.79	5	0.09	0.75	5	0.28	0.48	6	0.36	0.89
VdG-83/152ab (11)	5	0.18	0.84	4	0.09	0.83	10	0.36	0.96	4	0.73b	0.71*	5	0.73	0.76*
VdG-97/205b (11)	7	0.09	0.76	6	0.18	0.88	7	0.18	0.87	6	0.82	0.84*	6	0.91b	0.74*
VdG-99/210 (10)	4	0.14	0.82	4	0.00	0.86	5	0.14	0.86	4	0.43	0.87	5	0.57	0.84

*significantemente não diferente de Ho ($p<0,05$)

a=amostras coletadas após tratamento quimioterápico

b=Ho>He

(N)=tamanho das amostras

As infrapopulações de parasitos foram comparadas par-a-par e estimou-se o RST entre elas em três situações diferentes: a) todas as 53 infrapopulações; b) as 44 infrapopulações amostradas apenas uma vez; e c) as 9 infrapopulações amostradas pré e pós-tratamento. A partir destes dados foram montadas matrizes de distância utilizando como método a soma dos quadrados das diferenças de tamanhos – RST (Step-wise-mutation model). Com base nestas matrizes foram gerados fenogramas utilizando o método de UPGMA com o programa MEGA 3.1. A Figura 8 mostra o fenograma construído com as 53 infrapopulações. A topologia do fenograma mostra as infrapopulações formando grupos distintos, porém não houve relação com a origem (subáreas) das amostras.

Os percentuais de variação, os índices de fixação em cada nível e os respectivos testes de significância estão apresentados na Tabela 5. Os testes de significância são relativos às comparações entre os valores obtidos a partir das permutações de dados (1023 permutações neste trabalho) e os valores observados nas análises. Os menores índices de fixação foram observados entre as subáreas, com percentagem de variação de 0,32% e FCT de 0,00316. As diferenças entre as infrapopulação foram da ordem de 19,49%, FSC de 0,19464 (p=0,00000). Entre cada parasito de uma infrapopulação as diferenças foram de 62,49% com FIS de 0,77835 (p=0,00000). Entre todos os parasitos as diferenças foram de 17,79% com FIT de 0,82205 (p=0,00000).

Tabela 5: Análise de variância molecular (AMOVA) e índices de fixação entre as 53 infrapopulações estudadas.

Fonte de variação	Percentagem de variação
Entre subáreas	0,32
Entre infrapopulações	19,40
Entre parasitos dentro das infrapopulações	62,49
Entre todos os parasitos	17,79
Índices de fixação	Testes de significância
FCT: 0,00316	P = 0,37146
FSC: 0,19464	P = 0,00000*
FIS: 0,77835	P = 0,00000*
FIT: 0,82205	P = 0,00000*

*significativo para p<0,05

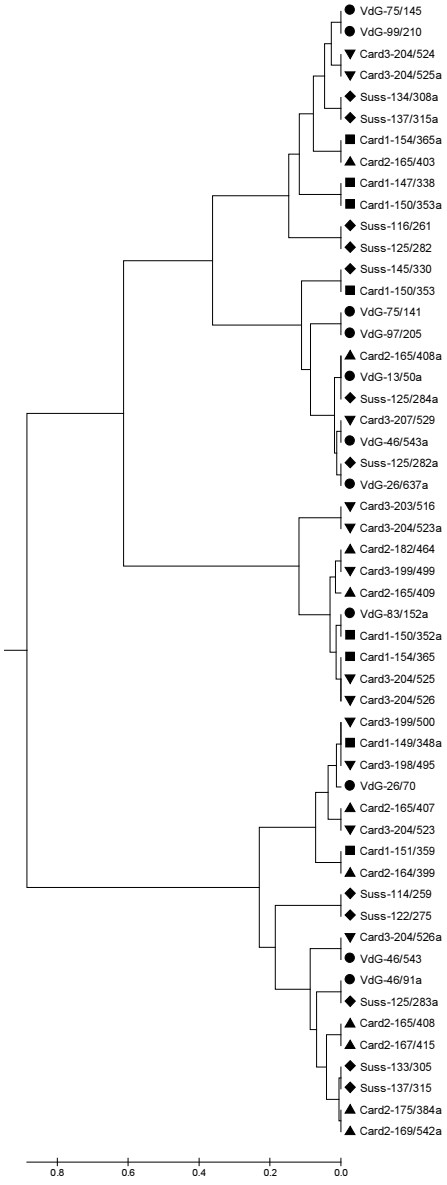


Figura 8 – Fenograma UPGMA baseado em Rst entre as 53 populações.

A Figura 9 mostra o fenograma construído a partir da análise das 44 infrapopulações de parasitos amostradas apenas uma vez. A topologia deste também mostra uma formação de grupos entre as populações, entretanto, sem relação com a origem geográfica das amostras.

Nesta análise também, os menores índices de fixação foram observados entre as subáreas com percentagem de variação de 0,09% e FCT de 0,00095. As diferenças entre as infrapopulação foram da ordem de 19,05%, FSC de 0,19071 (p=0,00000). Entre cada parasito de cada infrapopulação, as diferenças foram de 61,98% com FIS de 0,76657 (p=0,00000). Entre todos os parasitos as diferenças foram de 18,87% com FIT de 0,81127 (p=0,00000) (Tabela 6).

Tabela 6: Análise de variância molecular (AMOVA) e índices de fixação entre as 44 infrapopulações amostradas apenas uma vez.

Fonte de variação	Percentagem de variação
Entre Subáreas	0,09
Entre infrapopulações	19,05
Entre parasitos dentro das infrapopulações	61,98
Entre todos os parasitos	18,87
Índices de fixação	Testes de significância
FCT: 0,00095	P = 0,44770
FSC: 0,19071	P = 0,00000*
FIS: 0,76657	P = 0,00000*
FIT: 0,81127	P = 0,00000*

*significativo para p<0,05

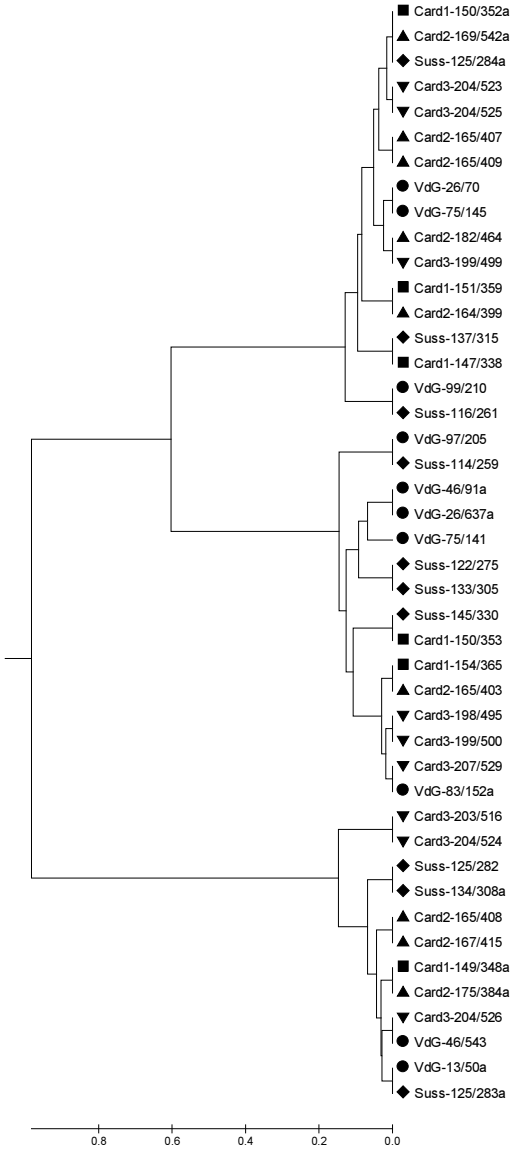


Figura 9 – Fenograma UPGMA baseada em RST entre as 44 populações pré-tratamento.

A Figura 10 mostra o fenograma construído a partir do resultado da análise das 18 infrapopulações de parasitos amostradas pré e pós-tratamento. A topologia deste fenograma também mostra que estas infrapopulações formam grupos distintos, entretanto, sem relação com o tratamento.

Novamente, os menores índices de fixação foram observados entre os grupos de populações (pré e pós-tratamento), com percentagem de variação de 0,47% e FCT de 0,00469, e significativo para $p<0,05$ ($p = 0,29814$). As diferenças entre as cada infrapopulação foram da ordem de 12,42%, FSC de 0,12480 ($p=0,00000$). Entre os parasitos de uma infrapopulação, as diferenças foram de 72,07% com FIS de 0,82730 ($p=0,00000$). Entre todos os parasitos as diferenças foram de 15,04% com FIT de 0,84956 ($p=0,00000$) (Tabela 7).

Tabela 7: Análise de variância molecular (AMOVA) e índices de fixação entre as 18 infrapopulações amostradas pré e pós-tratamento.

Fonte de variação	Percentagem de variação
Entre Grupos	0,47
Entre infrapopulações	12,42
Entre parasitos dentro das infrapopulações	72,07
Entre todos os parasitos	15,04
Índices de fixação	Testes de significância
FCT: 0,00469	P = 0,29814
FSC: 0,12480	P = 0,00000*
FIS: 0,82730	P = 0,00000*
FIT: 0,84956	P = 0,00000*

*significativo para $p<0,05$

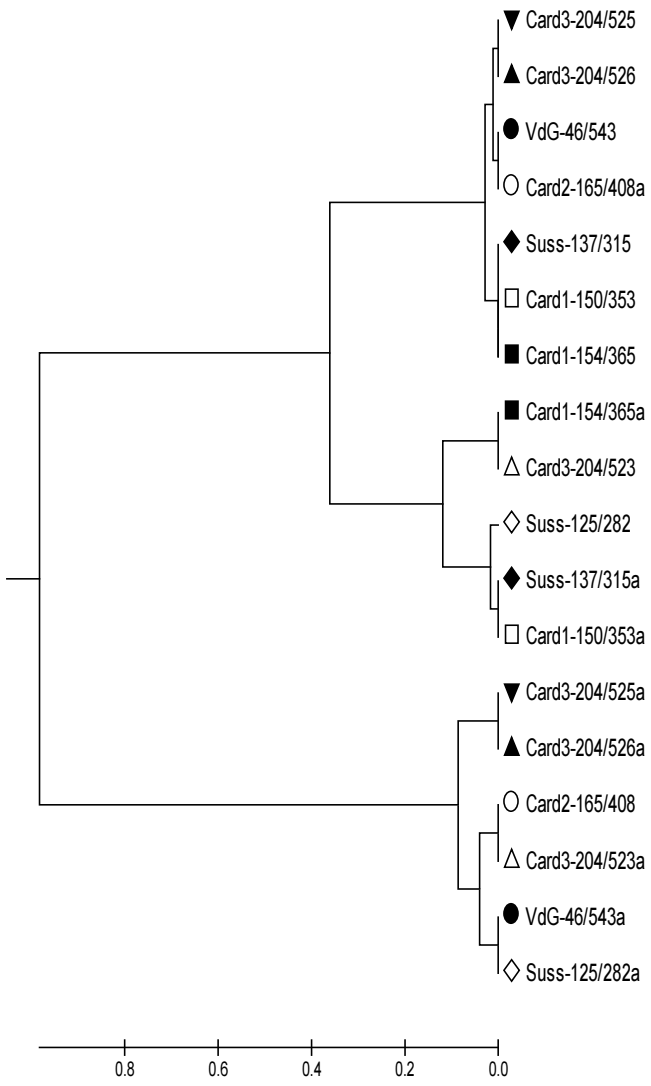


Figura 10 – Fenograma UPGMA baseada em RST entre as populações pré e pós-tratamento.

VI – DISCUSSÃO

Analisamos a estruturação e a variabilidade genética de infrapopulações de *S. mansoni* que infectavam moradores de uma área endêmica e tentamos avaliar o impacto ocorrido nesta variabilidade após o tratamento com a quimioterapia usual. Vimos que estas infrapopulações não apresentam uma estruturação que se possa relacionar à distribuição geográfica dos moradores dos quais as amostras foram isoladas. Os resultados indicam que não há restrição ao fluxo gênico dos parasitos nesta área endêmica. Também não foi possível observar relações entre tratamento quimioterápico e as distâncias entre as populações. Porem, já temos coletado, material biológico para a análise do impacto da quimioterapia após quatro anos de tratamento em massa realizado anualmente.

Com exceção de 13 das 53 infrapopulações, em todas as outras e para todos os 5 *loci* a heterozigosidade observada (Ho) foi sempre menor que a esperada (He). Em 39 das 53 infrapopulações esta diferença foi significativa para $p<0,05$, configurando desvio do equilíbrio de Hardy-Weinberg em pelo menos um *locus*. Estes resultados estão de acordo com os apresentados por Rodrigues e cols., (2002a), Stohler e cols., (2004) e Gower e cols., (2007), que reportaram um déficit de heterozigotos em populações de campo. Uma explicação plausível para este déficit pode ser o fato de as populações de parasitos de cada morador sofrer a influencia do efeito fundador ao se estabelecer neste hospedeiro. Outra explicação poderia ser o acasalamento entre vermes geneticamente relacionados (Gower et al., 2007). Uma terceira explicação poderia ser uma possível estruturação sexo-específica dos parasitos, fato já relatado por Prugnolle e cols.,(2002). Naquele trabalho os autores mostraram que parece haver uma competição entre fêmeas, eliminando aquelas com genótipos diferentes das já estabelecidas em determinado hospedeiro.

Em todos os níveis nos quais tentamos analisar a estruturação destas populações: entre as subáreas, entre as infrapopulações, entre os parasitos de cada infrapopulação, ou entre todos os parasitos; vimos que elas apresentam os maiores percentuais de distância quando comparamos os indivíduos internamente às populações (parasito a parasito) e os menores percentuais quando comparamos os grupos (as infrapopulações), mesmo quando se tratam de grupos pré e pós-tratamento. Em todos os nossos testes os índices de fixação entre as subáreas (FCT) foram inferiores a 0,05, com $p<0,05$ em todos os casos. Segundo Wright (Wright S, 1965) valores para os índices de fixação entre 0,05 e 0,15 indicam uma distância genética moderada, de 0,15 a 0,25, grande e maiores que 0,25, muito grande. Nossos resultados sugerem um grande fluxo gênico entre os parasitos das regiões estudadas (possivelmente um grande transito dos moradores e/ou vetores de uma região a outra) e que não existem barreiras físicas que impeçam a migração dos parasitos de uma subárea para a outra. Estes resultados

são contrários aos apresentados por Agola e cols.,(2006) para populações de parasitos do Kenia, entre os quais foi observada uma forte estruturação geográfica das populações. As amostras analisadas por Agola e cols, eram provenientes de pontos distantes vários quilômetros uns dos outros, nossas amostras não apresentavam um distanciamento geográfico tão grande com as subáreas não distanciando umas das outras mais que 4 ou 5 quilômetros (Figura 5).

Os indicadores FIS (0,77 a 0,83), que indicam as diferenças entre os parasitos de uma infrapopulação e FIT (0,81 a 0,85), que indicam as diferenças entre todos os parasitos, indicam um alto nível de endogamia entre os indivíduos das populações estudadas. Este fato pode ser relacionado às múltiplas infecções dos moradores, ocorridas em diferentes localidades nas subáreas, o que poderia ser explicado pelas pequenas diferenças encontradas entre estas subáreas.

VII – CONSIDERAÇÕES FINAIS

A proposta do nosso trabalho era desenvolver uma metodologia rápida e acurada que nos permitisse estudar a variabilidade genética de populações de *S. mansoni* infectando moradores de uma área endêmica. A acurácea dos microssatélites para estudos de variabilidade genética já foi extensivamente comprovada em diferentes organismos, incluindo o *S. mansoni*. De 2002, quando começaram a ser descritos, até o momento, apenas 30 *loci* polimórficos de microssatélites haviam sido descritos em *S. mansoni.* Em nosso trabalho acrescentamos à literatura, mais 11 *loci* polimórficos de microssatélites deste parasito, além de um 12º que pode ser funcional na identificação de outras espécies do gênero. Outra vantagem destes novos marcadores é que todos são de origem genômica, ao contrário dos anteriores que são originários de cDNA.

Quanto à rapidez, além de apresentarmos a utilização da PCR-multiplex na genotipagem de *S. mansoni*, minimizando o tempo de análise, apresentamos também, a utilização de ovos do parasito como fonte de DNA para esta genotipagem. O uso de ovos como material biológico reduz o tempo necessário entre a coleta de material e os resultados da análise em aproximadamente 90 dias, e elimina os filtros genéticos representados por caramujos e vertebrados de laboratório necessários à manutenção do ciclo do parasito.

O uso destas metodologias nos possibilitou fazer a análise genética de, 53 populações de parasitos coletadas de moradores de área endêmica para esquistossomose. Pudemos, verificar que não houve uma estruturação das populações em infrapopulações geograficamente definidas e ainda determinar que não houve influência da quimioterapia sobre a variabilidade dos 9 pares de populações de parasitos estudadas.

Temos, além dos marcadores já testados e citados na literatura, um grande número de outros marcadores que poderão ser de grande utilidade em futuras análises genotípicas deste parasito. Estes marcadores poderão ser ainda utilizados na confecção de mapas gênicos assim que os trabalhos, já em andamento, de montagem e anotação do genoma do *S. mansoni* estiverem concluídos, podendo num futuro de médio prazo, conduzir à descoberta de genes que poderão ser alvos de quimioterapias mais eficazes e mesmo de vacinas.

No tocante à diversidade de material biológico, temos, ainda por analisar, outras amostras coletadas em Virgem das Graças no decorrer deste projeto e amostras coletadas em outras áreas endêmicas do Estado de Minas Gerais e de outros estados do Brasil. Estes estudos nos darão uma melhor visão da dinâmica das populações de *S. mansoni* no país e até mesmo a possibilidade de compará-las com populações de outros paises onde a doença seja endêmica.

Este trabalho abre perspectivas promissoras no estudo e no entendimento de vários mecanismos biológicos envolvidos nas interações parasito-hospedeiro vertebrado, na patogenia e epidemiologia, através de estudos de ligação com genes de resistência, genes responsáveis por traços patogênicos, ou alvos para uma futura quimioterapia da esquistossomose, doença que ainda hoje afeta, de maneira séria, e em certos casos, fatal, cerca de 200 milhões de pessoas em mais de 70 paises.

VIII – BIBLIOGRAFIA

BIBLIOGRAFIA

Agola LE, Mburu DN, Dejong RJ, Mungai BN, Muluvi GM, Njagi EN, Loker ES, Mkoji GM. Microsatellite typing reveals strong genetic structure of *Schistosoma mansoni* from localities in Kenya. Infect Genet Evol. 2006;6(6):484-490

Ali PO, Simpson AJG, Allen R, Waters AP, Humphries CJ, Johnston DA, Rollinson D. Sequence of a small subunit rRNA gene of *Schistosoma mansoni* and its use in phylogenetic analysis. Mol Biochem Parasitol. 1991;46,201-208

Araujo KC, Rosa e Silva, Barbosa CS, Ferrari TC. Clinical-epidemiological profile of children with schistosomal myeloradiculopathy attended at the Instituto Materno-Infantil de Pernambuco. Mem Inst Oswaldo Cruz. 2006;101 Suppl 1,149-156

Araujo N, Katz N, Dias EP, Souza CP. Susceptibility to chemotherapeutic agents of strains of *Schistosoma mansoni* isolated from treated and untreated patients. Am J Trop Med Hyg. 1980;29(5):890-894

Barral V, This P, Imbert-Establet D, Combes C, Delseny M. Genetic variability and evolution of the *Schistosoma* genome analysed by using random amplified polymorphic DNA markers. Mol Biochem Parasitol. 1993;59,211-222

Bennett JL, Day T, Liang FT, Ismail M, Farghaly A. The development of resistance to anthelmintics: a perspective with an emphasis on the antischistosomal drug praziquantel. Exp Parasitol. 1997;87(3):260-267

Bichara M, Pinet I, Schumacher S, Fuchs RP. Mechanisms of dinucleotide repeat instability in *Escherichia coli*. Genetics. 2000;154(2):533-542

Blair L, Webster JP, Barker GC. Isolation and characterization of polymorphic microsatellite markers in *Schistosoma mansoni* from Africa. Molecular Ecology Notes. 2001;1(1-2):93-95

Blanton RE, Salam EA, Ehsan A, King CH, Goddard KA. Schistosomal hepatic fibrosis and the interferon gamma receptor: a linkage analysis using single-nucleotide polymorphic markers. Eur J Hum Genet. 2005;13(5):660-668

Brindley PJ, Heath S, Waters AP, McCutchan TF, Sher A. Characterization of a programmed alteration in an 18S ribosomal gene that accompanies the experimental induction of drug resistance in *Schistosoma mansoni*. Proc Natl Acad Sci USA. 1991;88,7754-7758

Campos R, Moreira AA, Sette H, Jr., Chamone DA, Da Silva LC. Hycanthone resistance in a human strain of *Schistosoma mansoni*. Trans R Soc Trop Med Hyg. 1976;70(3):261-262

Cioli D. Praziquantel: is there real resistance and are there alternatives? Curr Opin Infect Dis. 2000;13(6):659-663

Cioli D & Pica-Mattoccia L. Praziquantel. Parasitol Res. 2003;90 Supp 1,S3-S9

Cioli D, Pica-Mattoccia L, Moroni R. *Schistosoma mansoni*: hycanthone/oxamniquine resistance is controlled by a single autosomal recessive gene. Exp Parasitol. 1992;75(4):425-432

Coles GC, Bruce JI, Kinoti GK, Mutahi WT, Dias EP, Katz N. Drug resistance in schistosomiasis. Trans R Soc Trop Med Hyg. 1986;80(2):347-347

Coles GC, Mutahi WT, Kinoti GK, Bruce JI, Katz N. Tolerance of Kenyan *Schistosoma mansoni* to oxamniquine. Trans R Soc Trop Med Hyg. 1987;81(5):782-785

Cummings CJ & Zoghbi HY. Trinucleotide repeats: mechanisms and pathophysiology. Annu Rev Genomics Hum Genet. 2000;1,281-328

Curtis J, Sorensen RE, Page LK, Minchella DJ. Microsatellite loci in the human blood fluke *Schistosoma mansoni* and their utility for other schistosome species. Mol Ecol Notes. 2001;1(3):143-145

Despres L, Imbert-Establet D, Combes C, Bonhomme F. Molecular evidence linking hominid evolution to recent radiation of schistosomes (Platyhelminthes: Trematoda). Mol Phyl Evol. 1992;1(4):295-304

Despres L, Imbert-Establet D, Combes C, Bonhomme F, Monneron M. Isolation and polymorphism is mitochondrial DNA from *Schistosoma mansoni*. Mol Biochem Parasitol. 1991;47,139-142

Dias LCS, Bruce JI, Coles GC. Strain variation in the infectivity of *Schistosoma mansoni* for *Biomphalaria glabrata*. Rev Inst Med trop Sao Paulo. 1988;30(2):86-90

Dias LCS, Jesus Pedro R, Deberaldini E. Use of praziquantel in patients with schistosomiasis mansoni previously treated with oxamniquine and/or hycanthone: resistance of *Schistosoma mansoni* to schistosomicidal agents. Trans R Soc Trop Med Hyg. 1982;76,652-659

Dias-Neto E, Santos FR, Pena SDJ, Simpson AJG. Sex determination by low stringency PCR (LS-PCR). Nucleic Acids Research. 1993a;21(3):763-764

Dias-Neto E, Souza CP, Rollinson D, Katz N, Pena SDJ, Simpson AJG. The random amplification of polymorphic DNA allows the identification of strains and species of schistosome. Mol Biochem Parasitol. 1993b;57,83-88

Doenhoff MJ & Pica-Mattoccia L. Praziquantel for the treatment of schistosomiasis: its use for control in areas with endemic disease and prospects for drug resistance. Expert Rev Anti Infect Ther. 2006;4(2):199-210

Drescher KM, Rogers EJ, Bruce JI, Katz N, Dias LC, Coles GC. Response of drug resistant isolates of *Schistosoma mansoni* to antischistosomal agents. Mem Inst Oswaldo Cruz. 1993;88(1):89-95

Durand P, Sire C, Theron A. Isolation of microsatellite markers in the digenetic trematode *Schistosoma mansoni* from Guadeloupe island. Mol Ecol. 2000;9,997-998

El Sayed NM, Bartholomeu D, Ivens A, Johnston DA, LoVerde PT. Advances in schistosome genomics. Trends Parasitol. 2004;20(4):154-157

Ellegren H. Microsatellite mutations in the germline: implications for evolutionary inference. Trends Genet. 2000;16(12):551-558

Everett CM & Wood NW. Trinucleotide repeats and neurodegenerative disease. Brain. 2004;127(Pt 11):2385-2405

Excoffier L, Lawlor S, Schneider S. Arlequin ver.3,0: An integrated software package for population genetics data analysis.2005;

Excoffier L, Smouse PE, Quattro JM. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics. 1992;131(2):479-491

Fantappie MR, Freebern WJ, Osman A, LaDuca J, Niles EG, LoVerde PT. Evaluation of *Schistosoma mansoni* retinoid X receptor (SmRXR1 and SmRXR2) activity and tissue distribution. Mol Biochem Parasitol. 2001;115(1):87-99

Ferrari TC. Involvement of central nervous system in the schistosomiasis. Mem Inst Oswaldo Cruz. 2004;99(5 Suppl 1):59-62

Field D & Wills C. Long, polymorphic microsatellites in simple organisms. Proc Biol Sci. 1996;263(1367):209-215

Fletcher M, LoVerde PT, Kuntz RE. Electrophoretic differences between *Schistosoma mansoni* and *S. rodhaini*. J Parasitol. 1981a;67(4):593-595

Fletcher M, LoVerde PT, Richards CS. *Schistosoma mansoni*: electrophoretic characterization of strains selected for different levels of infectivity to snails. Exp Parasitol. 1981b;52(3):362-370

Fletcher M, LoVerde PT, Woodruff DS. Genetic variation in *Schistosoma mansoni*: enzyme polymorphisms in populations from Africa, southwest Asia, south America, and the west indies. Am J Trop Med Hyg. 1981c;30(2):406-421

Franco GR, Valadao AF, Azevedo V, Rabelo EML. The *Schistosoma* gene discovery program: state of the art. Inter J Parasitol. 2000;30,453-463

Gasser RB, Morahan G, Mitchell GF. Sexing single larval stages of *Schistosoma mansoni* by polymerase chain reaction. Mol Biochem Parasitol. 1991;47,255-258

Gazzinelli A, Velasquez-Melendez G, Crawford SB, LoVerde PT, Correa-Oliveira R, Kloos H. Socioeconomic determinants of schistosomiasis in a poor rural area in Brazil. Acta Trop. 2006;99(2-3):260-271

Gibbs HL, Brooke ML, Davies NB. Analysis of genetic differentiation of host races of the common cuckoo *Cuculus canorus* using mitochondrial and microsatellite DNA variation. Proc R Soc Lond B. 1996;263,89-96

Gower CM, Shrivastava J, Lamberton PH, Rollinson D, Webster BL, Emery A, Kabatereine NB, Webster JP. Development and application of an ethically and epidemiologically advantageous assay for the multi-locus microsatellite analysis of *Schistosoma mansoni*. Parasitology. 2007;134(Pt 4):523-536

Grady WM. Genomic instability and colon cancer. Cancer Metastasis Rev. 2004;23(1-2):11-27

Groppe K, Sanders I, Wiemken A, Boller T. A microsatellite marker for studying the ecology and diversity of fungal endophytes (*Epichloe* spp.) in grasses. Applied and Environmental Microbiology. 1995;61(11):3943-3949

Gryseels B, Mbaye A, de Vlas SJ, Stelma FF, Guisse F, Van Lieshout L, Faye D, Diop M, Ly A, Tchuem-Tchuente LA, Engels D, Polman K. Are poor responses to praziquantel for the treatment of *Schistosoma mansoni* infections in Senegal due to resistance? An overview of the evidence. Trop Med Int Health. 2001;6(11):864-873

Gryseels B, Polman K, Clerinx J, Kestens L. Human schistosomiasis. Lancet. 2006;368(9541):1106-1118

Hartl DL & Clark AG. Principles of population genetics1997;3ed. 23 Plumtree Road, Sunderland, MA, 01375 USA: Sinauer Associates, Inc., 542p, 1997.

Henegariu O, Heerema NA, Dlouhy SR, Vance GH, Vogt PH. Multiplex PCR: critical parameters and step-by-step protocol. BioTechniques. 1997;23(3):504-511

Hentschel CC. Homocopolymer sequences in the spacer of a sea urchin histone gene repeat are sensitive to S1 nuclease. Nature. 1982;295(5851):714-716

Hoekstra R, Criado-Fornelio A, Fakkeldij J, Bergman J, Roos MH. Microsatellites of the parasitic nematode *Haemonchus contortus*: polymorphism and linkage with a direct repeat. Mol Biochem Parasitol. 1997;89,97-107

Hotez PJ, Molyneux DH, Fenwick A, Ottesen E, Ehrlich SS, Sachs JD. Incorporating a rapid-impact package for neglected tropical diseases with programs for HIV/AIDS, tuberculosis, and malaria. PLoS Med. 2006;3(5):e102-e102

Ishibashi Y, Saitoh T, Abe S, Yoshida MC. Sex-related spatial kin structure in a spring population of grey-sided voles *Clethrionomys rufocanus* as revealed by mitochondrial and microsatellite DNA analyses. Mol Ecol. 1997;6,63-71

Jannotti-Passos LK, Vidigal THDA, Dias-Neto E, Pena SDJ, Simpson AJG, Dutra WO, Souza CP, Carvalho-Parra JF. PCR amplification of the mitochondrial DNA minisatellite region to detect *Schistosoma mansoni* infection in *Biomphalaria glabrata* snails. J Parasitol. 1997;83(3):395-399

Jarne P & Lagoda JL. Microsatellites, from molecules to populations and back. Trends in Ecology and Evolution. 1996;11,424-429

Jarne P, Viard F, Delay B, Cuny G. Variable microsatellites in the highly selfing snail *Bulinus truncatus* (Basommatophora: Planorbidae). Mol Ecol. 1994;3,527-528

Johnston DA, Dias-Neto E, Simpson AJG, Rollinson D. Opening the can of worms: molecular analysis of schistosome populations. Parasitology Today. 1993;9(8):286-291

Jourdane J & Theron A. Larval Development: Eggs to Cercariae.1987;London Academic Press(3):83-113. London

Kalendar R. Fast PCR2003;(3.6). Finland. Institute of Biotechnology. University of Helsinki.

Katz N, Dias EP, Araujo N, Souza CP. Estudos de uma cepa humana de *Schistosoma mansoni* resistente a agentes esquistossomicidas. Rev Soc Bras Med Trop. 1973;7,381-387

Katz N & Peixoto SV. Análise crítica da estimativa do número de portadores de esquistossomose mansoni no Brasil. Rev Soc Bras Med Trop. 2000;33(3):303-308

King DG, Soller M, Kashi Y. Evolutionary tuning knobs. Endeavour. 1997;21(1):36-40

Kokoska RJ, Stefanovic L, Tran HT, Resnick MA, Gordenin DA, Petes TD. Destabilization of yeast micro- and minisatellite DNA sequences by mutations affecting a nuclease involved in okazaki fragment processing (*rad*27) and DNA polymerase δ (*pol*3-*t*). Mol Cell Biol. 1998;18(5):2779-2788

Kolodner RD & Marsischky GT. Eukaryotic DNA mismatch repair. Curr Opin Genet Dev. 1999;9(1):89-96

Kumar S, Tamura K, Nei M. MEGA3: Integrated software for Molecular Evolutionary Genetics Analysis and sequence alignment. Briefings in Bioinformatics. 2004;5,150-163

Le TH, Blair D, Agatsuma T, Humair P-F, Campbell NJH, Iwagami M, Littlewood DTJ, Peacock B, Johnston DA, Bartley J, Rollinson D, Herniou EA, Zarlenga DS, McManus DP. Phylogenies inferred form mitochondrial gene orders - a cautionary tale from the parasitic flatworms. Mol Biol Evol. 2000;17(7):1123-1125

Levinson G & Gutman GA. High frequencies of short frameshifts in poly-CA/TG tandem repeats borne by bacteriophage M13 in *Escherichia coli* K-12. Nucleic Acids Res. 1987;15(13):5323-5338

Lopes R, Lopes MT, Carneiro MS, Matta FP, Camargo LE, Vieira ML. Linkage and mapping of resistance genes to *Xanthomonas axonopodis* pv. passiflorae in yellow passion fruit. Genome. 2006;49(1):17-29

LoVerde PT, DeWald J, Minchella DJ. Further studies of genetic variation in *Schistosoma mansoni*. J Parasitol. 1985a;71(6):732-734

LoVerde PT, DeWald J, Minchella DJ, Bosshardt SC, Damian RT. Evidence for host-induced selection in *Schistosoma mansoni*. J Parasitol. 1985b;71(3):297-301

Marx KA, Bizzaro JW, Blake RD, Hsien TM, Feng TL. Experimental DNA melting behavior of the three major *Schistosoma* species. Mol Biochem Parasitol. 2000;107(2):303-307

McCutchan TF, Simpson AJG, Mullins JA, Sher A, Nash TE, Lewis F, Richards C. Differentiation of schistosomes by species, strain, and sex by using cloned DNA markers. Proc,Natl Acad Sci USA. 1984;81,889-893

McManus DP & Hope M. Molecular variation in the human schistosomes. Acta Tropica. 1993;53,255-276

Morgan JA, Dejong RJ, Adeoye GO, Ansa ED, Barbosa CS, Bremond P, Cesari IM, Charbonnel N, Correa LR, Coulibaly G, D'Andrea PS, DE Souza CP, Doenhoff MJ, File S, Idris MA, Incani RN, Jarne P, Karanja DM, Kazibwe F, Kpikpi J, Lwambo NJ, Mabaye A, Magalhaes LA, Makundi A, Mone H, Mouahid G, Muchemi GM, Mungai BN, Sene M, Southgate V, Tchuente LA, Theron A, Yousif F, Zanotti-Magalhaes EM, Mkoji GM, Loker ES. Origin and diversification of the human parasite *Schistosoma mansoni*. Mol Ecol. 2005;14(12):3889-3902

Navarro MC, Cesari IM, Incani RN. Isoenzyme studies in one brazilian and two venezuelan strains of *Schistosoma mansoni*. Comp Biochem Physiol. 1992;102B(3):471-474

Oliveira ED, Pádua JG, Zucchi MI, Vencovsky R, Vieira MLC. Origin, evolution and genome distribution of microsatellites. Gen and Mol Biol. 2006;29(2):294-307

Oliveira GC & Bahia D. The genome of *Schistosoma mansoni*. Proceedings of the Third Brazilian Syposium of Mathematical and Computational Biology. 2004;1,101-115

Oliveira GC & Johnston DA. Mining the schistosome DNA sequence database. TRENDS in Parasitol. 2001;17(10):501-503

Oliveira GC, Rodrigues NB, Romanha AJ, Bahia D. Genome and Genomics of Schistosomes. Canadian Journal of Zoology. 2004;82,375-390

Oliveira RP, Melo AI, Macedo AM, Chiari E, Pena SD. The population structure of *Trypanosoma cruzi*: expanded analysis of 54 strains using eight polymorphic CA-repeat microsatellites. Mem Inst Oswaldo Cruz. 1999;94 Suppl 1,65-70

Paraense WL & Correa LR. Susceptibility of Australorbis tenagophilus to infection with *Schistosoma mansoni*. Rev Inst Med Trop Sao Paulo. 1963a;5,23-29

Paraense WL & Correa LR. Variation in susceptibility of populations of Australorbis glabratus to a strain of *Schistosoma mansoni*. Rev Inst Med Trop Sao Paulo. 1963b;5,15-22

Pearson CE, Edamura KN, Cleary JD. Repeat instability: mechanisms of dynamic mutations. Nat Rev Genet. 2005;6(10):729-742

Pellegrino J & Coelho PMZ. *Schistosoma mansoni*, wandering capacity of a worm couple. J Parasitol. 1978;64(1):181-182

Pellegrino J & Siqueira AF. A perfusion technic for recovery of *Schistosoma mansoni* from experimentally infected guinea pigs. Rev Bras Malariol Doencas Trop. 1956;8(4):589-597

Pena HB, Souza CP, Simpson AJG, Pena SDJ. Intracellular promiscuity in *Schistosoma mansoni*: nuclear transcribed DNA sequences are part of a mitochondrial minisatellite region. Proc,Natl Acad Sci USA. 1995;92,915-919

Pinheiro F, Pontes L, Gene M, Huguet E, Costa JP, Moreno P. Population study of the HUMTH01, HUMVWA31A, HUMF13A1, and HUMFES/FPS STR polymorphisms in the north of Portugal. Journal of Forensic Sciences. 1997;42(1):121-124

Pinto PM, Brito CFA, Jannoti Passos LK, Tendler M, Simpson AJG. Contrasting genomic variability between clones from field isolates and laboratory populations of *Schistosoma mansoni*. Mem Inst Oswaldo Cruz, Rio de Janeiro. 1997;92(3):409-414

Prugnolle F, Choisy M, Theron A, Durand P, de Meeus T. Sex-specific correlation between heterozygosity and clone size in the trematode *Schistosoma mansoni*. Mol Ecol. 2004a;13(9):2859-2864

Prugnolle F, de Meeus T, Durand P, Sire C, Theron A. Sex-specific genetic structure in *Schistosoma mansoni*: evolutionary and epidemiological implications. Mol Ecol. 2002;11(7):1231-1238

Prugnolle F, de Meeus T, Pointier JP, Durand P, Rognon A, Theron A. Geographical variations in infectivity and susceptibility in the host-parasite system *Schistosoma mansoni* / *Biomphalaria glabrata* : no evidence for local adaptation. Parasitology. 2006;133(Pt3):313-319

Prugnolle F, Theron A, Durand P, de Meeus T. Test of pangamy by genetic analysis of *Schistosoma mansoni* pairs within its natural murine host in Guadeloupe. J Parasitol. 2004b;90(3):507-509

Prugnolle F, Theron A, Pointier JP, Jabbour-Zahab R, Jarne P, Durand P, de Meeus T. Dispersal in a parasitic worm and its two hosts: consequence for local adaptation. Evolution Int J Org Evolution. 2005;59(2):296-303

Remme JH, Blas E, Chitsulo L, Desjeux PM, Engers HD, Kanyok TP, Kengeya Kayondo JF, Kioy DW, Kumaraswami V, Lazdins JK, Nunn PP, Oduola A, Ridley RG, Toure YT, Zicker F, Morel CM. Strategic emphases for tropical diseases research: a TDR perspective. Trends Parasitol. 2002;18(10):421-426

Rodrigues NB, Silva SA, Pucci MM, Minchella DJ, Sorensen RE, LoVerde PT, Romanha AJ, Oliveira GC. Microsatellite-enriched genomic libraries as a source of polymorphic loci for *Schistosoma mansoni*. Mol Ecol Notes. 2007;7,263-265

Rodrigues NB, Coura-Filho P, Souza CP, Jannoti Passos LK, Dias-Neto E, Romanha AJ. Populational structure of *Schistosoma mansoni* assessed by DNA microsatellites. Inter J Parasitol. 2002a;32(7):843-851

Rodrigues NB, LoVerde PT, Romanha AJ, Oliveira G. Characterization of new *Schistosoma mansoni* microsatellite loci in sequences obtained from public DNA databases and microsatellite enriched genomic libraries. Mem Inst Oswaldo Cruz. 2002b;97 Suppl 1,71-75

Rodriguez N, Lima H, Rodriguez A, Brewster S, Barker DC. Genomic DNA repeat from *Leishmania (Viannia) braziliensis* (Venezuelan strain) containing simples repeats and microsatellites. Parasitology. 1997;115,349-358

Rossi V, Wincker P, Ravel C, Blaineau C, Pages M, Bastien P. Structural organisation of microsatellite families in the *Leishmania* genome and polymorphisms at two $(CA)_n$ loci. Mol Biochem Parasitol. 1994;65,271-282

Rubinsztein DC. Trinucleotide expansion mutations cause diseases which do not conform to classical Medelian expectations1999;1(7):80-97 Oxford University Press.

Russell R, Iribar MP, Lambson B, Brewster S, Blackwell JM, Dye C, Ajioka JW. Intra and inter-specific microsatellite variation in the *Leishmania* subgenus *viannia*. Mol Biochem Parasitol. 1999;103,71-77

Sambrook J, Fritsch EF, Maniatis T. Molecular Cloning: a laboratory manual1989;2nd(2):8.3-15.109. New York, NY-USA. Cold Spring Harbor Laboratory Press.

Sanguinetti CJ, Dias-Neto E, Simpson AJG. Rapid silver staining and recovery of PCR products separated on polyacrylamide gels. BioTechniques. 1994;17(5):915-918

Santos FR, Pena SDJ, Epplen JT. Genetic and population study of a Y-linked tetranucleotide repeat DNA polymorphism with a simple non-isotopic technique. Hum Genet. 1993;90,655-656

Schlotterer C. Evolutionary dynamics of microsatellite DNA. Chromosoma. 2000;109,365-371

Schlotterer C & Pemberton J. The use of microsatellites for genetic analysis of natural populations. EXS. 1994;69,203-214

Schlotterer C & Tautz D. Slippage synthesis of simple sequence DNA. Nucleic Acids Res. 1992;20(2):211-215

Schwarzenbach H, Muller V, Stahmann N, Pantel K. Detection and characterization of circulating microsatellite-DNA in blood of patients with breast cancer. Ann N Y Acad Sci. 2004;1022,25-32

Shaikenov BS, Rysmukhambetova AT, Massenov B, Deplazes P, Mathis A, Torgerson PR. Short report: the use of a polymerase chain reaction to detect *Echinococcus granulosus* (G1 strain) eggs in soil samples. Am J Trop Med Hyg. 2004;71(4):441-443

Short RB & Grossman AI. Conventional giemsa and C-banded karyotypes of *Schistosoma mansoni* and *S. rodhaini*. J Parasitol. 1981;67(5):661-671

Short RB & Menzel MY. Chromosomes of nine species of schistosomes. J Parasitol. 1960;46,273-287

Short RB, Menzel MY, Pathak S. Somatic chromosomes of *Schistosoma mansoni*. J Parasitol. 1979;65(3):471-473

Shrivastava J, Gower CM, Balolong E Jr, Wang TP, Qian BZ, Webster JP. Population genetics of multi-host parasites--the case for molecular epidemiological studies of Schistosoma japonicum using larval stages from naturally infected hosts. Parasitology. 2005;131(Pt 5):617-626

Sia EA, Butler CA, Dominska M, Greenwell P, Fox TD, Petes TD. Analysis of microsatellite mutations in the mitochondrial DNA of *Saccharomyces cerevisiae*. Proc Natl Acad Sci U S A. 2000;97(1):250-255

Sia EA, Kokoska RJ, Dominska M, Greenwell P, Petes TD. Microsatellite instability in yeast: dependence on repeat unit size and DNA mismatch repair genes. Mol Cell Biol. 1997;17(5):2851-2858

Silva LK, Liu S, Blanton RE. Microsatellite analysis of pooled *Schistosoma mansoni* DNA: an approach for studies of parasite populations. Parasitology. 2005;1-8

Simões M, Bahia D, Zerlotini A, Torres K, Artiguenave F, Neshich G, Kuser P, Oliveira G. Single nucleotide polymorphisms identification in expressed genes of *Schistosoma mansoni*. Mol Biochem Parasitol. 2007;in press

Simpson AJG, Dame JB, Lewis FA, McCutchan TF. The arrangement of ribosomal RNA genes in *Schistosoma mansoni* identification of polymorphic structural variants. Eur J Biochem. 1984;139,41-45

Simpson AJG, Sher A, McCutchan TF. The genome of *Schistosoma mansoni*: isolation of DNA, its size, bases and repetitive sequences. Mol Biochem Parasitol. 1982;6,125-137

Sire C, Durand P, Pointier JP, Theron A. Genetic diversity of *Schistosoma mansoni* within and among individual hosts (*Rattus rattus*): infrapopulation differentiation at microspatial scale. Int J Parasitol. 2001;31(14):1609-1616

Slatkin M. Inbreeding coefficients and coalescence times. Genet Res. 1991;58(2):167-175

Smith AL, Huggins MC, Havercroft JC. *Schistosoma mansoni*: adult males and females differentially express antigens encoded by repetitive genomic DNA. Parasitology. 1992;105 (Pt 1),63-69

Sneath P & Sokal R. Numerical taxonomy. Nature. 1962;193,855-860

Solano P, Duvallet G, Dumas V, Cuisance D, Cuny G. Microsatellite markers for genetic population studies in *Glossina palpalis* (Diptera: Glossinidae). Acta Tropica. 1997;65,175-180

Sorensen RE, Curtis J, Minchella DJ. Intraspecific variation in the rDNA its loci of 37-collar-spined echinostomes from North America: implications for sequence-based diagnoses and phylogenetics. J Parasitol. 1998;84(5):992-997

Spotila LD, Hirai H, Rekosh DM, LoVerde PT. A retroposon-like short repetitive DNA element in the genome of the human blood fluke, *Schistosoma mansoni*. Chromosoma (Berl). 1989;97,421-428

Spotila LD, Rekosh DM, Boucher JM, LoVerde PT. A cloned DNA probe identifies the sex of *Schistosoma mansoni* cercariae. Mol Biochem Parasitol. 1987;26(1-2):17-20

Spotila LD, Rekosh DM, LoVerde PT. Polymorphic repeated DNA element in the genome of *Schistosoma mansoni*. Mol Biochem Parasitol. 1991;48,117-120

Steiner JJ, Poklemba CJ, Fjellstrom RG, Elliott LF. A rapid one-tube genomic DNA extraction process for PCR and RAPD analyses. Nucleic Acids Res. 1995;23(13):2569-2570

Stohler RA, Curtis J, Minchella DJ. A comparison of microsatellite polymorphism and heterozygosity among field and laboratory populations of *Schistosoma mansoni*. Int J Parasitol. 2004;34(5):595-601

Strand M, Prolla TA, Liskay RM, Petes TD. Destabilization of tracts of simple repetitive DNA in yeast by mutations affecting DNA mismatch repair. Nature. 1993;365(6443):274-276

Streisinger G & Owen J. Mechanisms of spontaneous and induced frameshift mutation in bacteriophage T4. Genetics. 1985;109(4):633-659

Su X & Wellems TE. Toward a high-resolution *Plasmodium falciparum* linkage map: polymorphic markers from hundreds of simple sequence repeats. Genomics. 1996;33,430-444

TDR Staff . TDR; Making health research work for poor people - Progress 2003-2004. Seventeenth Programme Report2005;17,51-53. Geneve - Switzerland WHO.

Thiele EA, Sorensen RE, Minchella DJ. Genetic diversity and population structuring of *Schistosoma mansoni* in a Brazilian village. Int J Parasitol. 2007;

Toth G, Gaspari Z, Jurka J. Microsatellites in different eukaryotic genomes: survey and analysis. Genome Res. 2000;10(7):967-981

Valadares TE, Coelho PMZ, Pellegrino J, Sampaio IBM. *Schistosoma mansoni*: aspectos da oviposicao da cepa LE em camundongos infectados com um casal de vermes. Rev Inst Med trop Sao Paulo. 1981;23(1):6-11

van Belkum A, Ramesar J, Trommelen G, Uitterlinden AG. Mini- and micro-satellites in the genome of rodent malaria parasites. Gene. 1992;118(1):81-86

van Keulen H, LoVerde PT, Bobek LA, Rekosh DM. Organization of ribosomal RNA genes in *Schistosoma mansoni*. Mol Biochem Parasitol. 1985;15,215-230

Verjovski-Almeida S, DeMarco R, Martins EA, Guimaraes PE, Ojopi EP, Paquola AC, Piazza JP, Nishiyama MY, Jr., Kitajima JP, Adamson RE, Ashton PD, Bonaldo MF, Coulson PS, Dillon GP, Farias LP, Gregorio SP, Ho PL, Leite RA, Malaquias LC, Marques RC, Miyasato PA, Nascimento AL, Ohlweiler FP, Reis EM, Ribeiro MA, Sa RG, Stukart GC, Soares MB, Gargioni C, Kawano T, Rodrigues V, Madeira AM, Wilson RA, Menck CF, Setubal JC, Leite LC, Dias-Neto E. Transcriptome analysis of the acoelomate human parasite *Schistosoma mansoni*. Nat Genet. 2003;35(2):148-157

Verjovski-Almeida S, Venancio TM, Oliveira KC, Almeida GT, DeMarco R. Use of a 44k oligoarray to explore the transcriptome of *Schistosoma mansoni* adult worms. Exp Parasitol. 2007;

Vieira LQ, Correa-Oliveira R, Katz N, Souza CP, Carvalho OS, Araujo N, Sher A, Brindley PJ. Genomic variability in field populations of *Schistosoma mansoni* in Brazil as detected with a ribosomal gene probe. Am J Trop Med Hyg. 1991;44(1):69-78

Weber JL. Human DNA polymorphisms and methods of analysis. Curr Opin Biotechnol. 1990;1(2):166-171

Webster P, Mansour TE, Bieber D. Isolation of a female-specific, highly repeated *Schistosoma mansoni* DNA probe and its use in an assay of cercarial sex. Mol Biochem Parasitol. 1989;36,217-222

WHO. Prevention and control of schistosomiasis and soil-transmitted helminthiasis. World Health Organ Tech Rep Ser. 2002;912,1-57

Wilson IJ & Balding DJ. Genealogical inference from microsatellite data. Genetics. 1998;150,499-510

Wright S. The interpretation of population structure by F-statistics with special regard to systems of mating. Evolution. 1965;19,395-420

Zarlenga DS, Aschenbrenner RA, Lichtenfels JR. Variations in microsatellites sequences provide evidence for population differences and multiple ribosomal gene repeats within *Trichinella pseudospiralis*. J Parasitol. 1996;82(4):534-538

Zheng W, Stacey NE, Coffin J, Strobeck C. Isolation and characterization of microsatellite loci in the goldfish *Carassius auratus*. Mol Ecol. 1995;4,791-792

IX – ANEXOS

9 – ANEXOS

9.1 – Anexo I:

Genome and genomics of schistosomes.

Publicado na revista Canadian Journal of Zoology

Oliveira GC, Rodrigues NB, Romanha AJ, Bahia D. Genome and Genomics of Schistosomes. Canadian Journal of Zoology. 2004;82,375-390

REVIEW / SYNTHÈSE

Genome and genomics of schistosomes[1]

Guilherme Oliveira, Nilton B. Rodrigues, Alvaro J. Romanha, and Diana Bahia

Abstract: Schistosomes infect over 200 million people and 600 million are at risk. Genomics and post-genomic studies of schistosomes will contribute greatly to developing new reagents for diagnostic purposes and new vaccines that are of interest to the biotechnology industry. In this review, the most recent advances in these fields as well as new projects and future perspectives will de described. A vast quantity of data is publicly available, including short cDNA and genomic sequences, complete large genomic fragments, and the mitochondrial genomes of three species of the genus *Schistosoma*. The physical structure of the genome is being studied by physically mapping large genomic fragments and characterizing the highly abundant repetitive DNA elements. Bioinformatic manipulations of the data have already been carried out, mostly dealing with the functional analysis of the genes described. Specific search tools have also been developed. Sequence variability has been used to better understand the phylogeny of the species and for population studies, and new polymorphic genomic markers are currently being developed. The information generated has been used for the development of post-genomic projects. A small microarray detected genes that were differentially expressed between male and female worms. The identification of two-dimensional spots by mass spectrometry has also been demonstrated.

Résumé : Deux cent millions de personnes sont infectées par les schistosomes et 600 millions sont à risque. Les études des schistosomes en génomique et en postgénomique vont contribuer de façon significative à l'élaboration de nouveaux réactifs diagnostiques et de vaccins qui intéresseront l'industrie de la biotechnologie. Notre synthèse décrit les plus récentes percées dans ces domaines, les projets innovateurs et les perspectives futures. Des quantités considérables de données, soit de courtes séquences d'ADNc et du génome, de grands fragments génomiques entiers et les génomes mitochondriaux complets de trois espèces de la genre *Schistosoma* sont disponibles dans le domaine public. La cartographie physique de grands fragments du génome et la caractérisation des éléments répétitifs très nombreux de l'ADN permettent de décrire la structure physique du génome. Des manipulations des données faites par bioinformatique ont servi à l'analyse fonctionnelle des gènes décrits. Des outils exploratoires ont aussi été développés. La variabilité des séquences a permis de mieux comprendre la phylogénie des espèces et elle a été utilisée dans des études démographiques. De nouveaux marqueurs génomiques polymorphes sont actuellement en voie de développement. Ces informations ont servi à la mise au point de projets de postgénomique. L'utilisation de puces d'ADN a servi à reconnaître des gènes dont l'expression est différente chez les vers mâles et femelles. La spectroscopie de masse a aussi permis l'identification de points (spots) bidimensionnels.

[Traduit par la Rédaction]

Introduction

Schistosomiasis is a disease caused by a digenetic trematode parasite of the genus *Schistosoma* that is endemic in over 76 countries, placing over 600 million individuals at risk and with over 200 million individuals infected mostly in sub-Saharan Africa (World Health Organization Partners for Parasite Control 2004). It is estimated that 20 000 schistosomiasis-related deaths occur yearly. According to The World Health Organization (WHO), the disease burden of schistosomiasis calculated by the disability-adjusted life years (DALYs or the number of healthy years of life lost

Received 5 February 2003. Accepted 9 October 2003. Published on the NRC Research Press Web site at http://cjz.nrc.ca on 16 April 2004.

G. Oliveira.[2,3] Avenida Francisco Sales 1111 9º andar, Ala C, Programa de Pós-Graduação e Pesquisa, Santa Casa de Misericórdia de Belo Horizonte, MG 30150-221, Brazil.
N.B Rodrigues, A.J. Romanha, and D. Bahia. Avenida Augusto de Lima 1715, Centro de Pesquisas René Rachou – Fundação Oswaldo Cruz (CPqRR–FIOCRUZ), Belo Horizonte, MG 30190-002, Brazil.

[1]This review is one of a series dealing with aspects of the biology of the phylum Platyhelminthes. This series is one of several virtual symposia on the biology of neglected groups that will be published in the Journal from time to time.
[2]Corresponding author (e-mail: oliveira@cpqrr.fiocruz.br).
[3]Present address: Avenida Augusto de Lima 1715, CPqRR–FIOCRUZ, Belo Horizonte, MG 30190-002, Brazil.

 doi: 10.1139/Z03-220

owing to premature death and disability) is 1 760 000, which is higher than the disease burdens caused by African trypanosomiasis, Dengue, Chagas disease, leprosy, and onchocerciasis (The World Health Organization/Special Programme for Research and Training in Tropical Diseases 2004*a*). The main control strategy involves the use of the effective, safe, and single-dose drug of choice, Praziquantel (The World Health Organization Partners for Parasite Control 2004). The use of Praziquantel has also been effective in decreasing morbidity in endemic areas (Kloetzel 1990; Chitsulo et al. 2000; Kheir et al. 2000). However, the eradication of transmission has not been achieved and several directions on schistosomiasis research, important in achieving this goal, are listed by the WHO/Special Programme for Research and Training in Tropical Diseases (TDR). Among these research areas is the development of new diagnostic tests, drugs, and vaccines (The World Health Organization/Special Programme for Research and Training in Tropical Diseases 2004*b*). Despite accounting for 11.4% of worldwide disease burdens, very few new drugs have been developed for tropical diseases (i.e., only 1% of the total between 1975 and 1999), mainly because of the high cost of drug development in relation to the purchasing power of the endemic countries (Trouiller and Olliaro 1999; Trouiller et al. 2002).

Six vaccine candidates have been selected by WHO for further study: glutathione s-transferase (GST), paramyosin, irradiated vaccine antigen number five (IrV-5), triose phosphate isomerase (TPI), 23-kDa membrane antigen Sm-23/MAP3, and a 14-kDa membrane antigen Sm-14 (Bergquist 1995). There is skepticism related to the possible success of the current vaccine candidates and new approaches for obtaining new molecules are necessary (Gryseels 2000). Nevertheless, research relative to one of the candidates (GST) has advanced and it performed well in phases I and II of preclinical trials (Capron et al. 2001).

Although some advances were achieved, the goals of developing new drugs and vaccines have not yet been reached. These objectives, in addition to the understanding of the biology of the parasite, can potentially benefit significantly from modern technologies and the creative use of genomics and post-genomic approaches. In this review, we will describe the main advances in schistosome genomics. These developments will have a profound impact on our understanding of the biology of schistosomes and in the discoveries of new drug targets and vaccine candidates.

The nuclear genome

The karyotype of *Schistosoma mansoni* consists of eight chromosome pairs (Short and Menzel 1979), with the female forming a heterologous pair WZ (Grossman et al. 1980). The size of the genome is approximately 2.7×10^8 base pairs (bp) with 29.4% GC content (Simpson et al. 1982; Marx et al. 2000). Methylation was not observed in the DNA of *S. mansoni* (Fantappie et al. 2001). C_0t analysis showed that 4%–8% of the DNA was highly repetitive and 40% moderately repetitive (Simpson et al. 1982).

The ends of the chromosomes, or telomeres, are essential for cell viability and consists of hundreds to thousands of tandem repeats of the sequence TTAGGG (Moyzis et al. 1988; Meyne et al. 1989). Typical telomeres were identified by fluorescent in situ hybridization (FISH) assay, with the sequence TTAGGG as a probe, in all eight pairs of *S. mansoni* mitotic metaphase chromosomes (Hirai and LoVerde 1996). Links to other relevant information on the genome composition of *S. mansoni* are listed in Table 1.

Genome sequencing and mapping

Expressed sequence tags (ESTs)

To increase the number of new genes discovered, the initial sequencing efforts have concentrated on complementary DNAs (cDNAs) rather than genomic DNA, because it would be a fast course for the identification of novel genes (Franco et al. 2000). ESTs are short single-pass sequences of about 300 bp obtained from cDNAs (Adams et al. 1991). From the initial effort of generating just over 400 clones (Franco et al. 1994), there are 139 135 *S. mansoni* sequences available today in dbEST (Table 2). Recent reviews of the generated data are available (Franco et al. 2000; Oliveira 2001; Merrick et al. 2003). It is interesting to point out, however, that a large number of novel genes have been identified (i.e., over 80% of the unique genes are novel) and that the libraries still show a low level of redundancy, indicating that many more ESTs can be generated by the available libraries. Certainly the use of new cDNA libraries will enhance gene discovery. Clusterization and further analysis of these data will be discussed below.

A different strategy was applied by Zouain et al. (1998) to sequence clones that were recognized by antibodies in infected human sera, generating an antigenic sequence tag. Although the strategy results in a small number of clones being selected for sequencing, it adds value to the DNA sequence information because the clones code for known immunogens.

Despite the relatively small number of available ESTs for most species of the genus *Schistosoma*, the entire research community is very excited about recent advances that have generated a large number of ESTs and open-reading frame ESTs (ORESTES) of *Schistosoma japonicum* and *S. mansoni* (Verjovski-Almeida et al. 2003; Hu et al. 2003). *Schistosoma haematobium* cDNA libraries were used for the production of ESTs in a high throughput system in a collaboration between The Sanger Institute, The Natural History Museum at London, and The University of Leeds (Table 1; D. Johnston, personal communication). The Sanger Institute has, in collaboration with the University of York group, generated over 10 000 ESTs of a cDNA library of a *S. mansoni* lung stage (Table 1). A large sequencing project of *S. japonicum* has also been recently published by a Chinese consortium (Table 2). Over 43 000 ESTs were generated from the adult worm and egg stages and were grouped into 13 131 clusters, 75% of which were previously unknown (Hu et al. 2003). In Brazil, two large-scale *S. mansoni* transcriptome projects were funded. Several laboratories in the State of Minas Gerais have been equipped to sequence at least 100 000 ESTs of various developmental stages of *S. mansoni* (Table 1). The second project involving several laboratories in the State of São Paulo has been recently published (Verjovski-Almeida et al. 2003). The São Paulo group produced 163 000 sequences using both the conventional EST

Table 1. Useful links to information on the genome, sequences, and bioinformatic resources for schistosomes.

Description	Name	Resource	Accessed date
Schistosomiasis	World Health Organization/Special Programme for Research and Training in Tropical Diseases (WHO/TDR)	www.who.int/tdr/dw/schisto2004.htm	27 February 2004
Codon usage of *Schistosoma* species	Kazusa DNA Research Institute	www.kazusa.or.jp/codon/	27 February 2004
Microarray technology	Gene-Chips	www.gene-chips.com/	27 February 2004
Proteome technology	Proteome Sciences	www.proteome.co.uk/proteomics/	27 February 2004
Bacterial artificial chromosome (BAC) library	Children's Hospital Oakland Research Institute	www.bacpac.chori.org/schis103.htm	27 February 2004
Microsatellite, interproscan, and gene nomenclature of *S. mansoni*	Centro de Pesquisas René Rachou – Fundação Oswaldo Cruz (CpqRR–FIOCRUZ)	www.cpqrr.fiocruz.br/dna/	27 February 2004
Expressed sequence tag (EST) sequencing at Minas Gerais	The Minas Gerais Genome Network	bioinfo.cenapad.ufmg.br	27 February 2004
Schistosoma network information	The *Schistosoma* Genome Network	www.nhm.ac.uk/hosted_sites/schisto	27 February 2004
TIGR *S. mansoni* genome project	The Institute for Genomic Research (TIGR)	www.tigr.org/tdb/e2k1/sma1/intro.shtml	27 February 2004
TIGR *S. mansoni* gene index (SmGI)	TIGR	www.tigr.org/tdb/tgi/smgi/	27 February 2004
Schistosoma mansoni chromosome sequencing at TIGR	TIGR	www.tigr.org/tdb/e2k1/sma1/genome.shtml	27 February 2004
Parasite genome databases and genome research resources	European Bioinformatics Institute	www.ebi.ac.uk/parasites/parasite-genome.html	27 February 2004
FTP site of trematode databases	Sanger Institute	ftp://ftp.sanger.ac.uk/pub/databases/Trematode/	27 February 2004
Sanger *S. mansoni* genome project	Sanger Institute	www.sanger.ac.uk/Projects/S_mansoni/	27 February 2004
Sanger *S. mansoni* microarray	Sanger Institute	www.sanger.ac.uk/PostGenomics/PathogenArrays/Schisto/	27 February 2004
Schistosoma mansoni transcriptome	Universidade Federal de Minas Gerais (UFMG)	www.icb.ufmg.br/~lgb/schisto/	27 February 2004
Schmidtea mediterranea EST database	The University of Utah	planaria.neuro.utah.edu/	27 February 2004
Schistosoma mansoni EST genome project	Universidade de São Paulo	verjo18.iq.usp.br/schisto/	27 February 2004

Table 2. Number of sequences of *Schistosoma mansoni* and *Schistosoma japonicum* in the dbEST, Genomic Sequence Survey (dbGSS), High Throughput Genomic (HTG), and Genomes divisions of GenBank.

	Molecular database at National Center for Biotechnology Information			
	dbEST[a]	dbGSS[b]	HTG[c]	Genomes[d]
S. mansoni	139 135	42 017	4	1
S. japonicum	45 902	14	0	1

[a]Division that holds the EST sequences in GenBank.
[b]Division similar to the EST division, except that its sequences are genomic in origin.
[c]The HTG sequence division contains "unfinished" genomic sequence data. The *S. mansoni* sequences represent unfinished BAC sequences being sequenced at TIGR.
[d]Division that contains a variety of genomes, complete chromosomes, contiged sequence maps, and integrated genetic and physical maps. The sequences present in this division represent the mitochondrial genomes.

and ORESTES methods, which were grouped into 31 000 clusters (Tables 1 and 2). The ORESTES method is based on polymerase chain reaction (PCR) amplification of cDNAs using random primers (Dias Neto et al. 2000). Six different life-cycle stages were studied: adult worms, eggs, miracidia, germ balls, schistosomula, and cercariae. The large-scale sequencing projects provide insights into the physiology of the species by automated gene ontology assignments and point to a large number of genes that may be of interest to the understanding of host–parasite interactions and vaccine and drug developments.

Genomic sequencing

Efforts towards the genomic sequencing of *S. mansoni* are under way. Bacterial artificial chromosome (BAC) and yeast artificial chromosome (YAC) libraries, which were developed for genomic sequencing and other purposes, are available as biological resources (Tanaka et al. 1995; Le Paslier et al. 2000). The BAC libraries contain inserts shorter than the YAC library, average size of 100–140 and 358 kb, respectively. The sequencing of the 5′ and 3′ ends of the BAC library inserts (i.e., BAC end sequencing) should enhance gene discovery and be useful markers in creating a physical map using a map-as-you-go strategy (Venter et al. 1996). A group from France has generated a large number of BAC end sequences that were deposited in GenBank, dbEST, using the library described by Le Paslier et al. (2000) (Table 2). Another project being conducted at The Institute for Genome Research (TIGR) aims at generating 13 Mb of discontinuous single-pass sequence by sequencing BAC ends (www.tigr.org/tdb/e2k1/sma1/genome.shtml; Table 1). So far, almost 32 000 BAC end sequences have been generated from two libraries, Sm1 (Le Paslier et al. 2000) and CHORI-103 (www.bacpac.chori.org/schis103.htm; Table 1), which are roughly 69% of the total planned. Another aim of the same project is to generate 8 Mb of complete sequences. At present, four entire inserts of BAC clones are being sequenced at TIGR and the sequences can be found in the High Throughput Genomic (HTG) sequence division of GenBank (Table 2). By blast searching against the TIGR clustered ESTs database in addition to many reverse transcriptase, repeat elements, *gag* and *pol* transcripts, the BACs contained the genes for *Smad2* (Osman et al. 2001), retinoic acid receptor RXR (Freebern et al. 1999), and superoxide dismutase (Hong et al. 1992). Both TIGR and Sanger Institute have initiated large-scale sequencing of the *S. mansoni* genome (Table 1). Jointly, these groups have generated 7.5 times the coverage of the genome (P.T. LoVerde, personal communication).

Repeated and mobile elements

The ribosomal RNA (rRNA) gene complex consists of units that are repeated approximately 100 times per haploid genome (Simpson et al. 1984). Each unit, approximately 10 kb, codes for three highly conserved large and small rRNA subunits separated by less conserved, non-transcribed regions called spacers (van Keulen et al. 1985; Walker et al. 1989*b*). Variations in the spacer regions and ribosomal subunit sequences accessed by fragment length polymorphism or sequencing have been used to detect differences between *Schistosoma* species (Lockyer et al. 2003).

Highly polymorphic repeated elements were present in many *S. mansoni* mRNA transcripts and appeared to be similar to hypervariable regions in the human genome (Spotila et al. 1991). A cDNA clone (SM750), containing a peculiar repetitive sequence consisting of 62 bp units called the polymorphic repeated element, was found in small tandem arrays in many mRNA transcripts of diverse size and sequences. An oligonucleotide probe directed against this 62-pb repeat recognized multiple fragments in genomic restriction fragments, suggesting that it was widely distributed in the genome (Spotila et al. 1991). Curiously, a minisatellite region in the mitochondrial DNA (mtDNA) of *S. mansoni* consisting of a 558-bp DNA segment followed by a 62-bp repeat was found that was highly similar to SM750 (Pena et al. 1995). This minisatellite was possibly transferred from the nuclear genome (Pena et al. 1995).

Several other repetitive elements have been described. Two repetitive DNA sequences have been characterized from *S. mansoni* that were transcribed and translated into cross-reactive proteins (Smith et al. 1992). The first element was arranged in a tandem array of at least 17 copies, the other was dispersed in the genome. A tandemly repeated genomic sequence, Sjrh1.0, was found in *S. japonicum* that demonstrated inter- and intra-strain diversity (Drew et al. 1998). Another highly repeated element of 121 bp, which represents about 12% of the total DNA, has been used as a target to detect the parasites in snails (Hamburger et al. 1991) and in human serum and fecal samples (Pontes et al. 2002).

Mobile elements are also highly repeated in the genome of schistosomes. There are two main categories of mobile genetic elements: class I, which transposes through a RNA

Table 3. Repetitive and mobile elements deposited in the European Molecular Biology Laboratory (EMBL) database.

Description	EMBL accession No.	Reference
Schistosoma mansoni SM750 gene	M63265	Spotila et al. 1991
Schistosoma mansoni satellite DNA repeat	M19148	Spotila et al. 1987
Mitochondrion *S. mansoni* minisatellite DNA region	L27240	Pena et al. 1995
Schistosoma mansoni retroposon-like repetitive DNA element DNA (Smα)	M27676–M27681	Spotila et al. 1989
Schistosoma japonicum Sj-alpha-1 retroposon-like sequence	AF213692	Laha et al. 2000
Schistosoma mansoni SR1 non-LTR retrotransposon	U66331–U66339	Drew and Brindley 1997
Schistosoma japonicum non-LTR retrotransposon Sj-pido, partial sequence	AY034003	Laha et al. 2002*b*
Schistosoma mansoni SR2 subfamily A retrotransposon	AF025676	Drew et al. 1999
Schistosoma japonicum retrotransposon SjR2 polyprotein gene, complete cds	AY027869	Laha et al. 2002*a*
Schistosoma japonicum retrotransposon Gulliver, complete sequence	AF243513	Laha et al. 2001
Schistosoma mansoni tandem repeat units	M61098	Hamburger et al. 1991

Table 4. Female-specific DNA segments.

Description	EMBL accession No.	Reference
Schistosoma mansoni female-specific DNA repeat W1	J04665	Webster et al. 1989
Schistosoma mansoni female-specific retroposon SMAlphafem-1, repeat DE region	U12442	Drew and Brindley 1995
Schistosoma mansoni female-specific repetitive DNA, repeat W2	U10109	Drew and Brindley 1995
Schistosoma mansoni satellite DNA repeat (D9)	M19148	Spotila et al. 1987

intermediate, and class II, which transposes directly. Class I elements include long terminal repeat (LTR) retrotransposons and retroviruses, and non-LTR retrotransposons (LINEs) and short interspersed nuclear elements (SINEs). Class II elements are termed transposons. The mobile genetic elements that colonize parasites have been reviewed by Brindley et al. (2003). Table 3 summarizes the elements that have been found in schistosomes.

One of the first mobile elements to be described is a family of short repetitive DNA called Smα that is repeated 7 000 – 10 000 times per haploid genome (Spotila et al. 1987) and that has also been observed in transcribed sequences (Abath et al. 1993; Ferbeyre et al. 1998). Smα belongs to a SINE-like family of mobile elements (Spotila et al. 1989). Similar elements have also been observed in *S. japonicum* (Laha et al. 2000). These elements were probably acquired by vertical transmission from a common ancestor (Laha et al. 2000). Although Smα and Smα-like sequences have been localized in the W chromosome by in situ hybridization (Spotila et al. 1989; Hirai et al. 1993), Smα has also been mapped to many other regions of the genome and therefore is not only present in the female genome (Hirai et al. 1989).

The Smα, Sjα, and similar repeats from other species of the genus *Schistosoma* were found to contain a hammerhead catalytic domain by in silico search (Ferbeyre et al. 1998; Laha et al. 2000). The hammerhead ribozyme is a naturally catalytic RNA structure (Prody et al. 1986). One possible trans-cleavage site for the ribozyme is the gene coding for a synaptobrevin-like protein, which was cleaved in in vitro experiments (Ferbeyre et al. 1998), and trans-cleavage of a natural RNA target has been demonstrated in vivo (Vazquez-Tello et al. 2002). The presence of the hammerhead ribozyme in all species of the genus *Schistosoma* studied so far point to a vertical mode of transmission and they may be involved in trans-cleavage of mRNA targets (Ferbeyre et al. 1998; Laha et al. 2000).

SR1, a non-LTR retrotransposon from *S. mansoni*, has been described and is present in over 200 copies per genome (Drew and Brindley 1997). Pido, another non-LTR retrotransposon, was described from *S. japonicum* where it was detected in about 1000 copies. Although related, pido is not orthologous to SR1 (Laha et al. 2002*b*).

Another lineage of widely distributed non-LTR retrotransposons, termed as RTE-1-like element, was described in *S. mansoni* (Malick and Eickbush 1998). The fully characterized retrotransposon was named SR2 and short SINE-like SR2 elements were also described (Drew et al. 1999). *Schistosoma japonicum* was also found to contain RTE-like elements with approximately 10 000 copies per genome. As described for pido, SjR2 is homologous but not orthologous to SR2 (Laha et al. 2002*a*).

Gulliver, a LTR retrotransposon, was described in *S. japonicum* (Laha et al. 2001). It was detected in between 100 and several 1000 copies in the genome and may also be found in *S. mansoni*.

Female-specific sequences

By representational difference analysis, two female-specific repetitive elements were identified (Drew and Brindley 1995). One of them displayed a high level of identity to the Smα family of SINE elements (see below). The other element, termed W2, was a tandemly repeated segment in the W chromosome. Another fragment, termed W1, was previously identified as having over 500 copies per female genome (Webster et al. 1989). D9, another member of the Smα family of SINE-like sequences, occurred as a female-specific 38-kb cluster of tandem repeats (Spotila et al. 1987). The localization of Smα or Smα-like sequences and the W1 segment in the W chromosome of *S. mansoni* have been demonstrated by in situ hybridization (Spotila et al. 1989; Hirai et al. 1993). However, Smα also maps to other chromosomes (Hirai et al. 1989) and the presence of W1 and

W2, but not D9, in the male genome has been postulated, depending on the strain studied (Grevelding 1995; Quack et al. 1998). A list of female-specific DNA segments can be found in Table 4.

Repeated *S. mansoni* DNA sequences have also proved useful as probes in hybridization assays for sex determination of the parasites (Spotila et al. 1987; Webster et al. 1989; Walker et al. 1989*a*), and methods for sexing single larval stages of *S. mansoni* by PCR are available (Gasser et al. 1991; Dias et al. 1993).

Genomic mapping

The complete generation of a physical map by identifying overlapping clones is no longer essential for the complete sequencing of a genome by the whole genome shotgun approach (Holt et al. 2002). However, the complete physical map is still highly desirable because it would aid in assembling the sequenced genome by any approach and also in the identification of genes of interest (Hearne et al. 1992; McPherson et al. 2001). The initial efforts towards a physical map of *S. mansoni* started with the construction of a YAC library (Tanaka et al. 1995). The initial strategy identified YAC clones containing the SMα repeat (Spotila et al. 1989; McPherson et al. 2001). The selected clones were mapped by FISH assay and all of the seven pairs and the sexual chromosomes were shown to contain at least two copies of the repeat. A map of the observed hybridization pattern can be seen at the *Schistosoma* Genome Network Web site (Table 1).

More recently, a BAC library has been constructed (Le Paslier et al. 2000). A small number of the BAC clones were used for FISH analysis and were mapped to chromosomes Z, W, 1, and 2. BAC libraries will provide a valuable resource for the generation of a complete map. Some researchers, however, are also using this resource to map clones of interest, which will provide additional sequence-tagged sites that are short genomic landmark sequences useful for physical mapping work (Olson et al. 1989; Mei et al. 1995; Foulk et al. 2002).

Microsatellites

Microsatellites are small tandemly repeated sequences that are 1–6 bp long and are widely dispersed in eukaryotic genomes. They are diploid, inherited in a Mendelian fashion, expressed co-dominantly, and are selectively neutral (Tautz and Renz 1984; Weber and May 1989; Ashley and Dow 1994). Microsatellites display high polymorphism owing to the variation in the number of repeated units (Schlotterer 2000). The polymorphism and wide genomic distribution make microsatellites one of the most useful genetic markers available for typing of individuals (Hagelberg et al. 1992) or for population studies (Jarne and Lagoda 1996) and for the construction of genetic maps to identify loci involved in genetic diseases (Dietrich et al. 1996). Microsatellites can be easily scored with the use of PCR with well-defined alleles (Ashley and Dow 1994). As microsatellites are thought to occur in all eukaryotic organisms, it is reasonable to expect that they could play an important role in the study of parasite population genetics (Curtis and Minchella 2000; Barker 2002).

To develop variable microsatellite markers, it is necessary to perform library screens or to make use of DNA sequences available in public databases. The cost and labor involved in constructing and screening libraries in addition to the lack of a large number of deposited sequences have hindered the development of these markers for a number of parasitic organisms. For schistosomes, however, the existence of genomic or cDNA sequencing projects has provided a rich source of sequences to be screened for microsatellites (Franco et al. 2000; Oliveira 2001; Oliveira and Johnston 2001).

The first microsatellite markers were developed by Durand et al. (2000). The authors described 33 microsatellite markers from *S. mansoni* through library screening and database searching. These markers detected considerable variability in parasites obtained from wild rats in Guadeloupe. In addition, a sex-specific genetic structure of the population was unexpectedly observed (Prugnolle et al. 2002). Blair et al. (2001) developed another 20 markers for *S. mansoni* and verified that the African population was more diverse than what was reported by Durand et al. (2000) in Guadeloupe. By database searching, another six markers were developed by Rodrigues et al. (2002*a*). Using these loci, Rodrigues et al. (2002*a*) observed that Brazilian field populations were more diverse than a laboratory strain, but not significantly distinct from each other, which pointed to the occurrence of a high gene flow in the field. A complete search of all of the NCBI databases in addition to sequencing microsatellite-enriched genomic libraries extended the number of loci identified (Rodrigues et al. 2002*b*). The complete set of results obtained by the latter group is available at the Web site listed in Table 1. The high AT content (65%) reported for the *S. mansoni* genome was reflected in the types of repeats identified (Simpson et al. 1982; Marx et al. 2000).

Although the number of polymorphic microsatellite loci described is still small, the studies mentioned above conclusively demonstrate the usefulness of polymorphic microsatellite markers for genetic and population studies of schistosomes. The use of these markers will help in the understanding of the population dynamics in endemic areas.

The mitochondrial genome

The mitochondrial genomes of *S. mansoni*, *S. japonicum*, and *Schistosoma mekongi* (GenBank entries NC_002545, NC_002544, and NC_002529, respectively) have been completely sequenced along with the partial sequence of *S. haematobium* (Le et al. 2000). The mitochondrial genome was shown to contain 36 genes: 12 protein-encoding genes (*cox1*, *nad1*, *nad4L*, *atp6*, and *cob*); 2 rRNAs (*rrn*L is a large subunit rRNA and *rrn*S is a small subunit rRNA); and 22 transfer RNA (tRNA) genes. The *atp8* gene was missing (Le et al. 2001*b*). A map of the completed mitochondrial genomes can be seen by accessing the sequences at the Genomes division of GenBank (www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=Genome). One interesting observation was the presence of a leucine zipper DNA binding motif in the *nad4L* gene (Le et al. 2001*a*). The mitochondrial genomes of parasitic flatworms have been recently reviewed by Le et al. (2002).

These mitochondrial sequences are very useful for phylogenetic analysis by studying the order of genes. In fact, a large difference was observed between African and Asian

schistosomes, particularly between *S. mansoni* and *S. japonicum* where the genes *atp6* and *nad2* were in different positions, the genes for *nad3* and *nad1* had exchanged positions, and the variable non-coding regions were in different places along with several genes for tRNAs (Le et al. 2000). Mitochondrial DNA of *S. mansoni* was shown to be highly polymorphic and varied in size from 16.5 to 24.9 kb in different strains (Despres et al. 1991). The basis of the size variability was because of the presence of a complex minisatellite in the non-coding region (Pena et al. 1995). The high level of variability within the mitochondrial genome, including sequence variations, has also been exploited to study the level of genetic variation in schistosome populations and for phylogenetic analysis (Curtis et al. 2001; Lockyer et al. 2003). The mode of inheritance of the mitochondrial genome has been determined to be maternal, but the high rate of mutation may result in the appearance of new haplotypes (Minchella et al. 1994; Bieberich and Minchella 2001; Jannotti-Passos et al. 2001). The mitochondrial genome has also been used as a target for PCR amplification in an assay to detect *S. mansoni* infected snails (Jannotti-Passos et al. 1997).

Post genomics

Gene-expression analysis

Microarray technology is a revolutionary high throughput tool for the study of gene expression (Shalon et al. 1996). The main advantage of the technique is the ability to simultaneously analyse the transcriptional status of thousands of genes in a single experiment. The basic strategy is to create spots containing a DNA sequence that is representative of the gene of interest. Thousands of spots can be applied on a glass slide by robots. RNA obtained from an experimental sample is used to synthesize cDNA labeled with a chromophore. After hybridization, the amount of hybridizing cDNA, or fluorescence emitted by the specific spot, correlates to the initial amount of message for the specific gene (Hegde et al. 2000). This new technology has had a deep impact on the way biological systems are studied (Eisen et al. 1998; Chu et al. 1998).

The comparison of male versus female gene expression was the objective of the first report on the use of microarrays in schistosomes (Hoffman et al. 2002). In that paper, the authors used a 576 spot array and observed 12 new female-associated transcripts and 4 male-associated transcripts. Among the female-specific transcripts (in addition to others previously published) were ferritin-1, a tyrosinase ortholog, and several unknown genes. Among male-associated transcripts were tropomyosin, actin, and dynein light chain. We should expect more results from the use of microarrays from different laboratories. Hoffman et al. (2002) reported that a new array of 4000 elements was under construction. The Sanger Institute is also preparing microarrays of the lung-stage transcripts (Table 1).

Although progress has been made in the identification of female-specific transcripts, the signaling mechanisms involved in the activation of female-specific genes are just beginning to be studied (LoVerde 2002). The study of genes involved in female worm maturation is extremely important because the females are incapable of reaching sexual maturity in the absence of sexually mature males (Kunz 2001). Therefore, it may be possible to interfere with this physiological function to prevent the completion of the parasite's life cycle.

The information provided by genome sequencing and the pattern of gene expression do not provide a complete profile of the abundance, final structure, or the state of activity of the gene product. Proteomics has emerged to provide a more complete representation of the expression of the genome in combination with genomic sequence data and mRNA expression profile. The methods used are varied (Table 1), but most systems separate the protein extract by two-dimensional electrophoresis and identify the spots of interest by mass spectrometry (Edwards et al. 2000). The use of proteomic technologies for the study of schistosomes is just being initiated. Ashton et al. (2001) reviewed the subject and demonstrated its use for schistosomes. However, the biggest problem for the development of proteomics has been the lack of full-length sequences. The availability of a large number of ORESTES sequences (Verjovski-Almeida et al. 2003), the conventional ESTs being generated by the Minas Gerais Genome Network, and the genomic sequencing by TIGR and the Sanger Institute will overcome this problem.

Bioinformatics

The first type of analysis carried out with the sequence information is usually clusterization. At the *Schistosoma* Genome Network Web site listed in Table 1, information related to the clusterization of *S. mansoni*, *S. japonicum*, and *Biomphalaria glabrata* can be obtained. New efforts into aggregating value to the sequence information available have been carried out by TIGR, the Universidade Federal de Minas Gerais (UFMG), CPqRR–FIOCRUZ, and the Organization for Nucleotide Sequencing and Analysis groups. Researchers from these institutions conducted clusterization of the ESTs followed by functional analysis. Links to these studies are listed in Table 1.

The TIGR study constructed a *S. mansoni* gene index (SmGI) (Merrick et al. 2003). The goal of SmGI is to represent a non-redundant view of all *S. mansoni* genes and data on their expression patterns, cellular roles, functions, and evolutionary relationships. The data are available free of charge and can be downloaded after a licensing agreement is signed. The current release clustered 16 643 sequences in 11 072 tentative consensus sequences (TCs) and 347 curated mature transcripts (ETs). Blast searches can be conducted against all of the sequences in a species-specific manner and the database can be searched for specific sequences (TCs, ETs, and ESTs). Each sequence can be viewed by the EST annotator, which is also searchable. In the EST annotator, there is information about the EST and there are also links to a tentative ortholog group. Similarly, the TC annotator condensates the information on the consensus sequences and adds a gene onthology term, an ET name, and a tentative annotation, if possible. The gene onthology assignment divides the TCs by molecular function (209 TCs), cellular component (240 TCs), or biological process (182 TCs). By molecular function, 143 were enzymes or ligand-binding or carrier; all 240 TCs were in the cellular compartment; and 156 TCs involved in biological processes were related to cell growth and (or) maintenance. Links exist to expand on all of the as-

signed categories and SmGI is also searchable by words present in the annotations. All of the sequences in SmGI have the library and life-cycle stage from which they were obtained identified. This permits another very interesting type of approach to analyzing the data, i.e., the comparison of genes by library and life-cycle stage. This would allow the identification of transcripts that were found uniquely at a certain developmental stage or that were transcribed at different rates at a certain life-cycle stage.

The second study by the UFMG group aims mainly at providing an annotated version of the current *S. mansoni* transcriptome and at identifying orthologous genes and metabolic pathways present in the parasite (Prosdocimi et al. 2002). The group used a different approach from the TIGR work and the results can be viewed in a different format. All of the scripts and files containing sequence information are available for download directly from their Web site. The sequences were divided into four main functional categories: information storage and processing; cellular processes; metabolism; and poorly characterized sequences. By selecting the functional category from the Web site, a list of the putative genes appears; by selecting on the gene, information related to the contig and ESTs in the contig can be obtained. At the latter level, there are also links to the Kegg server with a view of the metabolic pathway. The authors observed that most of the genes in the information storage and processing category were related to translation, ribosomal structure, and biogenesis (9.57% of the total); in the cellular process category most were related to cell motility, secretion (cytoskeleton and intracellular transport), or post-translational modification, protein turnover, and chaperones (7.36% and 6.17%, respectively); in the metabolism category most were related to carbohydrate transport and metabolism or amino acid transport and metabolism (5.79% and 3.71%, respectively); and finally in the poorly characterized sequence category most of the sequences had general function prediction only (24.05%) or were of unknown function (18.13%). As a complement to this study, our group has used the Interproscan system to conduct a search against motif databases. The São Paulo group has provided a searchable database of the ESTs that they produced, including a blast server and another GO browser which was also subdivided into Eukaria- or Metazoa-specific sequences. Links to all of the mentioned information are provided in Table 1.

In addition to the above, there is a Web blast server at the European Bioinformatics Institute (EBI; Table 1). At this server, WU-blast searches can be conducted against sequences from a number of different organisms including the entire *Schistosoma* genus; the unique databases of each of the two species *S. mansoni* and *S. japonicum*; and *S. mansoni* cercaria sequences. The databases used are also available for download.

In addition to the EST bioinformatic resources, TIGR has made available a Web interface that allows blast searching against available BAC ends and the whole genome shotgun sequences. The Sanger Institute site also provides blast searching against the whole genome shotgun sequences. All of the sequences can be obtained through a file transfer protocol (ftp) site (Table 1).

A standardized system for gene and sequence variation nomenclature is not yet available. To fill this gap our group has developed a system that is open for discussion by the research community. The information can be obtained from our Web site listed in Table 1.

One drawback of working with schistosomes is the lack of model organisms against which data can be compared. One such model could be the platyhelminth planarian *Schmidtea mediterranea*. Recently, a relatively small number of about 3000 ESTs were deposited in GenBank (Alvarado et al. 2002), and computational analysis and expression data can be obtained from Alvarado's Web site at the University of Utah (Table 1).

Patents

Gene patenting is treated differently according to local legislation. For example, patenting genes is permitted in the United States but is not permitted in Brazil according to Brazilian patent laws (Eisenberg 1998; Zanotto 2000; Andrews 2002). However, the value of patenting genes, especially ESTs as is the case for schistosomes, has been the source of heated debate with strong arguments for and against (Eisenberg 1998; Arnold and Ogielska-Zei 2002). Until now not many schistosome-derived biotechnologically oriented patents have been granted (Table 5). Most are related to vaccine, diagnostic, immunomodulatory agents and anticoagulant reagents, and a transgene method. The approaches involved in vaccine and diagnostic have been extensively discussed elsewhere (Gryseels 2000; Hagan et al. 2000). It is noteworthy, however, that at least three patents were granted for the use of schistosome gene products as pharmaceutical reagents, namely a factor that induces IL-4 secretion from basophils (European Patent Office (EPO) patent No. WO02085930; Falcone et al. 1996), the anti-inflamatory protein Sm16.8 (Ramaswamy et al. 1995), and the anticoagulant protein HB14 (EPO patent No. 772679). However, none of these products were identified by in silico methods from sequence databases. At least 40% of the genes sequenced have no known homologues from other species that would permit functional inferences (Oliveira and Johnston 2001). Therefore, even if the full genome is sequenced, the identification of genes with interest to the biotechnology industry would still be a difficult task. Nevertheless, the search for genes of interest for the biotechnology industry will be made easier with the entire genome sequence available. The biology of the parasite with its ability to thrive in the circulatory system of the vertebrate host (Combes 1993) and to modulate the immune system (Correa-Oliveira et al. 1998) should provide the industry with a number of possible products of market interest.

A method of interest to the biotechnology industry that may result from genomic information is a transgene method. As described by Brindley et al. (2003), this task may be accomplished with the use of transposons. However, one method for transgenesis has been patented (Table 5). The patent describes methods that interfere with schistosome genes, with employment of genomic promoter regions to drive transcription, and with signal peptides for the secretion of the gene product. Although such a method would be highly desirable, there were no experimental results presented using transgenic worms (EPO patent No. WO9711191). Transient transformation could nevertheless be a very useful tool for studying the regulation of gene expression. The first

Table 5. Patents granted in the United States (United States Patent Office, USPTO) and Europe (European Patent Office, EPO) on genes or proteins of schistosomes and their use.

Patent No.[a]	Title	Inventors	Date of deposit
USPTO 5,730,984 and EPO US5730984	Vaccine against helminth infection comprising Sm-14 fatty acid binding protein of *S. mansoni*	Miriam Tendler, Naftale Katz, Andrew John Simpson	Mar. 1998
USPTO 5,597,570 and EPO US5597570	Protein recognized by antibodies raised against native P28 of *S. mansoni*	Paul Sondermeyer, Jean-Marc Balloul, Raymond Pierce, Jean-Marie Grzych, Marie-Paule Kieny, Gerard Loison, Andre Capron, Jean-Pierre Lecocq	Jan. 1997
USPTO 4,656,033 and EPO US4656033	Isolated, soluble immunogen against *S. mansoni* and a method of vaccination employing same	Stephanie James, Alan Sher	Apr. 1987
USPTO 4,384,992 and EPO US4384992	Novel peptide from cultures of *S. mansoni*, a process for producing it and pharmaceutical compositions containing the same	Andre Capron, Christine Mazingue, Daniel Camus	May 1983
USPTO 4,158,049 and EPO US4158049 and GB1543277	Antigen fraction of *S. mansoni* eggs suitable for testing for schistosomiasis	Ronald P. Pelley, Kenneth S. Warren	June 1979
EPO WO02085930	Immunomodulating agents from parasitic worms and method for isolation thereof (IL-4 inducing principle of *S. mansoni* eggs)	Gabrielle Schramm, Helmut Haas, Franco Falcone, Achim Gronow, Karin Haisch	Oct. 2002
USPTO 6,372,219 and US6372219 and WO9733613	Parasite-derived anti-inflammatory immunomodulatory protein (Sm16.8)	Takeshi Shibuta, Ramaswamy Kalyana Sundaram, Bernard Salafsky	Apr. 2002 and Sept. 1997
WO0175148	Method and kit for the detection of schistosomiasis through polymerase chain reaction (highly repeated fragment)	Luis Andre Pontes,Ana Lucia Teles Rabello, Emmanuel Dias Neto	Jan. 2001
US6261788 and WO0017654	Diagnostic assays for infectious parasitic helminthes and assays for helminth infections (detection of carbohydrates Lex, LDN, and LDNF in the serum)	Anthony Kwame Nyame, Richard D. Cummungs	July 2001 and Mar. 2000
CN1286261	Mimic peptide of *S. japonicum* ovum antigen and its screening and application (phage display peptide)	Yanjun Bai, Min Qian	Mar. 2001
EP0992582	*Schistosoma* recombinant fusion protein vaccine (27/28 kDa cercarial elastase)	Jon Sayers, Michael Doenhoff	Apr. 2000
CA2232514, EP0851936, AU7241196, and WO9711191	Method of transgene expression and secretion in *Schistosoma*	Ira Miller	Mar. 1997
AU2078697 and WO9733610	Dna encoding 28 kda glutathione s-transferase of *S. mansoni* and uses thereof	Stephen G. Kayes	Sept. 1997
NZ260649 and GB2285626	*S. mansoni* protein rSM-14 and its use in vaccines against *Fasciola hepatica*	Andrew John Simpson, Naftale Katz, Miriam Tendler	June 1995
EP0 772679, WO9535370, and AU2745695	*Schistosoma* protein and uses thereof (anti-coagulant or vaccine antigen HB14)	Alison Agnew, Volker Gutsmann	Dec. 1995
US5583011 and WO9705158	Compositions, treatments, and diagnostics for schistosomiasis and related diseases (phosphofructokinase)	Tag E. Mansour	Dec. 1996 and Feb. 1997
US5219566 and US5051254	Immunoprophylactic polypeptides for schistosomiasis (myosin)	Mette Strand	June 1993 and Sept. 1991
FR2689906	Nucleic acid encoding schistosome glutathione peroxidase - and derived peptide, useful in diagnostic reagents and vaccines, also related vectors, transformed cells, etc.	Andre Capron, David Williams, Raymond Pierce	Oct. 1993

Table 5 (*concluded*).

Patent No.[a]	Title	Inventors	Date of deposit
FR2657883	Method for the purification of GST P26	Transgene SA	Aug. 1991
FR2656626 and WO9109621	Peptide fragment obtained from the 28 kDa protein of *S. mansoni*, and vaccinating and (or) therapeutic compositions comprising said peptide fragment	Helene Gras-Masse, Andre Capron, Andre Tartar, Claude Auriault, Isabelle Woloxczuk	July 1991
FR2447194	Peptide substance with a molecular weight of between 500 and 1000, from cultures of *S. mansoni*	Immunologie Biolog Para Centre	Aug. 1980
WO9002563	Antigens of *S. mansoni* (surface membrane antigens)	Andrew John George Simpson, Sidney Ronald Smithers	Mar. 1990
EP0554064	Vaccine against schistosomiasis (SmIMP25)	Alexander Markovics, Daniela Ram, Israel Schechter	Aug. 1993
AU2064592	A synthetic *S. japonicum* antigen and used thereof (Sj23-like)	Kathleen Davern, Graham Mitchell, Mark Wright	July 1992
JP1139533	Immunogen fraction having activity to *S. haematobium* and its production and immunological agent containing said fraction (glycoprotein extracted from hemocyanin of a mollusk)	Jean-Marie Grzych, Andre Capron, Colette Dissous, Jean Montreuil, Genevieve Spik	June 1991

[a]Search in the USPTO was carried out using www.uspto.gov/main/sitesearch.htm and the search criteria TTL/schistosoma or ABST/schistosoma; search in the EPO was carried out using ep.espacenet.com and the keyword "schistosoma". Comments from the authors are in parentheses.

description of transient expression by particle bombardment was published by Davis et al. (1999). The construct consisted of the spliced leader RNA gene fused to the 5′ end of the luciferase ORF followed by the polyadenylation signal of the enolase gene. A 20-fold above background increase in luciferase activity was observed. Wippersteg et al. (2002) reported on the use of the *S. mansoni hsp70* gene promoter and terminator fused to the green fluorescent protein (*GFP*) reporter gene. The gene construct was shot into adult worms and sporocysts by particle bombardment and transcription was induced in both stages after heat shock. These results are very encouraging and will permit progress before transgenesis becomes a reality for schistosomes. Recently two other reports of transient expression of artificial genes in *S. mansoni* were published (Heyers et al. 2003; Wippersteg et al. 2003). Another technology, RNA interference, has been used in *S. mansoni* and may prove extremelly valuable for the understanding of gene function (Boyle et al. 2003; Skelly et al. 2003).

Concluding remarks and perspectives

The next few years should be very exciting for genomic research. The human genome and the genome of the most important experimental host (the mouse) have been published (Venter et al. 2001; Waterston et al. 2002). The study of the interplay of the genomes of the host and the parasite and its consequence should provide new insights into virulence and pathogenicity of schistosomiasis (Hirst and Stapley 2000).

New genomic sequence information is currently being generated (see above). It is important to note the relevant contribution of developing countries. This participation in data production of, especially ESTs and ORESTES, and use of the information are necessary for the establishment of modern technologies and the use of the information in public databases (Pena 1996; Maudlin and Welburn 2001). Specifically in Brazil, local researchers have been involved from the initial efforts. Although financed by the WHO/TDR, the EST sequencing project was coordinated in Brazil (Franco et al. 2000). Since then, a number of other projects (*Xylella fastidiosa* (Simpson et al. 2000), *Chromobacterium violaceum* (Brazilian National Genome Consortium 2003), the Human Cancer Genome Project (Dias Neto et al. 2000), including several regional genome projects involving parasitic organisms such as *S. mansoni*, *Leishmania chagasi*, and *Trypanosoma cruzi*, have been initiated with funding from State and Federal Brazilian agencies (aeg.lbi.ic.unicamp.br/xf/; bioinfo.cenapad. ufmg.br; www.brgene.lncc.br/; www.lbm.fmvz.usp.br/projetos/cancer/). All of the above projects can receive a significant contribution from developing countries with appropriate funding. Training in bioinformatics has been to some extent conducted by the WHO/TDR, which has been aimed specifically at developing country students and scientists (www.who.int/tdr/grants/workplans/pathogen.htm). However, it is important to maintain the participation of these countries and include others that are not yet involved. To incorporate technology-excluded countries and the engagement of companies, research establishments, the scientific community, and international agencies is pivotal (Maudlin and Welburn 2001). Particularly important is the establishment of the rapidly progressing bioinformatics field, which will be essential in understanding and handling of genomic information (Vukmirovic and Tilghman 2000; Oliveira and Johnston 2001). The debate should also include intellectual property rights. The patenting of genomic information by rich nations may hinder the development of new drugs and vaccines for the diseases that are important for developing nations (Andrews 2002; Trouiller et al. 2002).

One of the challenges of genomics and post genomics is functional analysis of the unknown transcripts of schistosomes, which comprise over 40% of what has been sequenced (Franco et al. 2000). Unfortunately, unlike filarids, schistosomes do not have a better studied model that we can compare the genomic information. It will also be possible to develop new methods for identifying vaccine candidates. The cellular compartment of the coded polypeptides can be identified by computer methods such as signal peptide prediction (Nielsen et al. 1997; Smyth et al. 2003). Methods for antigenicity prediction have also been developed and are constantly being updated (Doytchinova and Flower 2002; Gendel 2002). It is feasible that with new approaches new vaccine candidates will emerge, and after being thoroughly tested in vitro and with animal models, will reach human trials. Approaches that use molecular modeling for the design of new drugs should also receive more attention with the increasing number of sequences being available (McKerrow 1996; Ooms 2000). Promising new methodologies for transgenesis utilizing transposons or transient transformations will provide new approaches for the study of the biology of the parasite and the snail host (Wippersteg et al. 2002; Brindley et al. 2003).

Acknowledgements

G.C.O. received financial assistance from the United Nations Development Program/World Bank/WHO/TDR grants A00552 and A20399, Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) grants 521108 and 40.1979/2003-2, and Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG) grants 17001/01, 407/02, and 2409/03. N.B.R. is a fellow of FAPEMIG and D.B. is a fellow of CNPq.

References

Abath, F.G., Hagan, P., Jeffs, S.A., Schechter, I., Meadows, H.M., Holder, A.A., and Simpson, A.J. 1993. Structure of the gene encoding a putative Schistosoma mansoni tegumental antigen precursor. Mol. Biochem. Parasitol. **60**: 81–91.

Adams, M.D., Kelley, J.M., Gocayne, J.D., Dubnick, M., Polymeropoulos, M.H., Xiao, H., Merril, C.R., Wu, A., Olde, B., and Moreno, R.F. 1991. Complementary DNA sequencing: expressed sequence tags and human genome project. Science (Wash., D.C.), **252**: 1651–1656.

Alvarado, A.S., Newmark, P.A., Robb, S.M., and Juste, R. 2002. The *Schmidtea mediterranea* database as a molecular resource for studying platyhelminthes, stem cells and regeneration. Development (Camb.), **129**: 5659–5665.

Andrews, L.B. 2002. Genes and patent policy: rethinking intellectual property rights. Nat. Rev. Genet. **3**: 803–808.

Arnold, B.E., and Ogielska-Zei, E. 2002. Patenting genes and genetic research: good or bad for innovation? Annu. Rev. Genomics Hum. Genet. **3**: 415–432.

Ashley, M.V., and Dow, B.D. 1994. The use of microsatellite analysis in population biology: background, methods and potential applications. Mol. Ecol. Evolut. **69**: 185–201.

Ashton, P.D., Curwen, R.S., and Wilson, R.A. 2001. Linking proteome and genome: how to identify parasite proteins. Trends Parasitol. **17**: 198–202.

Barker, G.C. 2002. Microsatellite DNA: a tool for population genetic analysis. Trans. R. Soc. Trop. Med. Hyg. **96**(Suppl. 1): 21–24.

Bergquist, N.R. 1995. Schistosomiasis vaccine development: approaches and prospects. Mem. Inst. Oswaldo Cruz, **90**: 221–227.

Bieberich, A.A., and Minchella, D.J. 2001. Mitochondrial inheritance in *Schistosoma mansoni*: mitochondrial variable number tandem repeat mutation produces noise on top of the signal. J. Parasitol. **87**: 1011–1015.

Blair, L., Webster, J.P., and Barker, G.C. 2001. Isolation and characterization of polymorphic microsatellite markers in *Schistosoma mansoni* from Africa. Mol. Ecol. Notes, **1**: 93–95.

Boyle, J.P., Wu, X.J., Shoemaker, C.B., and Yoshino, T.P. 2003. Using RNA interference to manipulate endogenous gene expression in *Schistosoma mansoni* sporocysts. Mol. Biochem. Parasitol. **128**: 205–215.

Brazilian National Genome Project Consortium. 2003. The complete genome sequence of *Chromobacterium violaceum* reveals remarkable and exploitable bacterial adaptability. Proc. Natl. Acad. Sci. U.S.A. **100**: 11 660 – 11 665.

Brindley, P.J., Laha, T., McManus, D.P., and Loukas, A. 2003. Mobile genetic elements colonizing the genomes of metazoan parasites. Trends Parasitol. **19**: 79–87.

Capron, A., Capron, M., Dombrowicz, D., and Riveau, G. 2001. Vaccine strategies against schistosomiasis: from concepts to clinical trials. Int. Arch. Allergy Immunol. **124**: 9–15.

Chitsulo, L., Engels, D., Montresor, A., and Savioli, L. 2000. The global status of schistosomiasis and its control. Acta Trop. **77**: 41–51.

Chu, S., DeRisi, J., Eisen, M., Mulholland, J., Botstein, D., Brown, P.O., and Herskowitz, I. 1998. The transcriptional program of sporulation in budding yeast. Science (Wash., D.C.), **282**: 699–705.

Combes, C. 1993. *Schistosoma*, water and man. Rev. Prat. **43**: 405–415.

Correa-Oliveira, R., Malaquias, L.C., Falcão, P.L., Viana, I.R., Bahia-Oliveira, L.M., Silveira, A.M., Fraga, L.A., Prata, A., Coffman, R.L., Lambertucci, J.R., Cunha-Melo, J.R., Martins-Filho, O.A., Wilson, R.A., and Gazzinelli, G. 1998. Cytokines as determinants of resistance and pathology in human *Schistosoma mansoni* infection. Braz. J. Med. Biol. Res. **31**: 171–177.

Curtis, J., and Minchella, D. J. 2000. Schistosome population genetics structure: when clumping worms is not just splitting hairs. Parasitol. Today, **16**: 68–71.

Curtis, J., Fraga, L.A., de Souza, C.P., Correa-Oliveira, R., and Minchella, D.J. 2001. Widespread heteroplasmy in schistosomes makes an mtVNTR marker "nearsighted". J. Hered. **92**: 248–253.

Davis, R.E., Parra, A., LoVerde, P.T., Ribeiro, E., Glorioso, G., and Hodgson, S. 1999. Transient expression of DNA and RNA in parasitic helminths by using particle bombardment. Proc. Natl. Acad. Sci. U.S.A. **96**: 8687–8692.

Despres, L., Imbert-Establet, D., Combes, C., Bonhomme, F., and Monnerot, M. 1991. Isolation and polymorphism in mitochondrial DNA from *Schistosoma mansoni*. Mol. Biochem. Parasitol. **47**: 139–141.

Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W.J., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H., Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J., and Simpson, A.J. 2000. Shotgun sequencing of the human transcriptome with ORF expressed sequence tags. Proc. Natl. Acad. Sci. U.S.A. **97**: 3491–3496.

Dias, N.E., Santos, F.R., Pena, S.D., and Simpson, A.J. 1993. Sex determination by low stringency PCR (LS-PCR). Nucleic Acids Res. **21**: 763–764.

Dietrich, W.F., Miller, J., Steen, R., Merchant, M.A., Damron-Boles, D., Husain, Z., Dredge, R., Daly, M.J., Ingalls, K.A., and O'Connor, T.J. 1996. A comprehensive genetic map of the mouse genome. Nature (Lond.), **380**: 149–152.

Doytchinova, I.A., and Flower, D.R. 2002. Quantitative approaches to computational vaccinology. Immunol. Cell. Biol. **80**: 270–279.

Drew, A.C., and Brindley, P.J. 1995. Female-specific sequences isolated from *Schistosoma mansoni* by representational difference analysis. Mol. Biochem. Parasitol. **71**: 173–181.

Drew, A.C., and Brindley, P.J. 1997. A retrotransposon of the non-long terminal repeat class from the human blood fluke *Schistosoma mansoni*. Similarities to the chicken-repeat-1-like elements of vertebrates. Mol. Biol. Evol. **14**: 602–610.

Drew, A.C., Brindley, P.J., Lewis, F.A., Liang, Y.S., and Minchella, D.J. 1998. Tandemly repeated genomic sequence demonstrates inter- and intra-strain genetic variation in *Schistosoma japonicum*. Trop. Med. Int. Health, **3**: 373–380.

Drew, A.C., Minchella, D.J., King, L.T., Rollinson, D., and Brindley, P.J. 1999. SR2 elements, non-long terminal repeat retrotransposons of the RTE-1 lineage from the human blood fluke *Schistosoma mansoni*. Mol. Biol. Evol. **16**: 1256–1269.

Durand, P., Sire, C., and Théron, A. 2000. Isolation of microsatellite markers in the digenetic trematode *Schistosoma mansoni* from Guadeloupe island. Mol. Ecol. **9**: 997–998.

Edwards, A.M., Arrowsmith, C.H., and des Pallieres, B. 2000. Proteomics: new tools for a new era. Mod. Drug Discov. **5**: 41–44.

Eisen, M.B., Spellman, P.T., Brown, P.O., and Botstein, D. 1998. Cluster analysis and display of genome-wide expression patterns. Proc. Natl. Acad. Sci. U.S.A. **95**: 14 863 – 14 868.

Eisenberg, R.S. 1998. Do EST patents matter? Trends Genet. **14**: 379–381.

Falcone, F.H., Dahinden, C.A., Gibbs, B.F., Noll, T., Amon, U., Hebestreit, H., Abrahamsen, O., Klaucke, J., Schlaak, M., and Haas, H. 1996. Human basophils release interleukin-4 after stimulation with *Schistosoma mansoni* egg antigen. Eur. J. Immunol. **26**: 1147–1155.

Fantappie, M.R., Gimba, E.R., and Rumjanek, F.D. 2001. Lack of DNA methylation in *Schistosoma mansoni*. Exp. Parasitol. **98**: 162–166.

Ferbeyre, G., Smith, J.M., and Cedergren, R. 1998. Schistosome satellite DNA encodes active hammerhead ribozymes. Mol. Cell. Biol. **18**: 3880–3888.

Foulk, B.W., Pappas, G., Hirai, Y., Hirai, H., and Williams, D.L. 2002. Adenylosuccinate lyase of *Schistosoma mansoni*: gene structure, mRNA expression, and analysis of the predicted peptide structure of a potential chemotherapeutic target. Int. J. Parasitol. **32**: 1487.

Franco, G.R., Adams, M.D., Bento Soares, M., Simpson, A.J.G., Venter, J.C., and Pena, S.D.J. 1994. Identification of new *Schistosoma mansoni* genes by the EST strategy using a directional cDNA library. Gene (Amst.), **152**: 141–147.

Franco, G.R., Valadão, A.F., Azevedo, V., and Rabelo, E.M. 2000. The *Schistosoma* gene discovery program: state of the art. Int. J. Parasitol. **30**: 453–463.

Freebern, W.J., Osman, A., Niles, E.G., Christen, L., and LoVerde, P.T. 1999. Identification of a cDNA encoding a retinoid X receptor homologue from *Schistosoma mansoni*. Evidence for a role in female-specific gene expression. J. Biol. Chem. **274**: 4577–4585.

Gasser, R.B., Morahan, G., and Mitchell, G.F. 1991. Sexing single larval stages of *Schistosoma mansoni* by polymerase chain reaction. Mol. Biochem. Parasitol. **47**: 255–258.

Gendel, S.M. 2002. Sequence analysis for assessing potential allergenicity. Ann. N.Y. Acad. Sci. **964**: 87–98.

Grevelding, C.G. 1995. The female-specific W1 sequence of the Puerto Rican strain of *Schistosoma mansoni* occurs in both genders of a Liberian strain. Mol. Biochem. Parasitol. **71**: 269–272.

Grossman, A.I., McKenzie, R., and Cain, G.D. 1980. Sex heterochromatin in *Schistosoma mansoni*. J. Parasitol. **66**: 368–370.

Gryseels, B. 2000. Schistosomiasis vaccines: a devil's advocate view. Parasitol. Today, **16**: 46–48.

Hagan, P., Doenhoff, M.J., Wilson, R.A., Al Sherbiny, M., and Bergquist, R. 2000. Schistosomiasis vaccines: a response to a devils' advocate's view. Parasitol. Today, **16**: 322–323.

Hagelberg, E., Gray, I.C., and Jeffreys, A.J. 1992. Identification of the skeletal remains of a murder victim by DNA analysis. Nature (Lond.), **352**: 427–429.

Hamburger, J., Turetski, T., Kapeller, I., and Deresiewicz, R. 1991. Highly repeated short DNA sequences in the genome of *Schistosoma mansoni* recognized by a species-specific probe. Mol. Biochem. Parasitol. **44**: 73–80.

Hearne, C.M., Ghosh, S., and Todd, J.A. 1992. Microsattelites for linkage analysis of genetic traits. Science (Wash., D.C.), **8**: 288–294.

Hegde, P., Qi, R., Abernathy, K., Gay, C., Dharap, S., Gaspart, R., Hughes, J.E., Snesrud, E., Lee, N., and Quackenbush, J. 2000. A concise guide to cDNA microarray analysis. Biotechniques, **29**: 548–562.

Heyers, O., Walduck, A.K., Brindley, P.J., Bleiss, W., Lucius, R., Dorbic, T., Wittig, B., and Kalinna, B.H. 2003. *Schistosoma mansoni* miracidia transformed by particle bombardment infect *Biomphalaria glabrata* snails and develop into transgenic sporocysts. Exp. Parasitol. **105**: 174–178.

Hirai, H., and LoVerde, P.T. 1996. Identification of the telomeres on *Schistosoma mansoni* chromosomes by FISH. J. Parasitol. **82**: 511–512.

Hirai, H., Spotila, L.D., and LoVerde, P.T. 1989. *Schistosoma mansoni*: chromosomal localization of DNA repeat elements by *in situ* hybridization using biotinylated DNA probes. Exp. Parasitol. **69**: 175–188.

Hirai, H., Tanaka, M., and LoVerde, P.T. 1993. *Schistosoma mansoni*: chromosomal localization of female-specific genes and a female-specific DNA element. Exp. Parasitol. **76**: 175–181.

Hirst, S.I., and Stapley, L.A. 2000. Parasitology: the dawn of a new millennium. Parasitol. Today, **16**: 1–3.

Hoffman, K.F., Johnston, D.A., and Dunne, D.W. 2002. Identification of *Schistosoma mansoni* gender-associated gene transcripts by cDNA microarray profiling. Genome Biology [online], **3**(8): research0041.1–0041.12. Available from http://genomebiology.com/2002/3/8/research/0041.1 [cited 27 February 2004].

Holt, R.A., Subramanian, G.M., Halpern, A., Sutton, G.G., Charlab, R., Nusskern, D.R., Wincker, P., Clark, A.G., Ribeiro, J.M., Wides, R., Salzberg, S.L., Loftus, B., Yandell, M., Majoros, W.H., Rusch, D.B., Lai, Z., Kraft, C.L., Abril, J.F., Anthouard, V., Arensburger, P., Atkinson, P.W., Baden, H., de Bernardinis, V., Baldwin, D., Benes, V., Biedler, J., Blass, C., Bolanos, R., Boscus, D., Barnstead, M., Cai, S., Center, A., Chatuverdi, K., Christophides, G.K., Chrystal, M.A., Clamp, M., Cravchik, A., Curwen, V., Dana, A., Delcher, A., Dew, I., Evans, C.A., Flanigan, M., Grundschober-Freimoser, A., Friedli, L., Gu, Z., Guan, P., Guigo, R., Hillenmeyer, M.E.,

Hladun, S.L., Hogan, J.R., Hong, Y.S., Hoover, J., Jaillon, O., Ke, Z., Kodira, C., Kokoza, E., Koutsos, A., Letunic, I., Levitsky, A., Liang, Y., Lin, J.J., Lobo, N.F., Lopez, J.R., Malek, J.A., McIntosh, T.C., Meister, S., Miller, J., Mobarry, C., Mongin, E., Murphy, S.D., O'Brochta, D.A., Pfannkoch, C., Qi, R., Regier, M.A., Remington, K., Shao, H., Sharakhova, M.V., Sitter, C.D., Shetty, J., Smith, T.J., Strong, R., Sun, J., Thomasova, D., Ton, L.Q., Topalis, P., Tu, Z., Unger, M.F., Walenz, B., Wang, A., Wang, J., Wang, M., Wang, X., Woodford, K.J., Wortman, J.R., Wu, M., Yao, A., Zdobnov, E.M., Zhang, H., Zhao, Q., Zhao, S., Zhu, S.C., Zhimulev, I., Coluzzi, M., della, T.A., Roth, C.W., Louis, C., Kalush, F., Mural, R.J., Myers, E.W., Adams, M.D., Smith, H.O., Broder, S., Gardner, M.J., Fraser, C.M., Birney, E., Bork, P., Brey, P.T., Venter, J.C., Weissenbach, J., Kafatos, F.C., Collins, F.H., and Hoffman, S.L. 2002. The genome sequence of the malaria mosquito *Anopheles gambiae*. Science (Wash., D.C.), **298**: 129–149.

Hong, Z., LoVerde, P.T., Hammarskjold, M.-L., and Rekosh, D.M. 1992. *Schistosoma mansoni*: cloning of a complementary DNA encoding a cytosolic Cu/Zn superoxide dismutase and high-yield expression of the enzymatically active gene product in *Escherichia coli*. Exp. Parasitol. **75**: 308–322.

Hu, W., Yan, Q., Shen, D.K., Liu, F., Zhu, Z.D., Song, H.D., Xu, X.R., Wang, Z.J., Rong, Y.P., Zeng, L.C., Wu, J., Zhang, X., Wang, J.J., Xu, X.N., Wang, S.Y., Fu, G., Zhang, X.L., Wang, Z.Q., Brindley, P.J., McManus, D.P., Xue, C.L., Feng, Z., Chen, Z., and Han, Z.G. 2003. Evolutionary and biomedical implications of a *Schistosoma japonicum* complementary DNA resource. Nat. Genet. 35(2):139–147.

Jannotti-Passos, L.K., Vidigal, T.H.D.A., Dias Neto, E., Pena, S.D.J., Simpson, A.W.J.G., Dutra, W.O., Souza, C.P., and Carvalho-Parra, J.F. 1997. PCR amplification of the mitochondrial DNA minisatellite region to detect *Schistosoma mansoni* infection in *Biomphalaria glabrata* snails. J. Parasitol. **83**: 395–399.

Jannotti-Passos, L.K., Souza, C.P., Parra, J.C., and Simpson, A.J. 2001. Biparental mitochondrial DNA inheritance in the parasitic trematode *Schistosoma mansoni*. J. Parasitol. **87**: 79–82.

Jarne, P., and Lagoda, J.L. 1996. Microsatellites, from molecules to populations and back. Trends Ecol. Evol. **11**: 424–429.

Kheir, M.M., Baraka, O.Z., el Tom, I.A., Mukhtar, M.M., and Homieda, M.M. 2000. Effects of single-dose praziquantel on morbidity and mortality resulting from intestinal schistosomiasis. East. Mediterr. Health J. **6**: 926–931. Available from http://www.emro.who.int/publications/emhj/0605/09.htm [cited 27 February 2004].

Kloetzel, K. 1990. Reinfection after treatment of schistosomiasis: environment or "predisposition"? Rev. Inst. Med. Trop. Sao Paulo, **32**: 138–146.

Kunz, W. 2001. Schistosome male–female interaction: induction of germ-cell differentiation. Trends Parasitol. **17**: 227–231.

Laha, T., McManus, D.P., Loukas, A., and Brindley, P.J. 2000. Sjalpha elements, short interspersed element-like retroposons bearing a hammerhead ribozyme motif from the genome of the oriental blood fluke *Schistosoma japonicum*. Biochim. Biophys. Acta, **1492**: 477–482.

Laha, T., Loukas, A., Verity, C.K., McManus, D.P., and Brindley, P.J. 2001. Gulliver, a long terminal repeat retrotransposon from the genome of the oriental blood fluke *Schistosoma japonicum*. Gene (Amst.), **264**: 59–68.

Laha, T., Brindley, P.J., Smout, M.J., Verity, C.K., McManus, D.P., and Loukas, A. 2002*a*. Reverse transcriptase activity and untranslated region sharing of a new RTE-like, non-long terminal repeat retrotransposon from the human blood fluke, *Schistosoma japonicum*. Int. J. Parasitol. **32**: 1163–1174.

Laha, T., Brindley, P.J., Verity, C.K., McManus, D.P., and Loukas, A. 2002*b*. Pido, a non-long terminal repeat retrotransposon of the chicken repeat 1 family from the genome of the Oriental blood fluke, *Schistosoma japonicum*. Gene (Amst.), **284**: 149–159.

Le, T.H., Blair, D., Agatsuma, T., Humair, P.F., Campbell, N.J., Iwagami, M., Littlewood, D.T., Peacock, B., Johnston, D.A., Bartley, J., Rollinson, D., Herniou, E.A., Zarlenga, D.S., and McManus, D.P. 2000. Phylogenies inferred from mitochondrial gene orders — a cautionary tale from the parasitic flatworms. Mol. Biol. Evol. **17**: 1123–1125.

Le, T.H., Blair, D., and McManus, D.P. 2001*a*. A leucine zipper protein of mitochondrial origin. Biochim. Biophys. Acta, **1546**: 435–443.

Le, T.H., Humair, P.F., Blair, D., Agatsuma, T., Littlewood, D.T., and McManus, D.P. 2001*b*. Mitochondrial gene content, arrangement and composition compared in African and Asian schistosomes. Mol. Biochem. Parasitol. **117**: 61–71.

Le, T.H., Blair, D., and McManus, D.P. 2002. Mitochondrial genomes of parasitic flatworms. Trends Parasitol. **18**: 206–213.

Le Paslier, M.C., Pierce, R.J., Merlin, F., Hirai, H., Wu, W., Williams, D.L., Johnston, D., LoVerde, P.T., and Le Paslier, D. 2000. Construction and characterization of a *Schistosoma mansoni* bacterial artificial chromosome library. Genomics, **65**: 87–94.

Lockyer, A.E., Olson, P.D., Ostergaard, P., Rollinson, D., Johnston, D.A., Attwood, S.W., Southgate, V.R., Horak, P., Snyder, S.D., Le, T.H., Agatsuma, T., McManus, D.P., Carmichael, A.C., Naem, S., and Littlewood, D.T. 2003. The phylogeny of the Schistosomatidae based on three genes with emphasis on the interrelationships of *Schistosoma* Weinland, 1858. Parasitology, **126**: 203–224.

LoVerde, P.T. 2002. Presidential address. Sex and schistosomes: an interesting biological interplay with control implications. J. Parasitol. **88**: 3–13.

Malik, H.S. and Eickbush, T.H. 1998. The RTE class of non-LTR retrotransposons is widely distributed in animals and is the origin of many SINEs. Mol. Biol. Evol. **15**: 1123–1134.

Marx, K.A., Bizzaro, J.W., Blake, R.D., Tsai, M.-H., and Tao, L.-F. 2000. Experimental DNA melting behavior of the three major *Schistosoma* species. Mol. Biochem. Parasitol. **107**: 303–307.

Maudlin, I., and Welburn, S.C. 2001. Genome, which genome? Trends Parasitol. **17**: 50–52.

McKerrow, J.H. 1996. Parasitic infections — molecular diagnosis and rational drug design. Biologicals, **24**: 207–208.

McPherson, J.D., Marra, M., Hillier, L., Waterston, R.H., Chinwalla, A., Wallis, J., Sekhon, M., Wylie, K., Mardis, E.R., Wilson, R.K., Fulton, R., Kucaba, T.A., Wagner-McPherson, C., Barbazuk, W.B., Gregory, S.G., Humphray, S.J., French, L., Evans, R.S., Bethel, G., Whittaker, A., Holden, J.L., McCann, O.T., Dunham, A., Soderlund, C., Scott, C.E., Bentley, D.R., Schuler, G., Chen, H.C., Jang, W., Green, E.D., Idol, J.R., Maduro, V.V., Montgomery, K.T., Lee, E., Miller, A., Emerling, S., Kucherlapati, Gibbs, R., Scherer, S., Gorrell, J.H., Sodergren, E., Clerc-Blankenburg, K., Tabor, P., Naylor, S., Garcia, D., de Jong, P.J., Catanese, J.J., Nowak, N., Osoegawa, K., Qin, S., Rowen, L., Madan, A., Dors, M., Hood, L., Trask, B., Friedman, C., Massa, H., Cheung, V.G., Kirsch, I.R., Reid, T., Yonescu, R., Weissenbach, J., Bruls, T., Heilig, R., Branscomb, E., Olsen, A., Doggett, N., Cheng, J.F., Hawkins, T., Myers, R.M., Shang, J., Ramirez, L., Schmutz, J., Velasquez, O., Dixon, K., Stone, N.E., Cox, D.R., Haussler, D., Kent, W.J., Furey, T., Rogic, S., Kennedy, S., Jones, S., Rosenthal, A., Wen, G., Schilhabel, M., Gloeckner, G., Nyakatura, G., Siebert, R.,

Schlegelberger, B., Korenberg, J., Chen, X.N., Fujiyama, A., Hattori, M., Toyoda, A., Yada, T., Park, H.S., Sakaki, Y., Shimizu, N., Asakawa, S., Kawasaki, K., Sasaki, T., Shintani, A., Shimizu, A., Shibuya, K., Kudoh, J., Minoshima, S., Ramser, J., Seranski, P., Hoff, C., Poustka, A., Reinhardt, R., and Lehrach, H. 2001. A physical map of the human genome. Nature (Lond.), **409**: 934–941.

Mei, H., Hirai, H., Tanaka, M., Hong, Z., Rekosh, D., and LoVerde, P.T. 1995. *Schistosoma mansoni*: cloning and characterization of a gene encoding cytosolic Cu/Zn superoxide dismutase. Exp. Parasitol. **80**: 250–259.

Merrick, J.M., Osman, A., Tsai, J., Quackenbush, J., LoVerde, P.T., and Lee, N.H. 2003. The *Schistosoma mansoni* gene index: gene discovery and biology by reconstruction and analysis of expressed gene sequences. J. Parasitol. **89**: 261–269.

Meyne, J., Ratliff, R.L., and Moyzis, R.K. 1989. Conservation of the human telomere sequence (TTAGGG)n among vertebrates. Proc. Natl. Acad. Sci. U.S.A. **86**: 7049–7053.

Minchella, D.J., Lewis, F.A., Sollenberger, K.M., and Williams, J.A. 1994. Genetic diversity of *Schistosoma mansoni*: quantifying strain heterogeneity using a polymorphic DNA element. Mol. Biochem. Parasitol. **68**: 307–313.

Moyzis, R.K., Buckingham, J.M., Cram, L.S., Dani, M., Deaven, L.L., Jones, M.D., Meyne, J., Ratliff, R.L., and Wu, J.R. 1988. A highly conserved repetitive DNA sequence, (TTAGGG)n, present at the telomeres of human chromosomes. Proc. Natl. Acad. Sci. U.S.A. **85**: 6622–6626.

Nielsen, H., Engelbrecht, J., Brunak, S., and von Heijne, G. 1997. A neural network method for identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. Int. J. Neural Syst. **8**: 581–599.

Oliveira, G. 2001. *Schistosoma* gene discovery project update. Trends Parasitol. **17**: 108–109.

Oliveira, G., and Johnston, D.A. 2001. Mining the schistosome DNA sequence database. Trends Parasitol. **17**: 501–503.

Olson, M., Hood, L., Cantor, C., and Botstein, D. 1989. A common language for physical mapping of the human genome. Science (Wash., D.C.), **245**: 1434–1435.

Ooms, F. 2000. Molecular modeling and computer aided drug design. Examples of their applications in medicinal chemistry. Curr. Med. Chem. **7**: 141–158.

Osman, A., Niles, E.G., and LoVerde, P.T. 2001. Identification and characterization of a Smad2 homologue from *Schistosoma mansoni*, a transforming growth factor-beta signal transducer. J. Biol. Chem. **276**: 10 072 – 10 082.

Pena, H.B., de Souza, C.P., Simpson, A.J., and Pena, S.D. 1995. Intracellular promiscuity in *Schistosoma mansoni*: nuclear transcribed DNA sequences are part of a mitochondrial minisatellite region. Proc. Natl. Acad. Sci. U.S.A. **92**: 915–919.

Pena, S.D.J. 1996. Third world participation in genome projects. Trends Biotechnol. **14**: 74–77.

Pontes, L.A., Dias-Neto, E., and Rabello, A. 2002. Detection by polymerase chain reaction of *Schistosoma mansoni* DNA in human serum and feces. Am. J. Trop. Med. Hyg. **66**: 157–162.

Prody, G.A., Bakos, J.T., Buzayan, J.M., Schneider, I.R., and Brustoski, K. 1986. Autolytic processing of dimeric plant virus satellite RNA. Science (Wash., D.C.), **231**: 1577–1580.

Prosdocimi, F., Faria-Campos, A.C., Peixoto, F.C., Pena, S.D., Ortega, J.M., and Franco, G.R. 2002. Clustering of *Schistosoma mansoni* mRNA sequences and analysis of the most transcribed genes: implications in metabolism and biology of different developmental stages. Mem. Inst. Oswaldo Cruz, **97**(Suppl. 1): 61–69.

Prugnolle, F., De Meeus, T., Durand, P., Sire, C., and Théron, A. 2002. Sex-specific genetic structure in *Schistosoma mansoni*: evolutionary and epidemiological implications. Mol. Ecol. **11**: 1231–1238.

Quack, T., Doenhoff, M., Kunz, W., and Grevelding, C.G. 1998. *Schistosoma mansoni*: the varying occurrence of repetitive elements in different strains shows sex-specific polymorphisms. Exp. Parasitol. **89**: 222–227.

Ramaswamy, K., Salafsky, B., Potluri, S., He, Y.X., Li, J.W., and Shibuya, T. 1995. Secretion of an anti-inflammatory, immunomodulatory factor by Schistosomulae of *Schistosoma mansoni*. J. Inflamm. **46**: 13–22.

Rodrigues, N.B., Coura Filho, P., de Souza, C.P., Jannoti Passos, L.K., Dias-Neto, E., Romanha, A.J. 2002*a*. Populational structure of *Schistosoma mansoni* assessed by DNA microsatellites. Inter. J. Parasitol. **32**: 843–851.

Rodrigues, N.B., LoVerde, P.T., Romanha, A.J., and Oliveira, G. 2002*b*. Characterization of new *Schistosoma mansoni* microsatellite *loci* in sequences obtained from public DNA databases and microsatellite enriched genomic libraries. Mem. Inst. Oswaldo Cruz, **97**(Suppl. 1): 71–75.

Schlotterer, C. 2000. Evolutionary dynamics of microsatellite DNA. Chromosoma (Berl.), **109**: 365–371.

Shalon, D., Smith, S.J., and Brown, P.O. 1996. A DNA microarray system for analyzing complex DNA samples using two-color fluorescent probe hybridization. Genome Res. **6**: 639–645.

Short, R.B., and Menzel, M.Y. 1979. Somatic chromosomes of *Schistosoma mansoni*. J. Parasitol. **65**: 471–473.

Simpson, A.J., Sher, A., and McCutchan, T.F. 1982. The genome of *Schistosoma mansoni*: isolation of DNA, its size, bases and repetitive sequences. Mol. Biochem. Parasitol. **6**: 125–137.

Simpson, A.J., Dame, J.B., Lewis, F.A., and McCutchan, T.F. 1984. The arrangement of ribosomal RNA genes in *Schistosoma mansoni*. Identification of polymorphic structural variants. Eur. J. Biochem. **139**: 41–45.

Simpson, A.J., Reinach, F.C., Arruda, P., Abreu, F.A., Acencio, M., Alvarenga, R., Alves, L.M., Araya, J.E., Baia, G.S., Baptista, C.S., Barros, M.H., Bonaccorsi, E.D., Bordin, S., Bove, J.M., Briones, M.R., Bueno, M.R., Camargo, A.A., Camargo, L.E., Carraro, D.M., Carrer, H., Colauto, N.B., Colombo, C., Costa, F.F., Costa, M.C., Costa-Neto, C.M., Coutinho, L.L., Cristofani, M., Dias-Neto, E., Docena, C., El Dorry, H., Facincani, A.P., Ferreira, A.J., Ferreira, V.C., Ferro, J.A., Fraga, J.S., Franca, S.C., Franco, M.C., Frohme, M., Furlan, L.R., Garnier, M., Goldman, G.H., Goldman, M.H., Gomes, S.L., Gruber, A., Ho, P.L., Hoheisel, J.D., Junqueira, M.L., Kemper, E.L., Kitajima, J.P., Krieger, J.E., Kuramae, E.E., Laigret, F., Lambais, M.R., Leite, L.C., Lemos, E.G., Lemos, M.V., Lopes, S.A., Lopes, C.R., Machado, J.A., Machado, M.A., Madeira, A.M., Madeira, H.M., Marino, C.L., Marques, M.V., Martins, E.A., Martins, E.M., Matsukuma, A.Y., Menck, C.F., Miracca, E.C., Miyaki, C.Y., Monteriro-Vitorello, C.B., Moon, D.H., Nagai, M.A., Nascimento, A.L., Netto, L.E., Nhani, A., Jr., Nobrega, F.G., Nunes, L.R., Oliveira, M.A., de Oliveira, M.C., de Oliveira, R.C., Palmieri, D.A., Paris, A., Peixoto, B.R., Pereira, G.A., Pereira, H.A., Jr., Pesquero, J.B., Quaggio, R.B., Roberto, P.G., Rodrigues, V., de, M.R.A., de, R.V., Jr., de Sa, R.G., Santelli, R.V., Sawasaki, H.E., da Silva, A.C., da Silva, A.M., da Silva, F.R., da, S.W., Jr., da Silveira, J.F., Silvestri, M.L., Siqueira, W.J., de Souza, A.A., de Souza, A.P., Terenzi, M.F., Truffi, D., Tsai, S.M., Tsuhako, M.H., Vallada, H., Van Sluys, M.A., Verjovski-Almeida, S., Vettore, A.L., Zago, M.A., Zatz, M., Meidanis, J., and Setubal, J.C. 2000. The genome sequence of the plant pathogen *Xylella fastidiosa*. The *Xylella fastidiosa*

Consortium of the Organization for Nucleotide Sequencing and Analysis. Nature (Lond.), **406**: 151–157.

Skelly, P.J., Da'dara, A., and Harn, D.A. 2003. Suppression of cathepsin B expression in *Schistosoma mansoni* by RNA interference. Int. J. Parasitol. **33**: 363–369.

Smith, A.L., Huggins, M.C., and Havercroft, J.C. 1992. *Schistosoma mansoni*: adult males and females differentially express antigens encoded by repetitive genomic DNA. Parasitology, **105**(1): 63–69.

Smyth, D., McManus, D.P., Smout, M.J., Laha, T., Zhang, W., and Loukas, A. 2003. Isolation of cDNAs encoding secreted and transmembrane proteins from *Schistosoma mansoni* by a signal sequence trap method. Infect. Immun. **71**: 2548–2554.

Spotila, L.D., Rekosh, D.M., Boucher, J.M., and LoVerde, P.T. 1987. A cloned DNA probe identifies the sex of *Schistosoma mansoni* cercariae. Mol. Biochem. Parasitol. **26**: 17–20.

Spotila, L.D., Hirai, H., Rekosh, D.M., and Lo Verde, P.T. 1989. A retroposon-like short repetitive DNA element in the genome of the human blood fluke, *Schistosoma mansoni*. Chromosoma (Berl.), **97**: 421–428.

Spotila, L.D., Rekosh, D.M., and LoVerde, P.T. 1991. Polymorphic repeated DNA element in the genome of *Schistosoma mansoni*. Mol. Biochem. Parasitol. **48**: 117–120.

Tanaka, M., Hirai, H., LoVerde, P.T., Nagafuchi, S., Franco, G.R., Simpson, A.J., and Pena, S.D. 1995. Yeast artificial chromosome (YAC)-based genome mapping of *Schistosoma mansoni*. Mol. Biochem. Parasitol. **69**: 41–51.

Tautz, D., and Renz, M. 1984. Simple sequences are ubiquitous repetitive components of eukaryotic genomes. Nucleic Acids Res. **12**: 4127–4139.

The World Health Organization Partners for Parasite Control. 2004. Statistics and facts Available from http://www.who.int/wormcontrol/statistics/en/ [accessed 27 February 2004].

The World Health Organization/Special Programme for Research and Training in Tropical Diseases. 2004*a*. Disease portfolio: TDR diseases. Available from http://www.who.int/tdr/diseases/default.htm [accessed 27 February 2004].

The World Health Organization/Special Programme for Research and Training in Tropical Diseases. 2004*b*. Schistosomiasis: strategic emphasis on research. Available from http://www.who.int/tdr/diseases/schisto/strategy.htm [accessed 27 February 2004].

Trouiller, P., and Olliaro, P.L. 1999. Drug development output: what proportion for tropical diseases? Lancet, **354**: 164.

Trouiller, P., Olliaro, P., Torreele, E., Orbinski, J., Laing, R., and Ford, N. 2002. Drug development for neglected diseases: a deficient market and a public-health policy failure. Lancet, **359**: 2188–2194.

van Keulen, H., LoVerde, P.T., Bobek, L.A., and Rekosh, D.M. 1985. Organization of the ribosomal RNA genes in *Schistosoma mansoni*. Mol. Biochem. Parasitol. **15**: 215–230.

Vazquez-Tello, A., Castan, P., Moreno, R., Smith, J.M., Berenguer, J., and Cedergren, R. 2002. Efficient trans-cleavage by the *Schistosoma mansoni* SMalpha1 hammerhead ribozyme in the extreme thermophile *Thermus thermophilus*. Nucleic Acids Res. **30**: 1606–1612.

Venter, J.C., Smith, H.O., and Hood, L. 1996. A new strategy for genome sequencing. Nature (Lond.), **381**: 364–366.

Venter, J.C., Adams, M.D., Myers, E.W., Li, P.W., Mural, R.J., Sutton, G.G., Smith, H.O., Yandell, M., Evans, C.A., Holt, R.A., Gocayne, J.D., Amanatides, P., Ballew, R.M., Huson, D.H., Wortman, J.R., Zhang, Q., Kodira, C.D., Zheng, X.H., Chen, L., Skupski, M., Subramanian, G., Thomas, P.D., Zhang, J., Gabor Miklos, G.L., Nelson, C., Broder, S., Clark, A.G., Nadeau, J., McKusick, V.A., Zinder, N., Levine, A.J., Roberts, R.J., Simon, M., Slayman, C., Hunkapiller, M., Bolanos, R., Delcher, A., Dew, I., Fasulo, D., Flanigan, M., Florea, L., Halpern, A., Hannenhalli, S., Kravitz, S., Levy, S., Mobarry, C., Reinert, K., Remington, K., Abu-Threideh, J., Beasley, E., Biddick, K., Bonazzi, V., Brandon, R., Cargill, M., Chandramouliswaran, I., Charlab, R., Chaturvedi, K., Deng, Z., Di, F., V, Dunn, P., Eilbeck, K., Evangelista, C., Gabrielian, A.E., Gan, W., Ge, W., Gong, F., Gu, Z., Guan, P., Heiman, T.J., Higgins, M.E., Ji, R.R., Ke, Z., Ketchum, K.A., Lai, Z., Lei, Y., Li, Z., Li, J., Liang, Y., Lin, X., Lu, F., Merkulov, G.V., Milshina, N., Moore, H.M., Naik, A.K., Narayan, V.A., Neelam, B., Nusskern, D., Rusch, D.B., Salzberg, S., Shao, W., Shue, B., Sun, J., Wang, Z., Wang, A., Wang, X., Wang, J., Wei, M., Wides, R., Xiao, C., Yan, C., Yao, A., Ye, J., Zhan, M., Zhang, W., Zhang, H., Zhao, Q., Zheng, L., Zhong, F., Zhong, W., Zhu, S., Zhao, S., Gilbert, D., Baumhueter, S., Spier, G., Carter, C., Cravchik, A., Woodage, T., Ali, F., An, H., Awe, A., Baldwin, D., Baden, H., Barnstead, M., Barrow, I., Beeson, K., Busam, D., Carver, A., Center, A., Cheng, M.L., Curry, L., Danaher, S., Davenport, L., Desilets, R., Dietz, S., Dodson, K., Doup, L., Ferriera, S., Garg, N., Gluecksmann, A., Hart, B., Haynes, J., Haynes, C., Heiner, C., Hladun, S., Hostin, D., Houck, J., Howland, T., Ibegwam, C., Johnson, J., Kalush, F., Kline, L., Koduru, S., Love, A., Mann, F., May, D., McCawley, S., McIntosh, T., McMullen, I., Moy, M., Moy, L., Murphy, B., Nelson, K., Pfannkoch, C., Pratts, E., Puri, V., Qureshi, H., Reardon, M., Rodriguez, R., Rogers, Y.H., Romblad, D., Ruhfel, B., Scott, R., Sitter, C., Smallwood, M., Stewart, E., Strong, R., Suh, E., Thomas, R., Tint, N.N., Tse, S., Vech, C., Wang, G., Wetter, J., Williams, S., Williams, M., Windsor, S., Winn-Deen, E., Wolfe, K., Zaveri, J., Zaveri, K., Abril, J.F., Guigo, R., Campbell, M.J., Sjolander, K.V., Karlak, B., Kejariwal, A., Mi, H., Lazareva, B., Hatton, T., Narechania, A., Diemer, K., Muruganujan, A., Guo, N., Sato, S., Bafna, V., Istrail, S., Lippert, R., Schwartz, R., Walenz, B., Yooseph, S., Allen, D., Basu, A., Baxendale, J., Blick, L., Caminha, M., Carnes-Stine, J., Caulk, P., Chiang, Y.H., Coyne, M., Dahlke, C., Mays, A., Dombroski, M., Donnelly, M., Ely, D., Esparham, S., Fosler, C., Gire, H., Glanowski, S., Glasser, K., Glodek, A., Gorokhov, M., Graham, K., Gropman, B., Harris, M., Heil, J., Henderson, S., Hoover, J., Jennings, D., Jordan, C., Jordan, J., Kasha, J., Kagan, L., Kraft, C., Levitsky, A., Lewis, M., Liu, X., Lopez, J., Ma, D., Majoros, W., McDaniel, J., Murphy, S., Newman, M., Nguyen, T., Nguyen, N., and Nodell, M. 2001. The sequence of the human genome. Science (Wash., D.C.), **291**: 1304–1351.

Verjovski-Almeida, S., DeMarco, R., Martins, E.A., Guimaraes, P.E., Ojopi, E.P., Paquola, A.C., Piazza, J.P., Nishiyama, M.Y., Kitajima, J.P., Adamson, R.E., Ashton, P.D., Bonaldo, M.F., Coulson, P.S., Dillon, G.P., Farias, L.P., Gregorio, S.P., Ho, P.L., Leite, R.A., Malaquias, L.C., Marques, R.C., Miyasato, P.A., Nascimento, A.L., Ohlweiler, F.P., Reis, E.M., Ribeiro, M.A., Sa, R.G., Stukart, G.C., Soares, M.B., Gargioni, C., Kawano, T., Rodrigues, V., Madeira, A.M., Wilson, R.A., Menck, C.F., Setubal, J.C., Leite, L.C., and Dias-Neto, E. 2003. Transcriptome analysis of the acoelomate human parasite *Schistosoma mansoni*. Nat. Genet. 35(2): 148–157

Vukmirovic, O.G., and Tilghman, S.M. 2000. Exploring genome space. Nature (Lond.), **405**: 820–822.

Walker, T.K., Rollinson, D., and Simpson, A.J. 1989*a*. A DNA probe from *Schistosoma mansoni* allows rapid determination of the sex of larval parasites. Mol. Biochem. Parasitol. **33**: 93–100.

Walker, T.K., Simpson, A.J., and Rollinson, D. 1989*b*. Differentiation of *Schistosoma mansoni* from S. *rodhaini* using cloned DNA probes. Parasitology, **98**: 75–80.

Waterston, R.H., Lindblad-Toh, K., Birney, E., Rogers, J., Abril, J.F., Agarwal, P., Agarwala, R., Ainscough, R., Alexandersson, M., An, P., Antonarakis, S.E., Attwood, J., Baertsch, R., Bailey, J., Barlow, K., Beck, S., Berry, E., Birren, B., Bloom, T., Bork, P., Botcherby, M., Bray, N., Brent, M.R., Brown, D.G., Brown, S.D., Bult, C., Burton, J., Butler, J., Campbell, R.D., Carninci, P., Cawley, S., Chiaromonte, F., Chinwalla, A.T., Church, D.M., Clamp, M., Clee, C., Collins, F.S., Cook, L.L., Copley, R.R., Coulson, A., Couronne, O., Cuff, J., Curwen, V., Cutts, T., Daly, M., David, R., Davies, J., Delehaunty, K.D., Deri, J., Dermitzakis, E.T., Dewey, C., Dickens, N.J., Diekhans, M., Dodge, S., Dubchak, I., Dunn, D.M., Eddy, S.R., Elnitski, L., Emes, R.D., Eswara, P., Eyras, E., Felsenfeld, A., Fewell, G.A., Flicek, P., Foley, K., Frankel, W.N., Fulton, L.A., Fulton, R.S., Furey, T.S., Gage, D., Gibbs, R.A., Glusman, G., Gnerre, S., Goldman, N., Goodstadt, L., Grafham, D., Graves, T.A., Green, E.D., Gregory, S., Guigo, R., Guyer, M., Hardison, R.C., Haussler, D., Hayashizaki, Y., Hillier, L.W., Hinrichs, A., Hlavina, W., Holzer, T., Hsu, F., Hua, A., Hubbard, T., Hunt, A., Jackson, I., Jaffe, D.B., Johnson, L.S., Jones, M., Jones, T.A., Joy, A., Kamal, M., Karlsson, E.K., Karolchik, D., Kasprzyk, A., Kawai, J., Keibler, E., Kells, C., Kent, W.J., Kirby, A., Kolbe, D.L., Korf, I., Kucherlapati, R.S., Kulbokas, E.J., Kulp, D., Landers, T., Leger, J.P., Leonard, S., Letunic, I., Levine, R., Li, J., Li, M., Lloyd, C., Lucas, S., Ma, B., Maglott, D.R., Mardis, E.R., Matthews, L., Mauceli, E., Mayer, J.H., McCarthy, M., McCombie, W.R., McLaren, S., McLay, K., McPherson, J.D., Meldrim, J., Meredith, B., Mesirov, J.P., Miller, W., Miner, T.L., Mongin, E., Montgomery, K.T., Morgan, M., Mott, R., Mullikin, J.C., Muzny, D.M., Nash, W.E., Nelson, J.O., Nhan, M.N., Nicol, R., Ning, Z., Nusbaum, C., O'Connor, M.J., Okazaki, Y., Oliver, K., Overton-Larty, E., Pachter, L., Parra, G., Pepin, K.H., Peterson, J., Pevzner, P., Plumb, R., Pohl, C.S., Poliakov, A., Ponce, T.C., Ponting, C.P., Potter, S., Quail, M., Reymond, A., Roe, B.A., Roskin, K.M., Rubin, E.M., Rust, A.G., Santos, R., Sapojnikov, V., Schultz, B., Schultz, J., Schwartz, M.S., Schwartz, S., Scott, C., Seaman, S., Searle, S., Sharpe, T., Sheridan, A., Shownkeen, R., Sims, S., Singer, J.B., Slater, G., Smit, A., Smith, D.R., Spencer, B., Stabenau, A., Stange-Thomann, N., Sugnet, C., Suyama, M., Tesler, G., Thompson, J., Torrents, D., Trevaskis, E., Tromp, J., Ucla, C., Ureta-Vidal, A., Vinson, J.P., Von Niederhausern, A.C., Wade, C.M., Wall, M., Weber, R.J., Weiss, R.B., Wendl, M.C., West, A.P., Wetterstrand, K., Wheeler, R., Whelan, S., Wierzbowski, J., Willey, D., Williams, S., Wilson, R.K., Winter, E., Worley, K.C., Wyman, D., Yang, S., Yang, S.P., Zdobnov, E.M., Zody, M.C., and Lander, E.S. 2002. Initial sequencing and comparative analysis of the mouse genome. Nature (Lond.), **420**: 520–562.

Weber, J.L., and May, P.E. 1989. Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. Am. J. Hum. Genet. **44**: 388–396.

Webster, P., Mansour, T.E., and Bieber, D. 1989. Isolation of a female-specific, highly repeated *Schistosoma mansoni* DNA probe and its use in an assay of cercarial sex. Mol. Biochem. Parasitol. **36**: 217–222.

Wippersteg, V., Kapp, K., Kunz, W., Jackstadt, W.P., Zahner, H., and Grevelding, C.G. 2002. HSP70-controlled GFP expression in transiently transformed schistosomes. Mol. Biochem. Parasitol. **120**: 141–150.

Zanotto, E.D. 2000. Genes should not be patentable, but if they are, we all have to do it. Nature (Lond.), **408**: 17.

Zouain, C.S., Azevedo, V., Franco, G.R., Pena, S.D.J., and Goes, A.M. 1998. Identification of genes encoding *Schistosoma manosni* antigens using an antigenic sequence tag strategy. J. Parasitol. **84**: 1307–1310.