Bioinformatics DNA analysis of highly repetitive clones among *Schistosoma haematobium* and other related species

Prepared by:

Fatemah Mohamad Gaith Zhaykh

B.Sc. Biology

AL-Quds University / Palestine

Supervisor: Dr. Ibrahim Abbasi

Co-supervisor: Dr.Mutaz Akkawi

A thesis submitted in partial fulfillment of requirement for the degree of Master of Science in Biochemistry and Molecular Biology.

Faculty of Graduate Studies /AL-Quds University

1428/2007

AL-Quds University
Deanship of graduate studies
Biochemistry and Molecular Biology

Thesis Approval

Bioinformatics DNA analysis of highly repetitive clones among *Schistosoma haematobium* and other related species

Submitted by: Fatemah Mohamad Gaith Zhaykah

Registration No:20410237

Supervisor: Dr. Ibrahim Abbasi

Co-Supervisor: Dr. Mutaz Akkawi

Master thesis submitted and accepted Date:

The names and signatures of examining committee members:

1. Head of committee: Dr. Ibrahim Abbasi _____
2. Co-Supervisor: Dr. Muataz Akkawi _____
3. Internal Examiner: Dr. Samer Barghouthi _____
4. External Examiner: Dr. Khaldoun Niejem _____

Jerusalem- Palestine

1428/2007

Dedication

This thesis is dedicated to:

My parents

My sisters

My brothers, especially to my lovely brother Ahmad

Declaration:

I Certify that this thesis submitted for the degree of Master is the result of my own research, except where otherwise acknowledged, and that this thesis (or any part of the same) has not been submitted for a higher degree to any other university or Institution.

Signed: Fatemah

Fatemah Mohamad Gaith Zhaykh

Date: 30/12/2008

Acknowledgments

After a great thanks to God, I would like first of all to take the opportunity to express my personal appreciation and sincere gratitude to my supervisor Dr. Ibrahim Abbasi from science and technology collage, department of Biology, for his special concern, support, continuous experienced advice, and for his recommendation on my experiments.

I am also grateful to Dr. Mutaz Akkawi for his support, continuous suggestions, and help during this study.

My thanks extend to all Doctors who tough me in the M. Sc. program.

Special thanks to my family for their great help, understanding, and support.

Finally I would like to express my sincere thanks to all friends, brothers, and sisters for their encouragement and support.

Abstract

Human schistosomiasis is one of the most prevalent parasitic diseases that caused by fluke worm (Trematoda, Platyhelminths) of the family Schistosomatidae, genus *Schistosoma.* The disease is mostly prevalent in tropical areas such as sub-Saharan Africa; it causes a morbidity burden of 3.5 million. It is estimated that 652 million people live at risk of infection and that 193 million people are actually infected. Of the 120 million symptomatic cases, 20 million are suffering from severe disease.

There are three species that are able to infect human being, and they cause either urinary or intestinal schistosomiasis. These species are: *Schistosoma haematobium, Schistosoma mansoni*, and *Schistosoma japonicum*. The life cycle start when eggs are excreted by adult worms through feces or urine sample of the infected person ,then an asexual phase in a fresh water snails which serve as an intermediate host. From snails cercariae are released into the surrounding water and can infected humans by skin penetration to continue its sexual phase in the definitive host.

Schistosoma haematobium the causative agent of Urinary schistosomiasis is prevalent in 54 countries in the Middle East and most of the African countries. *S. haematobium* is co-endemic with other related animal species that are known to develop in the same bulinid intermediate snail that transmit *S. haematobium*, also they have the same eggs and cercariae morphology as in *S. haematobium* . Therefore, their differential identification (which was the specific goal of this study) in snail intermediate host is critical when planning for a control program or when monitoring Urinary shcistosomiasis transmission in infested water. So, it is

possible to assess the distribution of *S. haematobium* and its transmission based on differential (diagnostic) secreening test.

In this work, Bioinformatics science (which is a new scientific discipline that merges DNA sequence information and computer science) was applied to identify new repetitive DNA sequences in *S. haematobium* genome. DNA analysis was done in comparison with Gene-Bank Schistosoma sequence information (Blast search), and Nucleotide-nucleotide similarities were determined for all the analyzed genomic DNA clones that belong to *S. haematobium* (191 clones) and for other related animal *Schistosoma* group: *S.bovis [58 clones], S. curassoni[45 clones], S. mattheei [94 clones], S. margebowiei[65 clones]*, and *S. intercalatum[61 clones]*. The DNA analysis for all the genomic DNA clones used in this study allowed the identification of new repetitive DNA clones among *S. haematobium [26clones]* and its other related groups (*S.bovis [5 clones], S. curassoni [5 clones], S. mattheei [8 clones], S. margebowiei [2 clones]*, and *S. intercalatum [7 clones]*). Further analysis were done for each selected repetitive DNA clones that belong to *S. haematobium*, in order to determine their specificity for *S. haematobium* or if they are groups specific.

Based on the sequence information obtained from the newly identified *S. haematobium* repeats [S.h. repeat 24 and S.h. repeat 25] new primers were designed (77f/77r, and 73f/73r). These primer pairs failed to differentiate *S. haematobium* from the other group by one step PCR assay. Therefore a new approach was carried out to achieve the overall aim of this study, the inter-repeat amplification approach was followed which is based on using different combinations of primers from different repetitive clones. The primer combinations (73f/4297) provided a suitable test for *S. haematobium* differentiation.

ملخص

البلهارسيا عبارة عن أحد الأمراض الطفيلية التي تسببها بعض أنواع الديدان الحلقية من عائلة Schistosomatidae وجنس Schistosoma. غالبا ما يسود هذا المرض في المناطق المداريه مثل منطقة جنوب الصحراء الأفريقية. من المعروف أن هناك 652 مليون نسمة معرضون لخطر الاصابة, و حوالي 193 مليون شخص مصاب. بالاضافة الى أن هناك 20 مليون شخص يعانون من أمراض حاده مرافقة للمراحل المتقدمة لهذا المرض.

هناك ثلاثة أنواع من البلهارسيا تكون قادرة على اصابة الانسان ، حيث انها من الممكن ان تسبب أمراض في الجهاز البولي ويسمى المرض البلهارسيا البوليه أو أنها تسبب أمراض في الجهاز الهضمي ويسمى المرض البهارسيا المعوية. وهذه الانواع هي البلهارسيا البولية(*Schistosoma haematobium*) والبلهارسيا المعوية مثل (*Schistosoma mansoni*, and *schistosoma japonicum*). تبدأ دورة حياة البلهارسيا عندما يلوث براز أو بول الشخص المصاب البرك المائية المستخدمة من قبل سكان تلك المنطقة, وغالبا ما توجد هذه البرك في المزارع و أماكن استيطان اللاجئين التي لا تتوفر لها التصريف الصحي . بعد اختراق الجلد ينطلق الطفيلي إلى موقعه حسب نوعه (إلى الامعاء أو إلى المثانة) ومن ثم يبدأ بوضع البيض بكميات كبيرة يتراكم جزء كبير منها في الكبد والمثانة مسببا تليفاً لانسجتهما، أما الجزء القليل جدا من البيض يمكن أن يخرج مع فضلات الشخص المصاب و الذي يمكن له أن يتخلص منها بجانب مصدر مائي راكد وقوقعه حلزونية مخصصة لنمو هذا الطفيل, وبذلك تبدأ دورة حياة جديدة لهذا الطفيل مع إنسان آخر لامس جلده هذه المياه.

تعتبر البلهارسيا البولية هي المسبب الرئيسي لبلهارسيا الجهاز البولي, حيث أنها توجد في 54 بلد من بلاد الشرق الأوسط و في معظم أنحاء القارة الافريقيه, ومن المعروف أن هذه المناطق أيضا موبوءة بأنواع أخرى من البلهارسيا التي تصيب الحيوانات والتي تعتبر من المجموعات التابعة للبلهارسيا البولية نظرا لمدى التشابه الكبير بينهم بالنسبة للشكل الخارجي الخاص بالبويضة والطفيلي التابعه للبلهارسيا البولية، وبالاضافة لذلك فان هذه الطفليات تستخدم أيضا نفس النوع من الحلزونات

التي تستخدمها البلهارسيا البولية. لذلك فان من الضروري ايجاد الفحص المناسب لتميز ولتشخيص البلهارسيا البولية (وهذا هو أحد اهم الأهداف العامة لهذه الدراسة) عن المجموعات الأخرى التابعة لها والتي تصيب الحيوانات حيث أن لذلك الأهمية الكبرى لتخطيط برامج لمراقبة المرض لأطول فترة ممكنة وبتكاليف أقل, حيث أن هذا الفحص في العادة يعتمد بشكل كبير على تقويم وتحديد عدد الحلزونات المصابة في المناطق التي تعتبر موبوءة بشكل كبير جدا وأيضا في المناطق التي يعتبر معدل انتشار المرض فيهامنخفضا.

بالاستناد الى تحليل المادة الوراثية والذي تم بواسطة استخدام علم المعلومات الحيوية (Bioinformatics Science), والذي هو عبارة عن علم يستخدم أحدث تقنيات علم الحاسوب وتكنولوجيا المعلومات لحل كثير من المشكلات البيولوجية والتي تتمثل في عمليات تخزين البيانات، تحليل سلاسل الحمض النووي, وايجاد المورثات, وبذلك كان من الممكن التعرف على الكثير من المقاطع المتكررة في المادة الوراثية الخاصة بالبلهارسيا البولية و المجموعات التابعة لها مثل

S.bovis [58 clones], S. curassoni [45 clones], S. mattheei [94 clones], S. margebowiei [65 clones], and S. *intercalatum[61 clones]*.

تم تحليل الحمض النووي بواسطة استخدام برامج كثيرة على شبكة الانترنت و التي نخص بالذكر منها بنك المعلومات

(Gene-Bank (Blast search) and Nucleotide-nucleotide similarities)

وبهذا التحليل للمادة الوراثية التي تم الحصول عليها من مكتبات المورثات (genomic libraries) التي تم تحضيرها في دراسات سابقة. تم تحديد المقاطع المتكررة في الجينوم الخاص بالبلهارسيا البولية (26 clones) والمجموعات الأخرى التابعة لها وذلك على النحو التالي:

(*S.bovis [5 clones], S. curassoni [5 clones], S. mattheei [8 clones], S. margebowiei [2 clones]*, and S. *intercalatum [7 clones]*)

بالاضافة لذلك التحليل تم إجراء تحليل أخر لهذه المقاطع المتكررة من أجل تحديد عدد المقاطع التي تخص البلهارسيا البولية فقط ولتحديد عدد المقاطع الاخرى التي تخص المجموعة الاخرى.

إعتمادا على هذه المقاطع التي تخص البلهارسيا البولية بشكل منفرد تم تصميم بوادئ جديدة (77f/77r, and 73f/73r) لتستخدم في تحقيق الهدف العام من هذه الدراسة.

بعد ان تم تطبيق هذا الفحص بشكل عملي كانت النتائج سلبية لان استخدام هذا الزوج من البوادئ لم يكن كافيا لتميز بين البلهارسيا البولية عن المجموعات الأخرى التابعة لها, لذلك كان من الضروري اتباع طريقة اخرى من اجل تحقيق الهدف المطلوب والتي سميت ب (inter-repeat amplification), والتي كان مبدأها العمل على استخدام بوادئ كثيرة تم الحصول عليها من مقاطع متكررة تم التعرف عليها في هذه الدراسة ومقاطع اخرى تم التعرف عليها مسبقا مثل(DraI) , وبذلك تم الحصول على بادئ جديد(73f/4297) استخدم لتحقيق الهدف المطلوب .

Chapter One

Introduction and Literature review

1.1: Schistosomiasis:

Human schistosomiasis or Bilharziasis is one of the most prevalent parasitic diseases in the tropics with a huge impact on public health. The World Health Organization (WHO) estimated that 652 million people live at risk of infection and that 193 million people are actually infected. The disease is mostly prevalent in sub-Saharan Africa and cause a morbidity burden of 3.5 million(1), also It is the major health risk in the rural areas of Central China, African countries including Egypt, South America, and Asia (2). Bilharziasis was named after Theodor Bilharz, German pathologist who identified the worms that cause this disease in 1851. The disease is caused by an infection with fluke worms (that belongs to phylum Platyhelminths, class Trematoda, family Schistosomatidae, and genus *Schistosoma*) (3), and according to the location site of the parasite in the final host, it is divided into two types: intestinal and urinary schistosomaisis. There are three major species able to infect human being and these are prevalent in different areas of the world; these are (*Schistosoma mansoni*, *Schistosoma japonicum*, *Schistosoma haematobium*).

The disease is often associated with open water bodies or water resource development projects, such as dams and irrigation schemes, these parasite use fresh water snail as intermediate host which are infected by larvae hatching from the parasite eggs that are excreted with the faeces or urines of schistosomiasis patients. People are infected by contact with water used in normal daily activities such as personal or domestic hygiene and swimming, or by professional activities such as fishing, rice cultivation, and irrigation (4). The endemic cycle of schistosomiasis transmission is affected by different factors, that include: the rate of human contact with transmitting water bodies, the proliferation of

intermediate host snails within these water bodies, and the rate of successful entry and development of schistosome stages in both molluscan and human hosts(5). At present, tools that are used for estimating schistosomiasis transmission potential focus on monitoring the components of the human-snail-water interface, where transmission occurs. Surveillance techniques include monitoring schistosome egg output by humans, human water contact activities, snail infection rates, and, in some cases, numbers of cercariae in the water. However, the effective level of water-contamination with schistosome ova, which culminates in the infection of host-snails by miracidia, has not yet been directly determined, because the conditions leading to transmission of schistosomiasis vary greatly between countries and schistosome species (6, 7).

Schistosomiasis control now is being the most important project for the WHO, because schistosomiasis is a major source of morbidity and mortality for many developing countries in Africa, South America, Caribbean, Middle East, and Asia.Therefore schistosomiasis being the second disease after malaria which is the most common infectious disease worldwide. The control projects of schistosomiasis that were applied to reduce the rate of disease distribution include: prevention of human from infections, prevention of snail infections, control of snail distribution and public health measures. Although these control efforts have been considerable in many endemic areas, but in others control is difficult or impossible due to financial constraints, poor living conditions, human behavior, and due to the strong influence of agricultural development on the generation of snail habitats (5, 4).

1.1.1 Geographical distribution of Schistosomes:

The distribution of Schistosomiasis is largely influenced by the distribution of the intermediate snails hosts whose distribution, in turn is dependent on environmental and

ecological factors, although these worms have a wide geographical distribution that causing schistosomiasis the most significant helminth disease in different countries, because the number of people estimated to be infected or at risk of infection has not been reduced and may be increasing due to population growth and increased water development projects in endemic areas. Below is a summary of the geographical distribution of human schistosomes:

1- *Schistosoma manssoni*: This species has a widespread distribution, being found in 55 countries (figure1.1.1.1a) such as parts of the Middle East, Africa(Djibouti, Mauritania, Senegal, Madagascar, Arabian peninsula, Egypt, Libya, Sudan, and Somalia), South America (Brazil, Surinam and Venezuela), and in the Caribbean islands (including Puerto Rico, St Lucia, Guadeloupe, Martinique, Dominican Republic, Antigua and Montserat),(8).

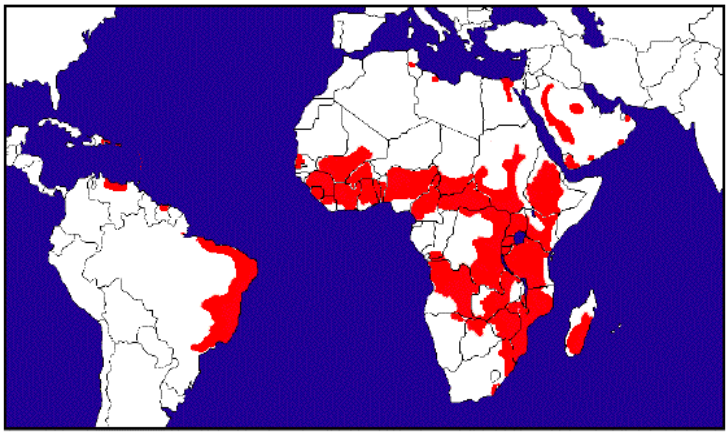
2- *Schistosma haematobium*, It considered as endemic parasite in large parts of Africa (figure1.1.1.1b)such as Kenya, Chana, Mali, Cameroon, Nigeria, Egypt, Libya, Yemen, Sudan, Khuzestan, Province in Iran, Madagascar, around the southern shores of the Mediterranean and Mauritius, parts of Saudi Arabia, Latin America. In addition to that, these area are co-endemic with other related schistosoma parasites that are infect different animal species, i.e; *S. bovis*, *S. mattheei, S. curassoni, S. intercalatum,* and *S. margrebowiei* (9).

3- *Schistosoma japonicum:* This species can be found in the Far East (8), particularly in China, the Philippines, Thailand, Southeast Asia, India, and Indonesia (figure1.1.1.1b).

1.1.2 Schistosomes life cycle:

Schistosomes have a typical trematode vertebrate-invertebrate lifecycle that utilizing humans being the definitive host. The life cycles of all human schistosomes (fig.1.1.2.1) are broadly similar, but they different in their intermediate host. For *S. haematobium* the *Bulinus* Snails serve as the intermediate host, Biomphalaria Snails for *S. manssoni*, and Oncomelania

A)



B)

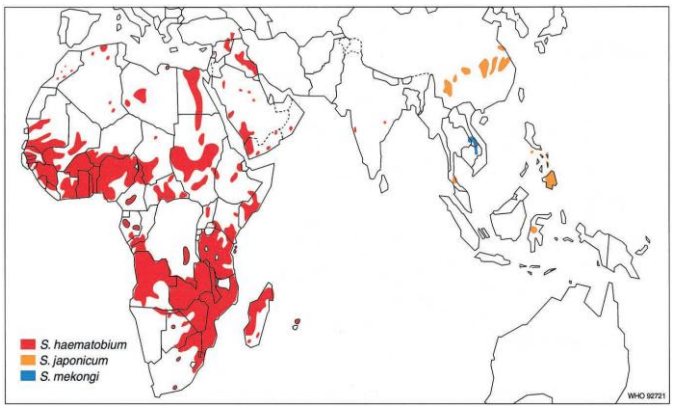


Figure1.1.1.1Geographical distribution of *S. manssoni* (A), *S. haematobium* and *S. japonicum* (B). (Adapted from Tropical Medicine Central Resource).

Snails for *S.japonicum*. The life cycle starts when the parasite eggs are released into the water coming from fecal or urine material of infected individuals, hatching on contact with fresh water to release the free-swimming miracidium, then these miracidia will infect fresh-water snails by penetrating the snail's foot. After infection, close to the site of penetration, the miracidium transforms into a primary (mother) sporocyst. Germ cells within the primary sporocyst will then begin dividing to produce secondary (daughter) sporocysts, which migrate to the snail's hepatopancreas. Once at the hepatopancreas, germ cells within the secondary sporocyst begin to divide again, producing thousands of new parasites, known as cercariae; which are capable of infecting mammals. Cercariae emerge daily from the snail host they are affected with different factor including temperature and light. Young cercariae are highly motile, alternating between strong upward movements and sinking to maintain their position in the water. Cercarial activity is particularly stimulated by water turbulence, shadows and human skin chemicals (10). When cercaria have attached to human skin the parasite secretes enzymes that break down the skin's protein to enable penetration of the cercarial head through the skin. As the cercaria penetrates the skin it transforms into a migrating schistosomulua stage.

The newly transformed schistosomulua may remain in the skin for 1-2 days before locating a post-capillary venule; from here the schistosomulua travels to the lungs where it undergoes further developmental changes necessary for subsequent migration to the liver. Eight to ten days after penetration of the skin, the parasite migrates to the liver sinusoids. Then the schistosomes worms develop an oral sucker after arriving at the liver. The nearly-mature worms pair, with the longer female worm residing in the gynecophoric channel of the male. Worm pairs of *S. mansoni* and *S. japonicum* relocate to the mesenteric or rectal veins,

while worms of *S. haematobium* migrate from the liver to the perivesical venous plexus of the bladder, urethras and kidneys.

Parasites reach maturity in 6-8 weeks; at that time they begin to produce eggs. Many of the eggs pass through the walls of the blood vessels, and through the intestinal wall, to be passed out of the body in faeces. *S. haematobium* eggs pass through the urethral or bladder wall and into the urine. Up to half the eggs released by the worm pairs become trapped in the mesenteric veins, or will be washed back into the liver, where they will become lodged, and they are responsible about all the pathological consequences of the disease. Worm pairs can live in the body for an average of four to five years, but may extend up to 20 years (11).

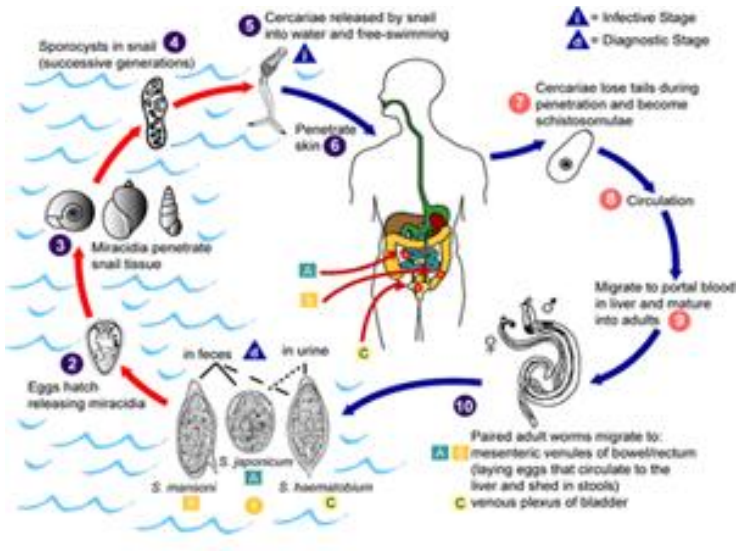


Figure1.1.2.1. The life cycle of schistosoma parasite (Adapted from CDC image sources)

1.1.3: Pathogenesis:

Intestinal schistosomiasis:

Intestinal schistosomiasis is the most common manifestation of infection with *S. mansoni*, and *S.japonicum* in endemic areas such as Africa, Eastern Mediterranean, Caribbean and South America. In the case of intestinal schistosomiasis, the worms reside in the blood vessels lining the intestine, especially in the mesenteric vessels where they establish a chronic infection. When eggs produced by adult female worm they are trapped in the intestinal wall and they cause an immune system reaction called a granulomatous reaction (granuloma formation). Granuloma will be followed by thickness and inflammation of the intestinal wall, that ends with a fibrotic reaction in the intestinal wall leads to mechanical obstruction. Granuloma formation is thought to facilitate egg migration into the intestinal lumen and eggs are eliminated in the stools.

Symptoms are mainly characterized by alternating bloody diarrhea, abdominal pain, and tenesmus associated with granular inflammation and polyps formation, these are more frequently reported as symptoms of acute infection, while symptoms of chronic *S. mansoni* infection can be describe as hepatoesplenic-megaly associated with granulomatous and fibrotic reactions of the parasitic eggs in the liver (12).

Urinary schistosomiasis:

S. haematobium the causative agent of urinary schistosomiasis, is endemic in the Middle East and most of the African continent, where the suitable habitats are abundant for *Bulinus* snails. The urinary schistosomiasis involves the infection of urinary tract and the bladder human (13). The produced eggs by adult females affect the lower ends of the urethras and seminal vesicles. However, other organs can be infected such as liver and lungs (14).

The first clinical indication for urinary schistosomiasis is haematuria which followed by tiredness and general ill health. When eggs are highly concentrated in the tissue of the bladder and lower urethra, granulomas are formed, the bladder wall become enlarged, irregularities bladder masses, and dilation of the upper urinary tract will occur, this will be followed by fibrosis and ulceration that will lead finally to loss the function of the bladder. While the early imaging diagnostic findings of the urethras can be diagnosed by contrast urogram that show persistent filling of the lower segment of the urethras, thickening and nodularity of the ureteric walls will be followed by dilation in the lower ureteric segment, due to constriction within the bladder wall (Fig. 1.1.3.1a). In the developing stage of the disease the nodularity in the lower urethras will be increased, and it become dilated in non regular form. As these changes progress, the dilatation may affect the whole urethra (Fig. 1.1.3.1b) and fine ureteric calcification may be seen on ultrasound. Later, these calcified areas coalesce and the urethras appear "calcified", even along their full length. The bladder changes at this stage include the presence of papillomas and the calcifications in the bladder wall, which result from the lying of eggs in the submucosa of the bladder.

If urinary schistostomiasis is untreated, the urethra obstruction and nodular changes will increase, and marked cystitis; patchy ureteric and bladder calcification may be occurred. On ultrasound, the bladder walls will be thickened and may be so heavily calcified that they appear solid, which mean that the bladder will lose its ability to empty (15).

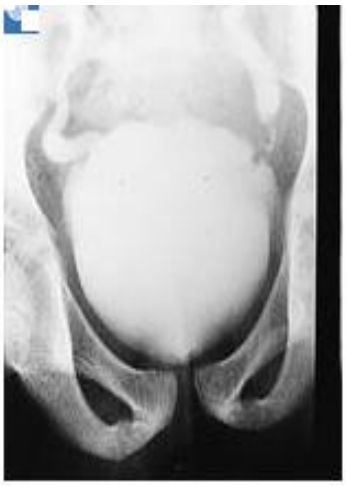
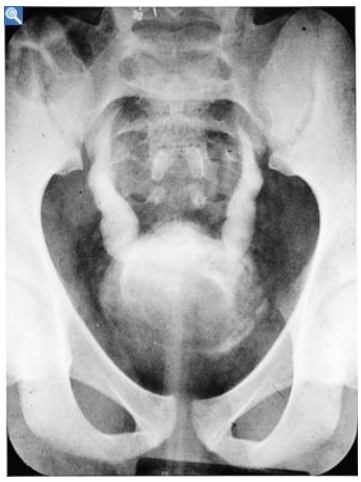
a)  b) 

Figure 1.1.3.1a: Early stages of *S. haematobium* infection; The Contrast Urogram show persistent filling in the bladder, also it could show strictures and dilatation in the ureter.
b) The late stages of *S. haematobium* infection. The ultrasound show markedly dilated in the urethras with lower third strictures and calcification in the walls of the bladder.

Liver involvement: Although the changes in the urinary tract and bowel which result from schistosomiasis is clinically important, still the damage which the parasite do to the liver is more likely to be fatal in the majority of cases. Clinically, the liver and spleen enlarge in the early stages. The portal pressure rises, but ascites does not immediately develop. When the fibrosis increases, it will develop a collateral circulation, which leads to ascites formation with its all clinical sequelae that will occur, which include portal tract thickened, and liver atrophy (16).

The pathogenesis of schistosomiasis is mostly due to trapping of eggs in the surrounding tissue and it can be divided into two main stages:

1- Acute schistosomiasis: Also known as (Katayama fever) is an early clinical manifestation of schistosomiasis that occurs several weeks after infection with Schistosoma species. It may

correspond to the first cycle of egg deposition and it is associated with marked increased in the peripheral eosinophilia and circulating immune complexes**.** The name of Katayama fever which is the second name of acute schistosomiasis came from Katayama valley in Japan which is a hyperendemic area for schistosomes. Acute schistosomiasis is mostly common with *S. japonicum*, *S. mansoni* and *S.haematobium* infections and it is most likely to occur in heavily infected individuals after primary infection. Symptoms usually resolve over several weeks, but the syndrome can be fatal if it was not treatment (17). The clinical presentation of Katayama fever is not specific for schistosomiasis, in spite of its involvement in respiratory and abdominal symptoms, and it can result with schistosomiasis misdiagnosis (18).

Mild skin lesions may develop in acute infection within hours after exposure to cercariae, but significant dermatitis is rare with the major human schistosomal pathogens, probably because the invading and developing cercariae are minimally immunogenic. However, abortive human infection with animal schistosomal groups may cause marked dermatitis or swimmer's itch(18).

2- Chronic schistosomiasis: The pathology of chronic schistosomiasis, which has clinical symptoms that is far more common than the acute form of the infection**,** because schistosomal eggs are highly immunogenic and it will induce local immune responses that involves granuloma formation associated with fibrotic changes, while cercarial and adult worms are minimally immunogenic (19). Egg retention and granuloma formation (fig.1.1.3.2) in the bowel wall (usually *S. mansoni* or *S. japonicum*) may cause bloody diarrhea, cramping, and eventually inflammatory colonic polyposis, while in *S. haematobium* infection, fibrosis of the bladder lead to ureteric obstruction, and renal failure due to hydronephrosis. Unshed

eggs, which are swept back to the portal circulation, lodge there and induce granulomatous reactions in the portal tracts.

Heavy infestations are more likely to produce hepatic disease, that include fibrosis of hepatic tissue, portal hypertension with the usual possible sequelae, such as splenomegaly, ascites, esophageal bleeding, also through all these pathogenic changes eggs can reach the pulmonary circulation and cause pulmonary granulomatosis associated with fibrosis which end by pulmonary hypertension(20).

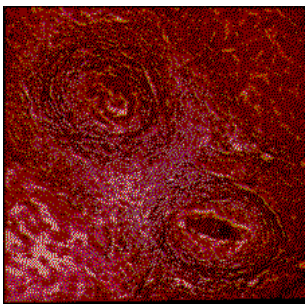


Figure 1.1.3.2: Granuloma formation in the liver, which result from eggs retention (after Liping et al.(20)).

1.1.4: Signs and symptoms:

Many individuals do not experience symptoms, but if present, it usually takes three to six months to appear, and it may take a year or more to develop. However within twelve hours of infection, the patient may be suffer from a light red rash, commonly referred to as "swimmer's itch," due to irritation at the point of entrance. The clinical symptoms include general ill feeling, bloody diarrhea, abdominal pain, eosinophilia, extremely high white blood cell count, Fatigue, Cystitis and ureteritis with hematuria, which can progress to bladder cancer, Cough (with pulmonary hypertension), a typical chest pain (with pulmonary hypertension), Hepatosplenomegaly, Ascites with portal hypertension figure.1.1.4.1 (21).

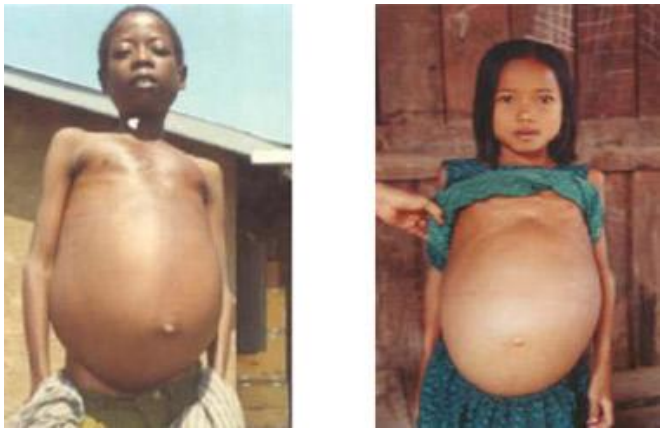


Figure1.1.4.1: Two patients suffering from hepatosplenomegaly associated with ascites (after Nguyen et al. (21)).

1.1.5. Diagnosis:

Schistosomiasis can be diagnosed in humans by different method such as:

1) Microscopic examination.

Microscopic identification of eggs in stool or urine is the most practical method for schistosomiasis diagnosis, although stool examination is more common of the two, and it should be performed when infection with *S. mansoni* or *S. japonicum* is suspected, while urine examination should be performed if *S. haematobium* is suspected. Differentiation of schistosomal species from other parasite can be done by using the microscope depending on the morphology of the eggs, for example urinary bladder schistosomes [*schistosoma haematobium*] have eggs with terminal spine (fig.1.1.5.1a), intestinal schistosomes [*S. mansoni*] have eggs with lateral spine, while eggs of *S.japonicum* have no spines (fig. 1.1.5.1b). Examination of stool sample can be done using a thick smear clarified with glycerol (Kato-Katz technique), and the scientific unit that used for eggs measurement in the feces of presenting patients is egg/gram (epg), but urine sample examination can be done qualitatively after centrifugation. However, eggs are not shed at a steady rate during the day,

and quantitative egg counts are useful for determining the degree of infestation and response to therapy. Therefore, 24-hour urine collections may be recommended (22).

a) **b)**

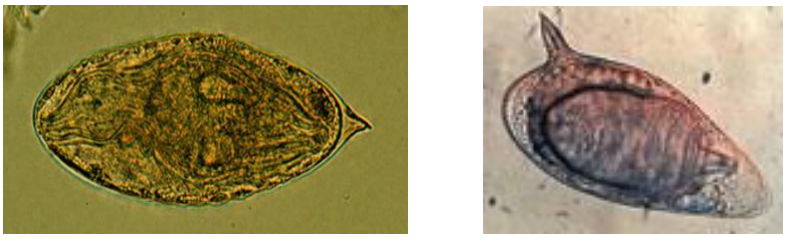


Figure 1.1.5.1: Shapes of certain species of schistosomes eggs .A). *Schistosoma haematobium* egg with terminal spine. B) *Schistosoma mansoni* egg with lateral spine (adapted from www.kstate.edu/).

2) Immunodiagnostic test:

Enzyme-linked immunosorbent assay (ELISA) is the most available test that can be used to confirm schistosomiasis. This test mainly depend on the detection of antibodies or antigen that obtained from the serum samples of infected individuals, however antibody detection tests provide only indirect proof of exposure, because they are molecules produced by the host's immune response to the parasite, so prepatent and early infections may not have stimulated a detectable antibody response, and this will result in obtaining false negative tests. Similarly ELISA test for antibody detection will not be able to differentiate between old and new schistosoma exposure. While antigen tests, by definition, detect molecules produced by the parasite and are therefore just as much direct proof of infection as finding eggs. Although antigen levels are often significantly lower than those of antibodies, but theoretically, are directly related to the number of worms present in the tested samples (23).

3) Ultrasonography and X-ray imaging methods:

This method has been applied to visualize lesions in bladder wall caused by trapped *S. haematobium* eggs, it is a sensitive means of assessing hepatosplenic conditions with periportal fibrosis or urinary obstruction. Also it can demonstrate portal collaterals, periportal adenopathy, urethral obstruction, and obstructive nephropathy. These imaging techniques are best to be used schistosoma diagnosis at late stages (22).

1.2: Snail intermediate host:

1.2.1: Biology of snails:

Many species of freshwater snail belonging to the family *Planorbidae* are considered to be schistosoma intermediate hosts because they harbor the asexual stages (24). Freshwater snails of the genus *Bulinus* act as the intermediate hosts for *Schistosoma haematobium* and other related animal schistosoma groups. This type of snails is present in different tropic regions such as many African countries. There are currently 37 species of *Bulinus* are recognized, but the specificity of the snail parasite interaction is such that only certain species are involved in transmission of the parasite. The genus can be divided into four major groups (species):

1- *B. forskalii* group contains 11 species with slender shells and usually high spires, and it is practically distributed in Africa and its surrounding islands.

2- *B. africanus* group has 10 species, it is practically distributed in many tropical regions, and it is the most common intermediate host of *schistosoma haematobium*.

3- *B. truncatus/tropicus* complex, which contains 14 species, is mainly distributed in Africa , but the distribution of some of these species can be extend to the Middle East, Mediterranean islands and some islands in Peninsula.

4- *B. reticulatus* group contains two species that have restricted distributions in central and southern Africa (25).

As intermediate host for *Schistosoma mansoni*, *Biomphalaria pfeifferi* plays a major role in the transmission of intestinal bilharziasis in the endemic areas of South Africa and Neotropical region. Ten species of *Biomphalaria* have been identified: The classical identification of these species is based on comparison of morphological characteristics of the shell and male and female reproductive organs. So all these snails have non-operculate with tightly coiled, flattened, and lens shaped shells, with diameters up to 7 to 22mm depending on the particular species (26).

S. japonicum utilizes amphibious snails of the genus *Oncomelania* as its intermediate hosts. These snails can survive periods of drought because they possess an operculum capable of closing the shell opening. They live both in and out of water in humid areas such as poorly tilled rice fields, sluggish streams, secondary and tertiary canals of irrigation systems, swamps and roadside ditches(27).

All species of *Biomphalaria* and *Bulinus* are hermaphrodite, possessing both male and female organs and being capable of self- or cross-fertilization. The single specimen can invade and populate a new habitat. The eggs are laid at intervals in batches of 5–40, each batch being enclosed in a mass of jelly-like material. The young snails hatch after 6–8 days and reach maturity in 4–7 weeks, depending on the species and environmental conditions such as temperature and food availability which considered as the most important limiting factors among snails life cycle. A snail lays up to 1000 eggs during its life, which may last more than a year (28).

1.2.2: Snails Ecology:

Snail habitats include almost all types of freshwater bodies ranging from small temporary ponds and streams to large lakes and rivers. Within each habitat, snail distribution may be patchy and detection requires examination of different sites. In general, the aquatic snail hosts of schistosomes occur in shallow water near the shores of lakes, ponds, marshes, streams and irrigation channels. They live on water plants and mud that is rich in decaying organic matter. They are most common in waters where water plants are abundant and in water moderately polluted with organic matter, such as faeces and urine, therefore they always found near human habitations. They can also be found on rocks, stones covered with algae. Furthermore plants and algae serve as substrates for feeding and oviposition, also it will provide the protection for the snail from high water velocities and predators such as fish and birds (29).

Snail densities vary significantly with the season, so their distribution and densities will be increased in the rainy season because water is the suitable habitat for the snail, while during the dry season their densities will be decreased , However, a proportion of some snail species are able to withstand desiccation for months while buried in the mud bottom by sealing their shell opening with a layer of mucus, and their oviposition can be protected by concentrated the eggs in clearly defined sites called breeding pockets, which protect the eggs from the lowering of the water level, as well as most species can survive outside water for short periods. For reproduction, temperatures between 22°C and 26°C are usually optimal, but *Bulinus* snails in Ghana and other hot places have a wider temperature range. The snails can easily survive between 10°C and 35°C. They are not found in salty or acidic water. In most areas, seasonal changes in rainfall, water level and temperature cause marked fluctuations in snail population densities and transmission rates (30).

1.3: Differentiation of *Schistosoma haematobium:*

1.3.1: Morphology and the Physical Description of *schistosoma haematobium*.

Human are the only significant definitive host of *Schistosoma haematobium*. The fully embryonated eggs has no operculum and measured about 112-170µm by 40-70 µm); and it can escape from the body in the urine. The eggs are light yellowish brown with terminal spine (31). The adult male worm contains minute integumentary tuberculations, smaller than those found on adult *mansoni.* Its adult males are around 10 mm and females are 15 mm in length. Both sexes of *S. haematobium* have a strong oral sucker and a smaller posterior ventral sucker. Males have a gynecophoral canal where females are usually located (figure1.3.1.1). This gynecophoral canal also was used to transfer nutrients and hormones from the male to the female. Males genital system have five to nine testes and no cirrus pouch, cirrus or prostate cells, and the genital pore is located directly behind the ventral sucker. While females have only one ovary near the center of the body that can contain 20 to 100 eggs in the uterus at one time (32).

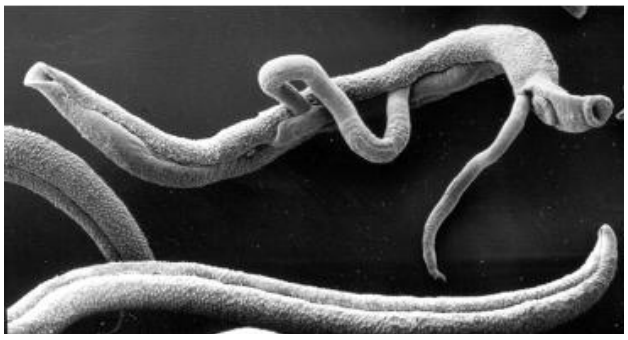


Figure 1.3.1.1.*Schistosoma haematobium* adult worm.
(Adapted from www.york.ac.uk/res/schisto/)

1.3.2: Schistosoma haematobium and terminal spined animal *Schistosoma* groups:

S. haematobium is sympatric with related schistosome parasites (most of other mammals) i.e., *S. bovis, S. mattheei, S. curassoni, S.intercalatum,* and *S. margrebowiei,* and

these related groups are known to develop in the same bulinid intermediate snails that transmit *S.haematobium* (33). Because *S. haematobium* and other related animal *Schistosoma* groups have the same morphology of eggs and cercaria, and they often infect snails inhabiting the same water bodies, the assessment of the transmission of S. *haematobium* by examining cercariae or infected snails is confounded by the need to differentially identify *S .haematobium* form other related animal *Schistosoma* groups (34).

S. bovis is the most widespread animal schistosome groups, with reported distribution in the Middle East, the Mediterranean basin, and most parts of Africa. While the other groups of schistosomes *S. mattheei, S. margrebowiei,* and *S.curassoni,* are present in various parts of Africa. In addition, *S. intercalatum*, which causes mixed intestinal and urinary schistosomiasis in humans, can also be found in transmission foci in Africa that are coinhabited by *S. haematobium* (35).

- ***Schistosoma bovis***: is a group whose final natural hosts are bovines, sheep, and goats. It is distributed throughout Africa, Southwest Asia, and Mediterranean Europe. Different studies have demonstrated the analogies existing between *S. bovis* and other *Schistosoma* groups which affect humans, these similarities include morphological, ecological, physiological, and genetic aspects. Furthermore, a high degree of cross-reactivity among *S. bovis*, *S. mansoni*, and *S. haematobium* has been demonstrated (36).

- ***Schistosoma intercalatum:*** was described in 1934 in Congo; this worm is highly detected in Gabon, Cameroon, Chad, Burkina Faso and Ghana and probably exists in the greater part of Central and West Africa. Adult worms and cercariae resemble those of *S. haematobium*. *S. intercalatum* is the causative agent of rectal schistosomiasis, because usually they are found

in the lower part of the rectum. The more severe forms present as a painful dysenteric syndrome with abdominal pains, but hepatic lesions are relatively minor and localization in the bladder wall is exceptional, however most infection are asymptomatic and discovered by chance(37).

- *Schistosoma curassoni*: is a parasite of sheep, cattle and goats in parts of West Africa. The epidemiology of *Schistosoma curassoni* in several areas of West Africa is similar to the epidemiology of other groups of schistosomes with terminal-spined eggs such as *S.haematobium* and *S. bovis* (38).

- *Schistosoma mattheei: is* primarily a parasite of animals, being found commonly in Cattle, sheep and goats, and it usually distributed in Central and Southern Africa. This worm can use *S. haematobium* intermediate host (*bulinus* snails). Also there are similarities between both worms include morphological (eggs and cercariae), ecological, and physiological aspects (39)**.**

- *Schistosoma margrebowiei*: is normally infecting antelope, buffalo and waterbuck, it is distributed in Southern and Central Africa and it utilizes *bulinus* snails as intermediate host (40).

1.3.3: Identification and characterization of *Schistosoma haematobium*:

The fundamental problem in identification and characterization of *S. haematobium* is due to the presence of other related animal *Schistosoma* groups that are closely related, and are often sympatric with *S. haematobium*; all are depend on freshwater snails of the genus *Bulinus* as intermediate hosts (41). Also, these groups have the same egg morphology with

terminal spine, and their geographical distribution is the same (42). For these reasons the differential identification of *S. haematobium* species from other related animal *Schistosoma* groups needs careful examination and testing, and it is not possible to identify specific shedding cercariae from infected snails.

Given that cercariae belonging to the *S. haematobium* and to other related animal *Schistosoma* groups are not readily distinguishable morphologically, a variety of approaches have been taken for differential identification, with the standard method being infection of laboratory animals and subsequent parasite species identification based on the morphology of the adult worms (43). The following section represents a summary for the major methods used in identification of *S. haematobium* infection in the human definitive host and in the snail intermediate host.

Methods used for identification of *Schistosoma haematobium* in field samples.

1- Laboratory test:

The most common test for the diagnosis of urinary schistosomiasis is the microscopical examination of urine samples looking for the parasites' eggs. However, this test is time consuming, very hazards, requires skilled personal (44), and it has limited sensitivity especially when parasitaemia is low. Moreover the similarity in the morphology of *schistosoma haematobium* egg with other related animal *Schistosoma* groups, microscope examination can't be used for groups differentiation of *shistosoma haematobium* (45).

2- Cercariae shedding:

Until recently, the impact of control programs on transmission potential has been determined by monitoring local rates of *S. haematobium* –infected snails shedding cercariae or by measurement of cercarial numbers at transmission sites (cercariometry).Cercariometry

has not been applied routinely and still presents problems in practice and in data analysis. However, as the rate of snails shedding cercariae is often low, even in areas of high transmission, large numbers of snails are required to detect statistically significant changes in transmission potential overtime, especially following implementation of control measures. In addition to that data on prepatency in field snails have not been widely studied because the detection methods that were employed such as cercariae shedding of field snails taken for the laboratory ,were unsuitable for large – scale ,sensitive monitoring(46).

3- Molecular techniques:

Molecular techniques are now being applied to examine the relationship between different species within *Schistosoma haematobium* group (47). During the last few years, several authors have tried to set new molecular diagnostic tools with higher specificity and sensitivity, some based on the detection of *Schistosoma* DNA by the polymerase chain reaction (PCR) technique (41). Once a fundamental problem presented by the abundance and frequent sympatry of *S. haematobium* with other related animal *Schistosoma* groups is to determine if a particular body of water harboring bulinid snails is actually a transmission focus for the human parasite which is *S. haematobium,* or for the livestock parasite which is mainly other related animal *Schistosoma* groups (*S. bovis*), or for both groups, and to determine which cercariae originating from bulinid snails, taking into account that morphologically differentiation between the two species is generally not feasible in the field (42). PCR, unlike other methods that previously used for identifying snails with prepatent infection; enabled detection of snail infection from its very earliest stages, and thus identifies the entire population of infected snails, regardless of whether they eventually shed cercariae, therefore we can represent in quantitative terms the direct outcome of human contamination of water bodies (43).

The molecular techniques have been applied to large-scale monitoring of infection in snail fields, because differential identification of snails infected with *S. haematobium* is important for further evaluation of integrated control approaches in other endemic regions, and this can be done by undertaken a search for DNA sequences from which primers can be designed for differential identification of *S. haematobium* by means of simple PCR techniques (43). The degree of sensitivity and specificity by using PCR, enabled the identification of individual cercariae and of the infected snails through prepatency, specially when the primers based on repeated DNA sequence that are specific for *schistosoma haematobium*.

1.3.4: *Schistosoma haematobium* repetitive DNA sequence:

Schistosomes have the largest genomic DNA in all parasitic organisms that is estimated to be about ~ 270 mega-bases for haploid genome of *S.haematobium*, arrayed on seven pairs of autosomes and one pair of sex chromosomes. Although no schistosome genome has yet been sequenced in its entirety, but several hundred thousand schistosome expressed sequence tags and genome survey sequences are in the Gene Bank (48). Between about 40%-60% of the *Schistosoma haematobium* genome seems to be repetitive sequences that are rich in AT base pairs (probably 60-70%), (49). These sequences normally do not code for protein, they are varied in size (70 up to 200bp) and they may be arranged in tandem patterns or randomly dispersed throughout *S. haematobium* genome (50).

Types of Repetitive Elements: Repetitive elements are divided into two classes according to the following difference: Their position in the genome, sequence, size, number of copies, and presence or absence of coding regions within them. These two major classes of repetitive elements are called interspersed elements and tandem arrays elements.

- Interspersed repeated elements are usually present as single copies and distributed widely throughout the genome. The interspersed repeats alone constitute about 45 percent of the genome. The best-characterized interspersed repeats are the transposable genetic elements, also called mobile elements or "jumping genes".

- Sequences that are "tandemly arrayed" are present as duplicates, either head to tail or head to head. So-called satellites, minisatellites, and microsatellites, and they are largely exist in the form of tandem arrays, these elements originally acquired their name as "satellites" because they separate from the bulk of nuclear DNA during gradient centrifugation. Sequences repeated in tandem are common at the centromere where the two arms of a replicated chromosome are held together, and at or near the telomeres (the chromosome tips). Because it is difficult to sequence the repetitive DNA sequence that are arranged in tandem arrayed at centromeres and telomeres, this makes it difficult to estimate the copy number, but they certainly represent at least 10 percent of the genome (50).

1.3.5: Bioinformatics science:

In the last few decades, advances in molecular biology and the equipment available for research in this field have allowed the increasingly rapid sequencing of large portions of the genomes of several species. In fact, to date, several bacterial genomes, as well as those of some simple eukaryotes (e.g., *Saccharomyces cerevisiae*, or baker's yeast) have been sequenced in full. Also all the 24 chromosomes of the human being were sequenced.

This increasing of sequence information has necessitated the careful storage, organization and indexing of these information in a suitable form , in order to facilitate the dealing with them in any time ,so for that purpose number of a popular sequence databases, such as GenBank and EMBL, have been growing at exponential rates, also an information

science has been applied to biology to produce the field called **Bioinformatics**, which can be defined as a pure and an applied science dealing with the collection, management, analysis and dissemination of biological data and knowledge, especially with respect to genetics and molecular biology.

The simplest tasks used in bioinformatics concern the creation and maintenance of databases of biological information, such as nucleic acid sequences (and the protein sequences derived from them) which comprise the majority of such databases. Because storage and organization of millions of nucleotides is far from trivial, this requires to design a database and to develop an interface whereby researchers can both access existing information and submit new entries data at any time, and this is only the beginning of bioinformatic science tasks. Bioinformatics analysis of sequence information is done by the aid of **computational** process, and it involves the following functions:

- Finding the genes in the DNA sequences of various organisms.

-Analysis of genome sequence data, particularly the analysis of the human genome project, is one of the main achievements of bioinformatics to date.

- Developing methods to predict the structure and function of newly discovered proteins and structural RNA sequences.

- Clustering protein sequences into families of related sequences and the development of protein models.

- Aligning similar proteins and generating phylogenetic trees to examine evolutionary relationships (51).

1.4: The Importance of this study:

Schistosomiasis is a wide spread disease that affecting many people in developing countries, thus the morbidity rate due to urinary schistosomiasis was estimated at 11,000 deaths per year, 120 million are symptomatic and up to 20 million suffer from the severe consequences of the infection. These calculations do not consider late sequelae and indirect morbidity /mortality due to schistosomiasis infection such as liver disease, portal hypertension, nonfunctional kidney, and squamous cell bladder carcinoma (52). Because chronic morbidity is the major impact of schistosomiasis on ill health, different exercises have been applied to determine the levels of morbidity due to schistosomiasis in many endemic regions especially in sub-Saharan Africa. These exercises have provided estimates of these clinical sequelae and morbidity out comes, as well as evidence of death rate to reach 200,000 deaths per year. For the above mentioned reasons schistosomiasis will be considered as the most important disease with TRD portfolio, and this is the major reason that leads the world health organization (WHO) expert committee to design a strategy for morbidity and transmission control (53).

The strategies that were followed to control morbidity and transmission rate include different methods, and the most important goal of these strategies was the development of more sensitive and specific diagnostic tools to be used in endemic regions with human and animal schistosomiasis. Many barriers make the application of schistosomiasis control strategies almost impossible; it needs a high cost to be implemented because most of the control projects depend on expensive chemotherapy [paraziquantal as the drug of choice] (54).

At the beginning schistosomiasis control programs in the rural areas were mainly focused on the elimination of snails (snail control programs), due to the absence of any anti-

schistosomal drugs in that period. For this purpose copper sulfate was used effectively in reducing snails' number. This approach was stopped later for different reasons such as: persistence of the chemical in the surroundings, high ability of snails to re-colonies and re-populate the treated areas, and also for the bad impact of copper sulfate on material human health, especially when the treated water is the major source for people in that areas (55).

Treatment of water bodies was not based on full knowledge regarding what type of schistosomal parasite found in the water (animal or human schistosomes), therefore and in order to reduce the cost of control strategies (morbidity and snail transmission control) there is a need for a diagnostic techniques to differentiate between *schistosoma haematobium* that cause human urinary schistosomiasis and other animal schistosomal groups (56). The use of a suitable technique for differentiation, the amount and the cost of the chemotherapy will be reduced in each endemic countries, and control projects can be applied for longer periods. The ability of differentiation between both types will facilitate monitoring of the continuous transmission potential (measured as snail pre-patent infection rate), and it could help to determine and to select the suitable time for water treatment. Marked fluctuations in snails population densities and transmission rates will be changed according to the seasonal changes, which are the most important factor that will affect the life cycle of snails and the parasite chance for survival. This fact can be used to draw recommendations for the best time for water treatment by molluscicides and the suitable time for chemotherapy treatment.

The development of molecular tools to be used in the determination of infected water will turn schistosoma monitoring easier, more sensitive and specific. Also they will avoid the use of expensive instruments and elaborate multi-step methods to detect the amplification products that's will provide many developments in the control programs(57).

1.5: Objectives:

The overall aim of the proposed study is to design a PCR test that can differentially identify *S. haemtobium* parasite from other related *schistosoma* animal species. The specific objectives are as follow:

1- To identify *S. haematobium* highly repetitive DNA sequences by the aid of bioinformatics DNA analysis tools.

2- To compare these newly identified *S. haematobium* repetitive DNA sequences with all DNA sequences obtained from other schsitosomal animal groups . This will enable the identification of *S. haematobium* specific DNA segments that provides sequence information for designing new primers suitable for *S. haematobium* PCR species specific test.

3- To compare the flanking regions of similar repeats among *S. haematobium* and the other *schistosoma* animal groups. Also, based on the found differences in these flanking regions a *S. haematobium* species specific PCR test can be designed.

4- To study the efficiency of the above designed primers for *S. haematobium* species specific PCR test in terms of specificity and sensitivity and their evaluation in field studies.

Chapter Two

Material and Methods

1- Shistosoma genomic DNA clones:

The various used DNA clones from *Schistosoma haematobium*, and other schistosomal animal groups *namely Schistosoma mattheei, Schistosoma intercalatum, Schistosoma curassoni, Schistosoma margebowiei,* and *Schistosoma bovis,* were obtained from previously prepared genomic DNA libraries. The selection criteria of these clones were based on the highest obtained signals after their hybridization with total genomic homologous DNA. Some of these clones were published in the Gene Bank (table 2.1).

Table 2.1: Number of the different Schistosoma clones that were analyzed in this study.

Species	Number of clones
S.haematobium	**191**
S.bovis	**58**
S.mattheei	**94**
S.margebowiei	**65**
S.curassoni	**45**
S.intercalatum	**61**
Total	**514**

2-Gene-Bank and analysis of sequence similarities:

All the DNA clones that are included in this study were analyzed for their similarities by the aid of the Gene-Bank information (BLAST search), Nucleotide-nucleotide similarities were determined. Clones that showed a similarity to known sequences to other than *Shcistosoma* parasite and clones that were similar to (*S.mansoni, and S. japonicum*) were

excluded from this study because if they are similar at the genus level then most probably are not *S. haematobium* species specific.

Similarity tables were created for each individual species and they contain the following information: the names of the clones, size, source of the clones, names of similar clones, the size of the similar clones, the range of similarity between the main clone and its similar clones. The generation of similarity tables between the analyzed clones enabled the identification of newly repetitive sequence in *S. haematobium*, and in other *Schistosoma* animal groups. [See table (3.1) and appendix A for summary]. The complete tables for each group specific independently analyzed clones are composed from many pages and were considered as a raw data that could not be included in the thesis.

3-Multiple sequence alignment:

All the identified repetitive DNA clones were compared by the aid of multiple sequence alignment process. This comparison (multiple sequence alignment) was done with the aid of BCM launcher program, together with a BLAST search in the gene bank. In this section the newly identified *S. haematobium* (26 clones) repeats were compared with similar repeats from animal schistosomes. This type of comparison permits the identification of DNA sequences (or segments) that are found in *S. haematobium* and not found in other schistosoma animal groups. Also, DNA analyses of the flanking regions located before and after the repetitive sequence in S. haematobium were analyzed.

6-Primer design and Polymerase Chain Reaction:-

Two pairs of Primers were designed based on sequence information obtained from newly identified repetitive DNA clones. By computer analysis these clones were shown to be *S. haematobium* specific. The first pair of primers was designed from clone S.h.repeat 24 which is 77 base pair (bp) in length; therefore the forward and reverse primers were named as 77f and 77r respectively. While the second pair of primers were design from clone S.h.repeat 25 which is 73 bp in long, and in the same manner the forward and reverse primers were named as 73f and 73r respectively. Figure 3.1 shows the sequence of these newly designed primers.

7- Parasites:

Adult worms that were used as positive control in this study were obtained from different scientists from working in the same field and they were kept in ethanol until subjected to DNA extraction. These included: *Schistosoma haematobium*, *S. b*ovis, *S. margrebowiei*, *S. curassoni*, S. *mattheei,* and S. *intercalatum*.

8- DNA extraction:

DNA was prepared from adult schistosome worms of all groups that include in this study described previously (36): For this purpose the worms were kept in lysis buffer containing (0.1M EDTA, pH 8.0, 0.1M Tris-HCl pH 7.5, 0.2M NaCl, 1% SDS, 0.2% 2-mercaptoethanol and 100 µg Proteinase K (Sigma, St. Louis, USA)) at 65^{o}C for 1-2 hours. This was followed by phenol and chloroform extraction, (an equal volume of phenol to the lysate was added; the mixture was vortexed for 2 min followed by centrifugation for 3 min at 14000 rpm. The same was done with chloroform extraction; the obtained supernatant was subjected to ethanol

precipitation using 5M NaCl and 2.5 volumes of cold ethanol. The precipitated DNA was collected after centrifugation at high speed (14,000 rpm) for 10 minutes. The precipitated DNA was washed by 200 ml of 70% ethanol to remove excess salts, the precipitate then was air dried for 15 minutes and later it was dissolved in 100 ml of TE buffer (10 mM Tris, 1 mM EDTA).

9- PCR assay:

PCR assays were carried out in a total volume of 50 µl containing 200 µM of each dNTPs, 25 pmol of the each primers, 2.5 units of *Taq* polymerase (Sigma-Aldrich, St. Louis, USA) and the target DNA. The reaction buffer was determined by a separate PCR experiment to specify the optimal buffer to be used for each of the different designed primers. A thermal cycler (Bio-Rad, California, USA) was used with a thermal profile involving 5 min at 95^{o}C, followed by 35 cycles each of 30 seconds at 95^{o}C, 30 seconds at 50 ^{o}C, followed by 1 min at 72^{o}C, and a final elongation step at 72 ^{o}C for 10 min.

10- Determination of optimal PCR buffer:

For each new designed PCR primers (77r/77f, and 73f/73r) the optimal PCR buffer was determined, because optimal amplification buffer can be varied if the primers and the type of the DNA was changed. Therefore, modifying the buffer components make it possible to determine the optimal buffer to be used, in order to improve the yield and the specificity of the desired PCR product, this step was done by the aid of Opti-prime PCR optimization Kit (stratagene, California, USA).

1-Materials for optimization include the preparation of 12 buffers (10x concentrated buffer) that vary in their pH (8.8, 8.3, and 9.2). Buffers were prepared according to the below table (Table 2.2) and following the instructions of the manufacturer.

Table 2.2: Optimal buffers components, each buffer composed from the indicated materials to which double distilled water is added up to 1 ml.

Buffer no.	Tris-HCl (1M)	$MgCl_2$ (1M)	KCl (1M)
1	100 µl (pH 8.3)	15 µl	250 µl
2	100 µl (pH 8.3)	15 µl	750 µl
3	100 µl (pH 8.3)	35 µl	250 µl
4	100 µl (pH 8.3)	35 µl	750 µl
5	100µl (pH8.8)	15µl	250µl
6	100µl (pH8.8)	15µl	750µl
7	100µl (pH8.8)	35µl	250µl
8	100µl (pH8.8)	35µl	750µl
9	100µl (pH9.2)	15µl	250µl
10	100µl (pH9.2)	15µl	750µl
11	100µl (pH9.2)	35µl	250µl
12	100µl (pH9.2)	35µl	250µl

Preparation of PCR mixture for optimization assay:

Enough Mixture for 13 PCR reactions was prepared in a final volume equal to 650 µl that contains the following materials:

- 65 µl of 10x buffer (5 µl of each buffer were added in one of the 12 PCR tube).

- 13 µl of 10mM dNTPs.

- 13 µl of each of the two oligonucleotide primers.

- 13 µl of Taq DNA polymerase.

- 10 µl of *S. haematobium* genomic DNA (1ng/µl).

- 528 µl of sterile distilled water sufficient to bring the final volume of the reaction up to 650 µl.

The reaction was kept in the PCR machine with a thermal profile involving 5 min at 95°C, followed by 35 cycles each of 30 seconds at 95°C, 1 seconds at 52°C, followed by 72°C for 1 minute, and a final elongation step at 72°C for 10 min. The PCR product was analyzed on agarose gel electrophoresis as indicated below.

Note: Optimization was done for each of the two pairs of the newly designed primers.

PCR sensitivity test:

This was done by amplifying of 10–fold dilution of total *S. haematobium* genomic DNA (1ng, 0.1ng, 10pg, and 0.1pg), and of other animal schistosome groups (10ng, 1ng, and 0.1ng). PCR reaction was done as described before, and for each of the newly designed pair of primers. This step will facilitate the determination of suitable concentration of *S. haematobium* genomic DNA that can be amplified by the aid of the tested primers, also it will be needed in the determination of the used DNA in the following specificity studies. In brief the PCR reaction was included the following materials as calculated for 10 reactions (final volume of 500µl, and contains 10µl of each primer (25pmoles/reaction), 10µl of Taq DNA polymerase, 50µl of 10x buffer, 10µl of dNTPs (10mM), 5µl of the diluted DNA/reaction, and 380µl of sterile distilled water).

PCR specificity test:

Based on the sensitivity results it was enabled to determine which DNA concentration is suitable for the specificity studies. For this purpose different concentrations of genomic

DNA from *haematobium* and other schistosomes animal groups were used. PCR reaction was done as described before, and for each of the newly designed pair of primers.

DNA analysis by Agarose gel electrophoresis:

- Agarose gel was prepared (1.2%: 1.2 grams in 100 ml of TAE buffer (40 mM Tris, 20 mM acetic acid, 1mM EDTA)). The solution was boiled to dissolve the agarose, then Ethedium Bromide was added in a concentration of 10µg/ml. The mixture was stirred to disperse Ethidium Bromide, and then poured it into the gel rack. A comb was inserted at one side of the gel, about 5-10 mm from the end of the gel and the gel was allowed to cool and solidify. Tested DNA samples (Genomic DNA or PCR products) were analyzed by mixing 2µl of the loading buffer and 10µl of the DNA sample, the mixture was then loaded in separate well for each of the tested samples. For DNA band size determination DNA marker pUC8 (MBI Fermentas, Germany) was used.

Chapter Three

Result

Computerized nucleotide sequence bioinformatics analysis:

By the aid of different bioinformatics searching engines it was possible to generate different tables that include repeated DNA clones belonging to *S. haematobium* [26 clones] and other animal schsitosoma groups (*S.bovis[5 clones], S. curassoni[5 clones], S. mattheei[8 clones], S. margebowiei[2 clones]*, and *S. intercalatum[7 clones]*). The complete tables that were created for each of the specific groups were considered as the raw data and they were not included in the theses.

DNA repetitive segments were determined after performing DNA analysis for the previously obtained genomic DNA clones for *S. haematobium* and for the other animal schistosmal groups. *S. haematobium* repeated DNA clones describe in table 3.1 were subjected to further computational analysis that were performed for the repeated sequence itself and for its flanking region in order to determine which of these repetitive sequences could be *S. haematobium* group specific. This analysis was done by searching for similar clones in the Gene-Bank by Blast search looking for nucleotide-nucleotide similarities. The similar clones were aligned together using BCM luncher multiple sequence alignment utilities **(http://searchlauncher.bcm.tmc.edu/multi-align/multi-align.html).** The result of this DNA analysis was summarized into two tables, the first summarizing the repeated DNA sequences that are specific for *S. haematobium* only, which may refer to the repetitive DNA sequence itself or to the flaking region of that repeat (Table3.2). The second table includes all the identified repeated DNA clones that are not specific for *S. haematobium*, and they can be used for all schistosoma groups identification (genus specific) (Table 3.3).

Table 3.1: Summary for newly identified repeats in *S. haematobium* and other related groups.

Clone Number	Clone name	Clone Source	clone size	Repeat size	Range	Appendix
1	*S.h.*repeat.1	*S. h.*	193 bp	118 bp	75-193	A.1
2	*S.h.*repeat.2	*S. h.*	313 bP	99 bp	213-312	A.1
3	*S.h.*repeat.3	*S. h.*	302 bp	277 bp	24-301	A.1
4	*S.h.*repeat.4	*S. h.*	160 bp	49 bp	111-160	A.1
5	*S.h.*repeat.5	*S. h.*	358 bp	46 bp	1-47	A.1
6	*S.h.*repeat.6	*S. h.*	392 bp	54 bp	18-72	A.1
7	*S.h.*repeat.7	*S. h.*	193 bp	113 bp	6-119	A.1
8	*S.h.*repeat.8	*S. h.*	287 bp	186 bp	1-287	A.1
9	*S.h.*repeat.9	*S. h.*	285 bp	284 bp	1-285	A.1
10	*S.h.*repeat.10	*S. h.*	143 bp	131 bp	12-143	A.1
11	*S.h.*repeat.11	*S. h.*	76 bp	75 bp	1-76	A.1
12	*S.h.*repeat.12	*S. h.*	512 bp	96 bp	299-395	A.1
13	*S.h.*repeat.13	*S. h.*	258 bp	172 bp	73-245	A.1
14	*S.h.*repeat.14	*S. h.*	170 bp	94 bp	76-170	A.1
15	*S.h.*repeat.15	*S. h.*	251 bp	250 bp	2-251	A.1
16	*S.h.*repeat.16	*S. h.*	420 bp	76 bp	2-78	A.1
17	*S.h.*repeat.17	*S. h.*	352 bp	253 bp	3-256	A.1
18	*S.h.*repeat.18	*S. h.*	307 bp	96 bp	76-172	A.1
19	*S.h.*repeat.19	*S. h.*	343 bp	59 bp	199-258	A.1
20	*S.h.*repeat.20	*S. h.*	108 bp	55 bp	53-108	A.1
21	*S.h.*repeat.21	*S. h.*	299 bp	57 bp	242-299	A.1
22	*S.h.*repeat.22	*S. h.*	111 bp	47 bp	64-111	A.1
23	*S.h.*repeat.23	*S. h.*	162 bp	161 bp	1-162	A.1
24	*S.h.*repeat.24	*S. h.*	77 bp	76 bp	1-77	A.1

25	*S.h.*repeat.25	*S. h.*	73 bp	73 bp	1-73	A.1
26	*S.h.*repeat.26	*S. h.*	174 bp	174 bp	1-714	A.1
27	*S.b.*repeat.1	*S.b*	409 bp	88 bp	80-168	A.2
28	*S.b.*repeat.2	*S.b*	444 bp	45 bp	181-226	A.2
29	*S.b.*repeat.3	*S.b*	363 bp	113 bp	178-291	A.2
30	*S.b.*repeat.4	*S.b*	245 bp	50 bp	44-94	A.2
31	*S.c.* repeat1	*S.c*	821 bp.	178 bp	366-544	A.3
32	*S.c.* repeat2	*S.c*	428 bp	79 bp	42-121	A.3
33	*S.c.* repeat3	*S.c*	471 bp	120 bp	347-467	A.3
34	*S.c.* repeat4	*S.c*	608 bp	440 bp	5-445	A.3
35	*S.c.* repeat5	*S.c*	872 bp	49 bp	823-872	A.3
36	*S.i* .repeat.1	*S.i*	849 bp	146 bp	391-537	A.4
37	*S.i* .repeat.2	*S.i*	1030 bp	106 bp	1-107	A.4
38	*S.i* .repeat.3	*S.i*	852 bp	198 bp	385-583	A.4
39	*S.i* .repeat.4	*S.i*	861 bp	57 bp	441-498	A.4
40	*S.i* .repeat.5	*S.i*	286 bp	49 bp	237-286	A.4
41	*S.i* .repeat.6	*S.i*	398 bp	40 bp	167-207	A.4
42	*S.marg.* repeat.1	*S.marg*	327 bp	39 bp	216-255	A.5
43	*S.marg.* repeat.2	*S.marg*	1051 bp	32 bp	196-228	A.5
44	*S.mat.* repeat.1	*S.mat*	362 bp	65 bp	200-265	A.6

45	*S.mat* repeat.2	*S.mat*	804 bp	85 bp	365-450	A.6
46	*S.mat.* repeat.3	*S.mat*	85 bp	59 bp	25-84	Mat
47	*S.mat* repeat.4	*S.mat*	541 bp	175 bp	166-341.	Mat
48	*S.mat* repeat.5	*S.mat*	285 bp	122 bp	72-194	Mat
49	*S.mat* repeat .6	*S.mat*	912 bp	78 bp	834-912	Mat
50	*S.mat* repeat.7	*S.mat*	656 bp	108 bp	548-656	Mat
51	*S.mat* repeat.8	*S.mat*	877 bp	779 bp	1-780	Mat

Table 3.2: Summary of the newly identified *S.haematobium* specific repeats.

Clone number	Clone name	Clone size (bp)	Species	Similar clone	Note	Size of similarity (bp)	Range of similarity	Range of Similarity in the main Clone	Specificity of the main clone
1	S.h.repeat 1*	71	*S.h 1-71*	|gb|DQ831697.1 gb|DQ831684.1| gb|DQ831686.1|	*S.b* *S.h* //	203bp 71bp 334bp	194-168 1-71 178-242	9-35 1-71 9-71	***S. haematobium*** **specific (Its flanking region contain sequence that are specific for *S. haematobium*)**
2	S.h.repeat 2*	517	*S.h* Length 1-300		It does not have any similarity with S, species				***S. haematobium*** **specific,as the result that was obtained from BCM luncher & from searching in the gene bank**
3	S.h.repeat 13	258	*S.h* Length 1-73		It does not haveany similarity with S, species				***S. haematobium*** **specific,as the result that was obtained from BCM luncher & from searching in the gene bank**
4	S.h.repeat 16	420	*S.h* Length 78-420		It does not have any similarity with S, species.				***S. haematobium*** **specific,as the result that was obtained from BCM luncher & from searching in the gene bank**
5	S.h.repeat 4	160	*S.h* Length 1-111		It does not have any similarity with S, species				***S. haematobium*** **specific,as the result that was obtained from BCM luncher & from searching in the gene bank**

Table 3.2 continue: Summary of the newly identified *S.haematobium* specific repeats.

6	S.h.repeat 3*	766	*S.h* Length 1-200		It does not has any similarity with S, species				***S. haematobium* specific,as the result that was obtained from BCM luncher & from searching in the gene bank**
7	S.h.repeat 1	193	*S.h* Length 120-193	gb|DQ831685.1| gb|DQ831687.1|	*S.h* *S.h*	509bp 209bp	432-509 2-76	47-124 49-124	***S. haematobium* specific,as the result that was obtained from BCM luncher & from searching in the gene bank**
8	S.h.repeat 17	352	*S.h* Length 257-352		It doesnot has any similarity with S, species.				***S. haematobium* specific,as the result that was obtained from BCM luncher & from searching in the gene bank**
9	S.h.repeat 4*	343	*S.h* Length 1-198		No significant similarity found				***S. haematobium* specific,as the result that was obtained from BCM luncher & from searching in the gene bank**
10	S.h.repeat 22	111	*S.h* Length 1-63		It does not has any similarity with S, species.				***S. haematobium* specific,as the result that was obtained from BCM luncher & from searching in the gene bank**

***: Clones that have a flanking rejoin that are specific for *S. haematobium*.**

Table3.3: Summary for newly identifed specific repeats in *S.haematobium* and other related groups only.

Clone number	Clone name	Clone size (bp)	species	Similar clone	Notes	Size of similarity (bp)	Range of similarity	Range of Similarity in the main Clone	Specificity of the main clone
1	S.h. repeat 24	77	S.h Length (1-77)	gb|DQ831691.1| gb|DQ831687.1| gb|DQ831685.1| gb|DQ831697.1|	*S.h* // // *S.b*	77bp 209bp 509bp 203bp	1-77 1-76 486-509 168-203	1-77 1-77 54-77 115-150	Group.specific (*S. bovis&S. h.*)
2	S.h. repeat 1	193	S.h Length (1-75)	gb|DQ831685.1| gb|DQ831687.1| gb|DQ831691.1|	*S.h* // //	509bp 209bp 77bp	437-509 3-76 3-77	1-73 1-75 1-75	Group.spesific (*S. bovis&S. h.*)
3	S.h. repeat 19	343	S.h Length (259-343) after		It dose nat has any similarity with S, species.				Group.spesific (*S. mattheei & S. h.*)
4	S.h repeat 25	73	S.h Length (1-73)	gb|DQ831686.1| gb|DQ831684.1|	S.h S.h	334 71	184-249 1-63	2-66 1-46	Group.spesific

Groups specific *Schistsoma haematobium* DNA repetitive sequences:

A newly identified repetitive sequences that were called S.h.repeat 24 and S.h.repeat 25 (Figure 3.1), were candidate repetitive sequences to be used for *S. haematobium* group specific identification using one step PCR assay. One clone is 73 bp and the other is 77 bp, two pairs of oligonucleotide primers were designed for amplification of the full of the identified two repeats. The primers (77f/77r, and 73f/73r) were named according to the repeat length and were designated by (d) or (r) to show if the primer is direct or reverse.

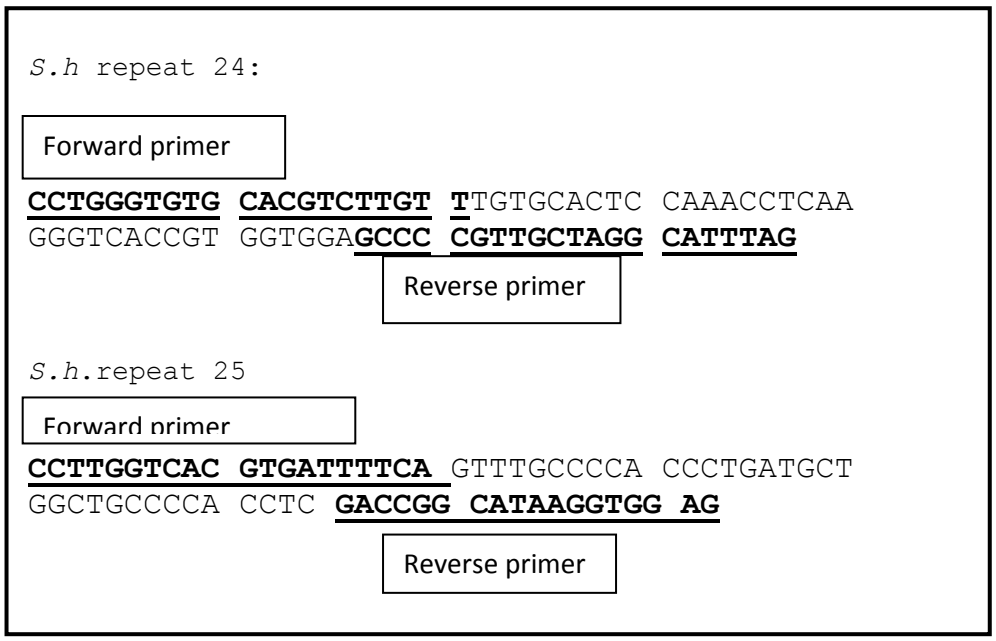
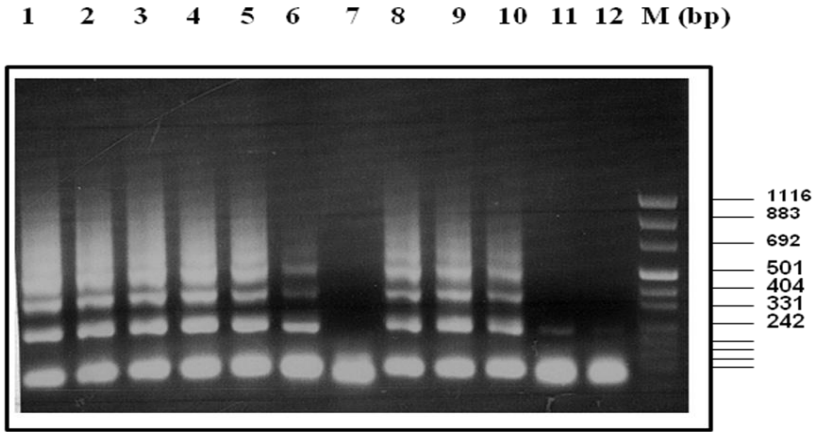


Figure3.1: The newly identified repetitive clones (S.h. repeat 24 and S.h.repeat25) for *S. haematobium* group identification. The identified primers were designed from these clones to be employed in PCR assay for *S. haematobium* group identification. Bold fonts indicating the direct and reverse primers.

Each specific PCR assay has its optimal buffer components and this is due to differences in the sequence of the primers or the purity and quality of these synthesized primers. Because this it is recommended to find up the optimal assay conditions for each new PCR assay. Therefore the optimization reactions were carried out for each of the newly designed primers

(77f/77r, and 73f/73r) (Figure 3.2 a and b). For both repeat different optimal buffers could be used to make the DNA amplification.

A)



B)

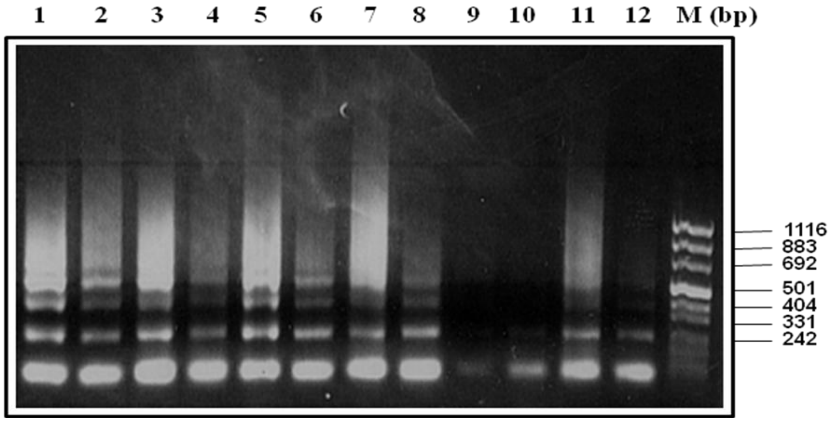


Figure 3.2: PCR assay optimization reaction of S.h.repeat 24 (A) and S.h.repeat25 (B) repeats. 1ng of *S. haematobium* genomic DNA was used per reaction. 1-12 represents the used 12 buffers. M: pUC 8 molecular size marker.

Sensitivity and specificity PCR test for the newly designed primers:

Sensitivity of detection by PCR were determine using different concentration of *S. haematobium* genomic DNA. As shown in figure 3.3a, and b, the lowest detectable concentration of *S. haematobium* DNA can be detected using the newly designed primers (77f/77r, and 73f/73r) based on S.h.repeat 24 and S.h. repeat 25 sequence information was 0.1pg . This result indicates that these primers are very sensitive to detecting *S. haematobium* DNA. While the detection sensitivity of the same primers used in PCR amplification sensitivity test of the other schistosoma animal groups were varied. The PCR assay could detect as low as 1ng of *S. mattheei* genomic DNA, and 0.1ng of *S. bovis* genomic DNA. This result indicates that a high concentration of genomic DNA from (*S. mattheei* and *S. bovis*) is needed to be detected with the PCR assay utilizing these newly designed primer pairs (S.h.repeat 24 and S. h.repeat 25). Based on the obtained results it can be concluded that it is possible to differentiate between the mentioned two groups and *S. haematobium* based on detection sensitivity of the test, which is by itself not sufficient for species differentiation.

Specificity test for the PCR assay was performed using various concentrations of genomic DNA from *S. haematobium* and other schistosomal groups. The amplified PCR products were analyzed by gel electrophoresis (Figure3.3). The results show that 0.1ng and 0.1pg of S. *haematobium* DNA were detectable by both sets of designed primers (77r/ 77f, and 73r/ 73f). On the other hand PCR amplification was also enabled from two other shcistosomal animal groups (*S. bovis* and *S. mattheei*).From the obtained results it is possible to conclude that; it is not possible to differentiate between *S. haematobium* and other schistosomal animal groups utilizing the above mentioned primers designed based on the sequence information of (S.h.repeat 24 and S.h.repeat 25) repetitive clones by simple one step PCR assay. The other tested species (S. *bovis* and *S. mattheei)*, also exhibited PCR

amplification banding pattern similar to that obtained upon using S. *haematobium* genomic DNA. Based on this information it is proven that the newly identified repeats (S.h.repeat 24 and S.h.repeat 25) are found at least in the above three tested schistosomal groups and they have similar distribution in the genome, but they may have different abundances.

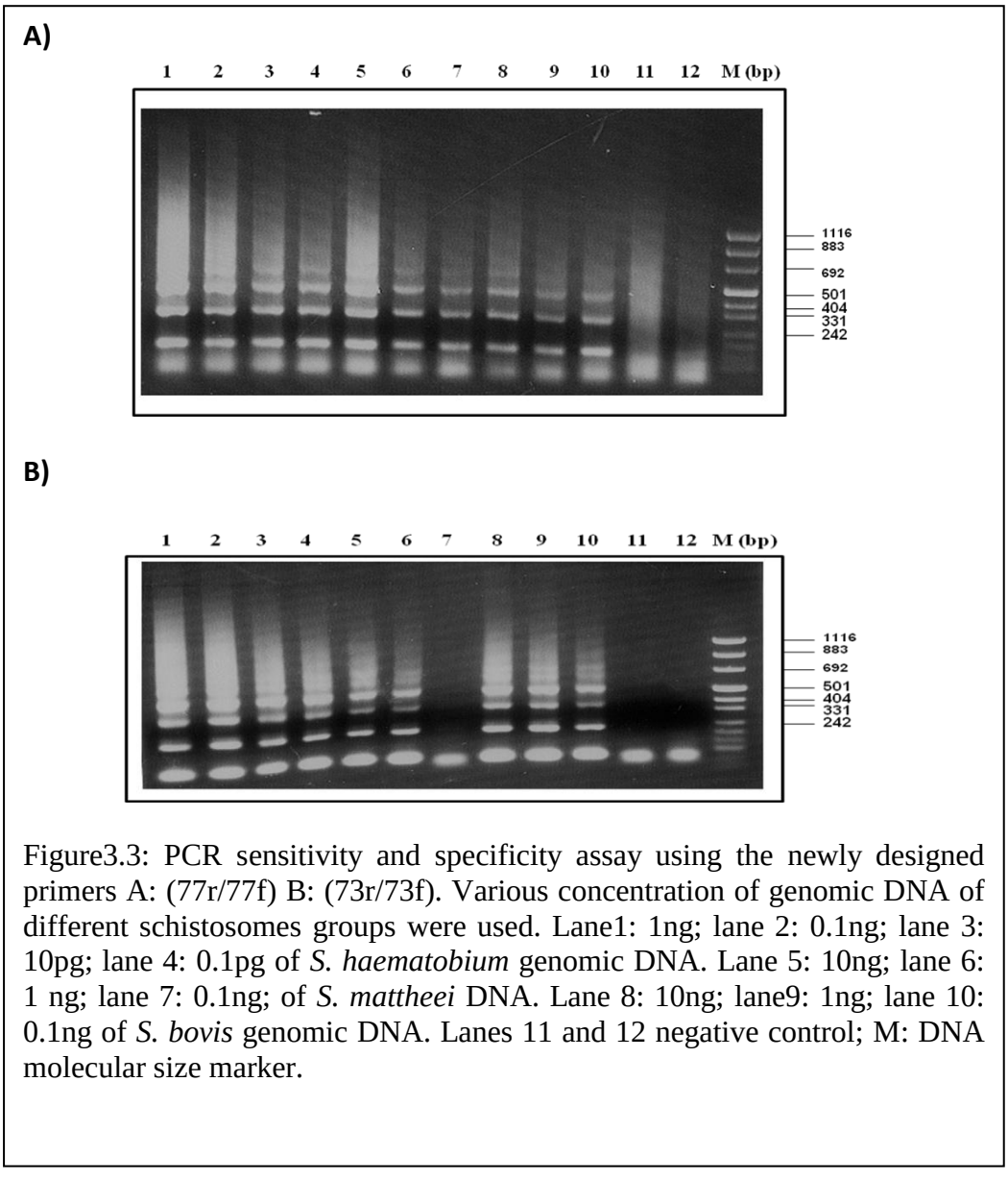


Figure3.3: PCR sensitivity and specificity assay using the newly designed primers A: (77r/77f) B: (73r/73f). Various concentration of genomic DNA of different schistosomes groups were used. Lane1: 1ng; lane 2: 0.1ng; lane 3: 10pg; lane 4: 0.1pg of *S. haematobium* genomic DNA. Lane 5: 10ng; lane 6: 1 ng; lane 7: 0.1ng; of *S. mattheei* DNA. Lane 8: 10ng; lane9: 1ng; lane 10: 0.1ng of *S. bovis* genomic DNA. Lanes 11 and 12 negative control; M: DNA molecular size marker.

Inter-repeat amplification and primers combination PCR assays:

Alternative possibilities were examined to achieve differentiation of *S. haematobium* from other schistosomal animal groups. One of the approaches that were followed was DNA amplification by one step PCR assay for sequences that located between the identified repetitive DNA sequences in *S. haematobium* genome; an approach that was called inter-repeat amplification. Inter-repeat amplification was carried out using different possible combination either (a) from the newly identified candidate repetitive clones (S.h.repeat 24 and S.h.repeat 25) or (b) from other already known repetitive *S. haematobium* clones such as Dra1 repetitive DNA sequence.

A- Inter- repeat combinations based on S.h.repeat 24 and S.h. repeat 25 repetitive clones:

A combination between the reverse and the forward Primer pairs that were designed from the newly identified repetitive DNA sequences of *S. haematobium* genomic DNA (S.h.repeat 24 and S.h.repeat 25) was done. From the designed primers (77r/77f) and (73r/73f); new combinations were created and were used in PCR reactions utilizing *S. haematobium* as the target DNA, the combinations were (77f X73f, 77 f X73r, 77rX73f, and 77rX73r). Taking into account that the correct orientation of these repetitive DNA clones was not determined originally in *S. haematobium* genome, which means from these four combinations only one will be in the right orientation and can results in PCR amplification. So it is very important to perform a PCR assay containing all the possible primer combinations in order to determine the correct orientation. This was done by optimization reaction using six different PCR buffers and *S. haematobium* total genomic DNA as a template (Figure 3.4).

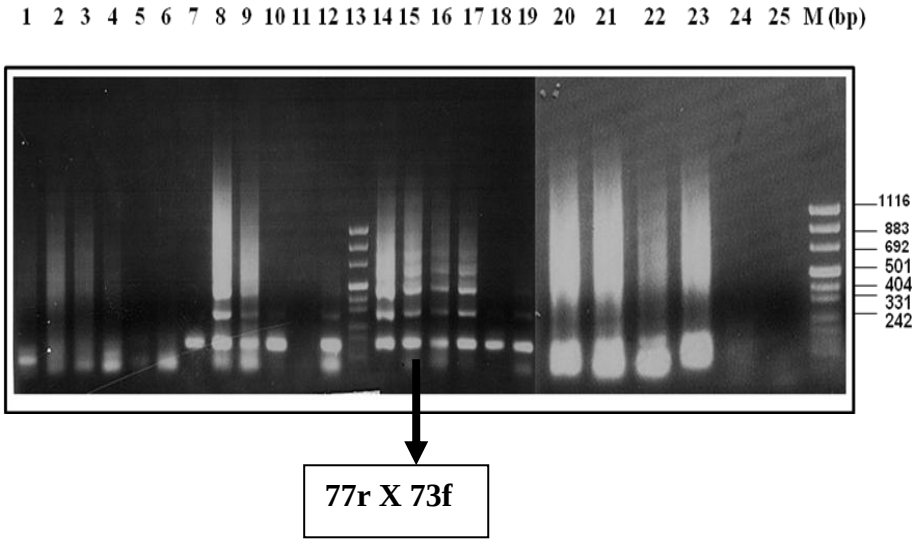


Figure 3.4: PCR assay optimization for different primer combinations obtained from S.h.repeat 24 and S.h.repeat 25 repetitive clones. Six different optimizations buffers were used (1, 3, 4, 5, 10, and 12). Lanes 1-6: 77fX73f primers; lanes 7-12: 77fX73r primers; lanes 14-19: 77rX73f primers; and lanes 20-25: 77rX73r primers; lanes13and 26: M: DNA molecular size markers.

From figure 3.4 it can clearly seen that the third primer combination (77r / 73f) gave the clearest PCR amplification represented by the clear banding pattern on agarose gel electrophoresis. So, inter-repeat amplification can be carried out using this primer pair.

PCR sensitivity and specificity using 77r x73f combined primers.

Various concentration of *S.haematobium* genomic DNA and other schistosomal animal groups were tested by simple one step PCR assay using the following primer combination (77r/77f). The amplified products from all schistosomal groups were analyzed by agarose gel electrophoresis to see the PCR banding pattern. Results shown in figure 3.5 did not show any specific amplification for *S.haematobium* that can be used in the parasite differentiation. All the used sources of template DNA produced the same PCR pattern.

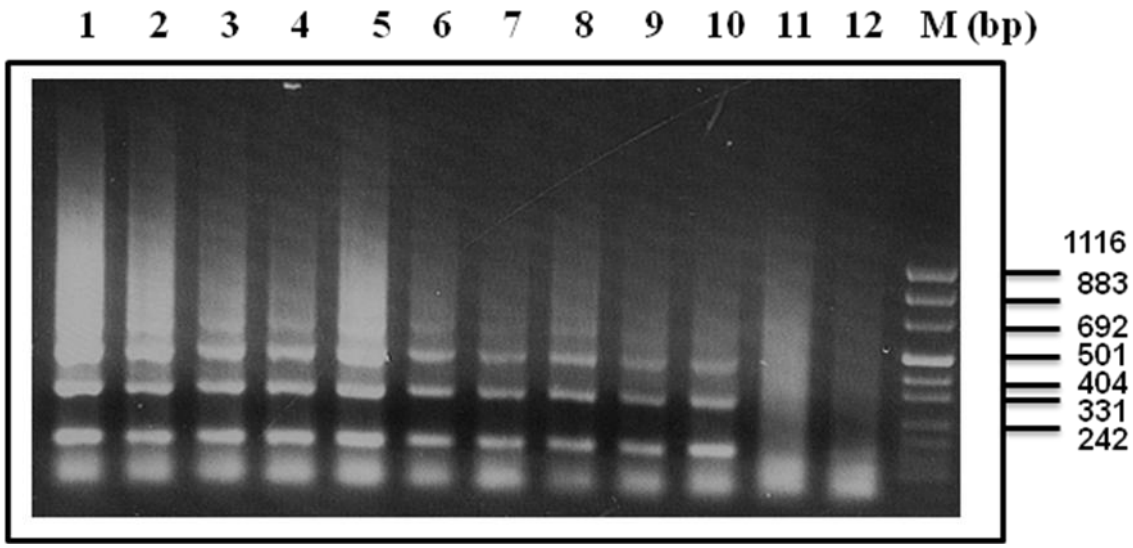


Figure3.5: PCR sensitivity and specificity assay using 77r x73f combined primers**.** Various concentration of genomic DNA of different schistosomes species groups genomic DNA. Lanes 1-4:1ng-0.1pg ;of *S.haematobium* DNA .Lane 5: 10ng; lane 6: 1 ng; lane 7: 0.1ng; of *S. mattheei* DNA. Lane 8: 10ng; lane9: 1ng; lane 10: 0.1ng of *S. bovis* genomic DNA. Lanes 11 and 12 negative control; M: DNA molecular size marker.

B- Primers combinations based on *S.h.*repeat 24, *S.h*. repeat 25 and Dra1 repetitive DNA clones.

In this approach primer combinations was chosen from the newly designed primers base on the identified repetitive DNA clones (S.h.repeat 24 and S.h.repeat 25) known as (77r/77f) and (73r/73f) from one side and on the other side ,already known primers(4297and 4296) were used .These are based on *S. haematobium* DraI repeat (42).*S. haematobium* DraI is a highly abundant repeated sequence that is tandemly arranged in the genome of *S. haematobium*, and in the other schsitosomal animal groups. This sequence was used in one step PCR test to identify *S. haematobium* in infected snails. Although it has very high sensitivity,DraI PCR assay cannot differentiate between *S. haematobium* and other schistosomal animal groups. DraI repetitive sequence was also abundant in the genomes of other *S. haematobium* related terminal spined animal groups.

Optimization reactions were carried out with different primers combinations as follow:

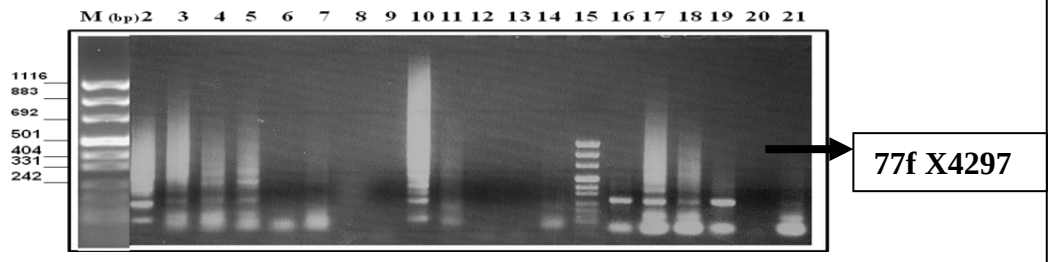
1) 77f x4296. 2) 77r x4296. 3) 77fx4297.

4) 77r x4297. 5) 73f x4296. 6) 73r x4296.

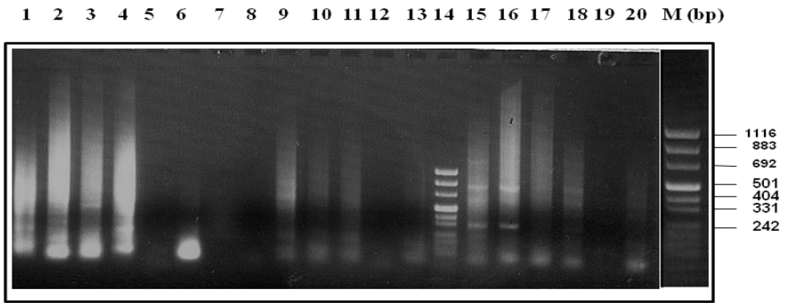
7) 73fx4297. 8) 73r x4297.

The PCR reaction was performed using these combinations and total genomic DNA of *S. haematobium* as target DNA. Six different optimization buffers were tested (buffer 1, 3, 4, 5, 10, 12). Figure 3.6 (a, b, and c) summarizes the obtained results using these different combinations. Only clear PCR product indicated by its clear banding pattern was obtained with the following pairs of primers ((77f /4297) and (73f /4297). These primers were tested for their sensitivity detection and for their specificity to *S. haematobium* as indicated by the following section.

A)



B)



C)

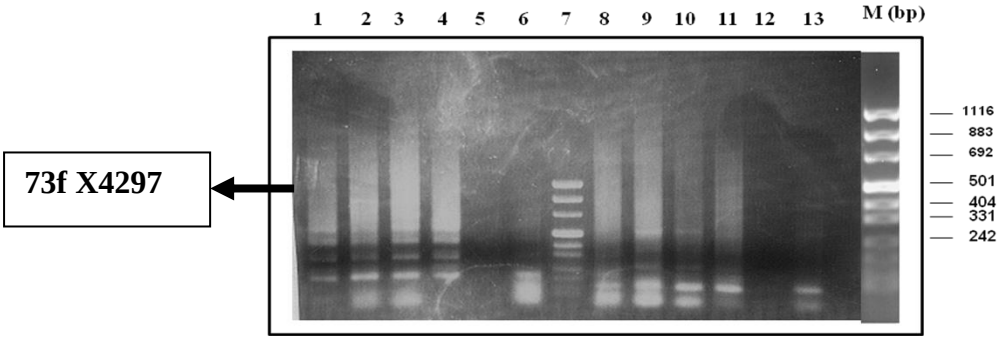


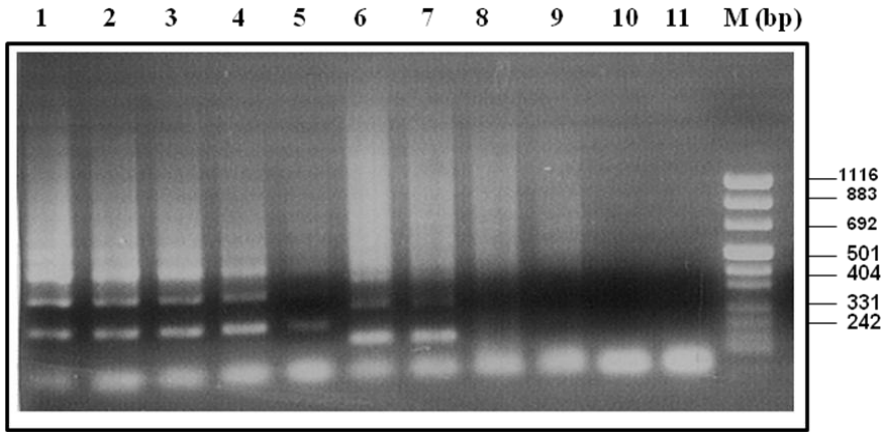
Figure 3.6.A: Optimization detection of the second combination reaction of S.h.repeat 24, S.h.repeat 25, and Dra1 primer pair. Lanes 2-7: 77f x4296; lanes 8-14:77r x4296; M; molecular size DNA; lanes 16-21: 77fx4297. B) Lanes 1-6:77r x4297; lanes 7-13: 73f x4296; M: DNA molecular size DNA; lanes 15-20:73r x4296 .C) lanes 1-6:73fx4297; lanes 8-13:73r x4297; lane7 and lane 15: M DNA molecular size marker.

Differential identification of *S.haematobium* by PCR: (sensitivity and specificity studies).

Based on the results obtained from previous PCR reactions the primers combinations of 77f/4297 and 73f/4297 were selected as candidate pairs to be used in one single step PCR. The results shown in figure 3.7a, represents PCR amplification reaction of different concentration of *S. haematobium* genomic DNA using the first pairs of primers 73fx4297. This newly combination can amplify at least 0.1pg of *S. haematobium* genomic DNA, indicating the high sensitivity of this PCR assay for amplifying *S. haematobium* genomic DNA. PCR assay using the same primers with other schistosomal animal groups was shown to be non sensitive, since it only could detect as low as 1ng of *S. bovis* genomic DNA (1000 times less sensitive as compared to *S. haematobium* DNA amplification), while for *S. mattheei* no DNA amplification was exhibited with 1ng and even with 10ng genomic DNA.

Sensitivity of DNA detection by PCR was determined using different concentrations of schistosoma genomic DNA utilizing the new primers combination (77fx4297) as shown in (Figure 3.7b). The obtained results reflects a low sensitivity test as revealed by low DNA amplification, since as low as 1ng of S. *haematobium* DNA was detected by this PCR assay and only 10ng of *S. bovis* and *S. mattheei* genomic DNA can be detected by the same PCR assay. The test was not approved for further analysis steps and this due to its low sensitivity and its non-specific amplification for S. *haematobium* DNA.

A)



B)

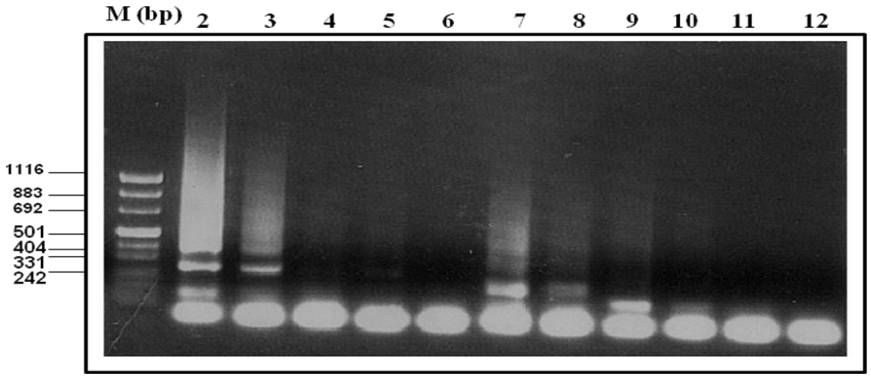


Figure 3.7: PCR assay sensitivity test for the new primers combinations. A) PCR detection using (73f/4297) combined primers. Different concentrations of genomic DNA from different species were used: Lanes 1-5: 1ng-0.1pg of *S. haematobium*; lanes 6 and 7: 10ng and 1ng of *S. bovis;* lanes 8 and 9: 10ng and 1ng of *S. mattheei*; lanes 10 and 11: negative control; M: DNA molecular size marker.

B) PCR detection using (77f/4297) combined primers. Lanes 2-6: 1ng-0.1pg of *S. haematobium*; lanes 7 and 8: 10ng and 1ng of *S. bovis;* lanes 9 and 10: 10ng and 1ng of *S. mattheei*; lanes 11and 12: negative control; M: DNA molecular size marker.

PCR specificity of new primer combination (73f/4297):

The primer combination (73f/4297) appears to be sensitive enough for *S. haematobium* DNA amplification. In this section the specificity of this PCR assay was further tested with different genomic DNA (*S. haematobium* DNA and other schistosomal animal groups). Various concentrations of genomic DNA of different schistosoma groups were analyzed by one step PCR assay. Results shown in Figure 3.8, indicate that 0.1pg of *S. haematobium* can be clearly detected, while there is DNA amplification was noticed with 10ng-10pg of *S. mattheei*, *S.curassoni*, *S. intercalatum* , and *S. margebowiei* genomic DNA, but this DNA amplification have different banding pattern that could be clearly differentiate from that of *S. haematobium*. On the other hand, the DNA amplification banding pattern using *S. bovis* genomic DNA was different compared to that obtained by the same PCR assay using *S. haematobium* genomic DNA as a template. The obtained banding pattern obtained from *S. haematobium* DNA starting from a larger band size (about 150bp) compared to the pattern obtained from *S. bovis* DNA in which the first band starts from about 100bp). This PCR test can be considered as quantitative and qualitative *S. haematobium* differential test.

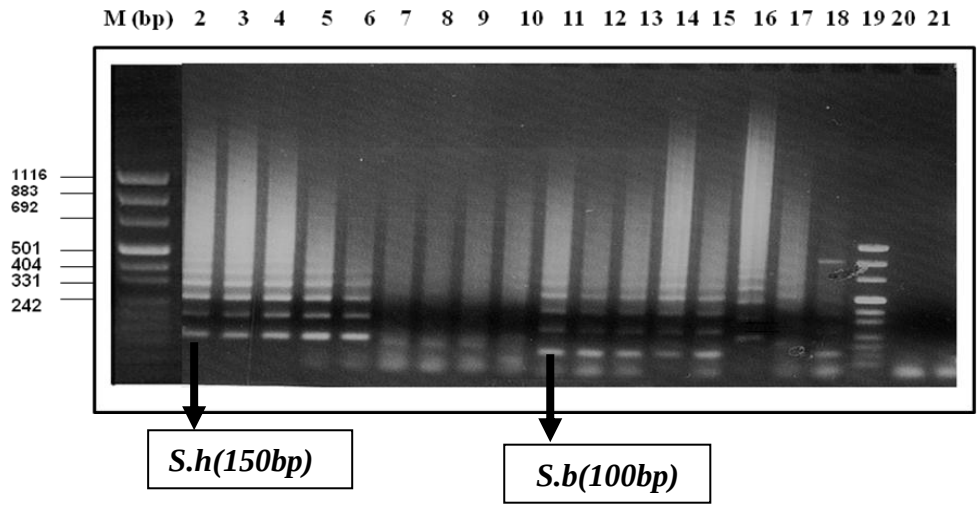


Figure 3.8: PCR Specificity of primer combination (73f/4297). Various concentrations of genomic DNA from different groups were analyzed. Lanes2-6: 1ng-0.1pg of *S. haematobium* DNA; Lanes 7-10: 10ng-10pg of *S. mattheei* DNA; lanes 11-15: 1ng-5ng of *S. bovis* DNA; lanes 16: 10ng of *S. curassoni* DNA; lane 17: 10ng of *S. intercalatum* DNA; lane 18: 10ng of *S. margebowiei* DNA; lane 19: M: DNA molecular marker; lanes 20-21: negative control.

For further clarification between the band size obtained after DNA amplification using (73f/4297), a separate run was performed using amplified DNA amplified products from *S. haematobium* and *S. bovis* (Figure 3.9)

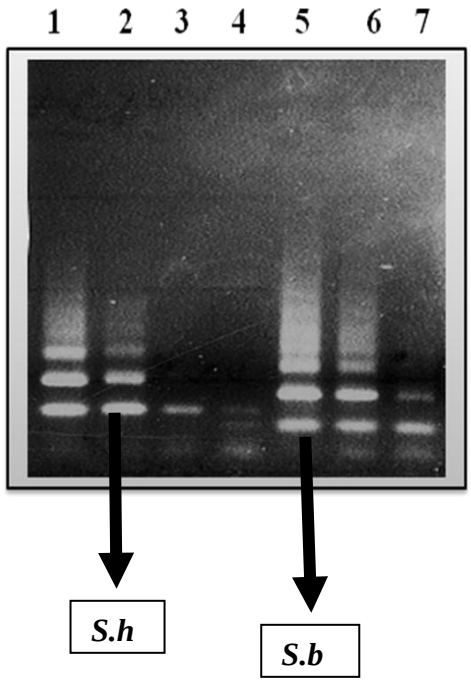


Figure3.9: PCR amplification from *S.haematobium* and *S. bovis* using 73f/4297 primers. Various concentrations of genomic DNA from *S.haematobium* and *S. bovis* were analyzed.lane1-4:1ng-0.1pg of *S. haematobium* DNA; Lanes 5-7: 10ng-10pg of *S.*bovis.

Chapter Four

DISCUSSION

Human schistosomiasis is a parasitic trematode infection that causes a major health burden in the developing world (58). *Schistosoma haematobium*, *S. mansoni*, and *S. japonicum* are the most common species involved in human disease. More than 600 million people live in schistosomiasis transmission zones, and an estimated 200 million people are infected (59). Contamination of water bodies by human excreta that contain the schistosomes eggs facilitate transmission of the disease by the aid of an intermediate snail host in which the parasite can propagate and produce the infective cercariae stage (60). A recent increase has been noticed in parasites range considering the new implementation methods of agricultural and water management. Several living conditions that favor transmission, such as daily use of streams and ponds for washing, laundering, swimming and irrigating, and through lack of sanitary systems that can prevent water contamination (60).

Schistosomiasis is considered as a chronic inflammatory disease that can last for decades. To date, most control efforts have focused on treating and preventing the advanced forms of the disease (62). Schistosomiasis treatment programs are a first step in reducing the global burden of Schistosoma-related disease, yet such programs may not significantly change parasite transmission especially in high-risk areas. In low endemic areas it is very difficult to assess the transmission changes from snails to human. Novel molecular tools that are sensitive for transmission monitoring will enable the gathering of much more information on schistosome propagation in an area. Such information will be very valuable in planning an effective transmission control of schistosomiasis morbidity (57).

The overall aim of the proposed study was to develop molecular tools that are sensitive and specific for *S. haematobium* detection in its snail intermediate host. The capacity of the developed molecular tool will allow efficient monitoring of schistosomiasis

transmission and it will be useful for large-scale monitoring of residual parasite transmission that is very difficult to detect by conventional methods.

Traditionally measuring Schistosoma transmission was performed through measuring the number and percentage of snails that are infected and shedding cercariae, or by measurement of cercariae numbers (cercariometry) in the water at human contact sites (30). Due to technical problems in practical implementation and data analysis, cercariometry has not been routinely applied. While, measurement of snail infection rates remains the golden test of monitoring schistosomiasis transmission rate (60). However, the rate of cercaria-shedding snails can be quite low (< 1%), even in areas of high transmission, so that large samples and extensive field monitoring have been required to detect statistically significant changes in transmission potential over time. Using PCR technology as a tool of molecular monitoring of prepatent infection in snails enables very early detection of contamination, and can be used for accurate and sensitive measurement of local human-to-snail schistosome transmission. The advantages of molecular monitoring of prepatent *Schistosoma* infection can be summarized in the following points:

1- Greater specificity of schistosome species identification

2- Greater sensitivity for detection of snail infection

3- Enhanced sensitivity for detection of area contamination as human-to-snail transmission

4- Minimal dependence on participation by local population

PCR- based assays are much less labor-intensive, tedious, and often more practical than the traditional method of snail dissection for large scale screening in endemic areas. PCR-based methods have the additional advantage that they can be used to detect single

parasite in a pool of field collected snails. This can make their field application more cost-effective than dissection, particularly in areas where schistosomiasis prevalence becomes very low after effective intervention. By using them to screen pools of field collected snails and water from potential transmission sites, such methods can provide useful information on active transmission in a particular setting, enabling remaining "hot spot" sites to be identified and transmission in areas where infection is rare (as might be expected after an effective intervention) to be monitored.

For epidemiological studies, transmission control and for schistosomiasis eradication in low endemic countries it is very important to have a tool to confirm the specific elimination of local human parasite transmission. For this, the development of suitable tools for monitoring transmission is urgently required. In the past, monitoring the impact of schistosomiasis control has often been based on counting schistosome eggs in human excreta in order to determine prevalence, intensity, and incidence of infection (53). Also, the impact of control programs on Schistosoma transmission potential can be determined by measuring the number and percentage of snails that are infected and shedding cercariae, or by measurement of cercariae numbers (cercariometry) in the water at human contact sites. Due to technical problems in practical implementation and data analysis, cercariometry has not been routinely applied (60). However, for large-scale control purposes, a more rapid and inexpensive monitoring method is needed and this can be specifically achieved by molecular monitoring.

The ability to detect snail *S. mansoni* infection from the first day after miracidial penetration was established by using PCR amplification of the Sm1-7 sequence in *S. mansoni*-exposed snails (63). For similar detection of *S. haematobium*, a *S. haematobium* genomic sequence, DraI was identified, that is a 121 bp long tandemly repeated sequence

comprising about 15% of the genome which, when amplified by PCR, enabled identification of 1 fg of *S. haematobium* DNA or a single cercaria (43). However, the DraI sequence was also found in the genomes of closely related schistosomes species belonging to the *S. haematobium* group, including *S. bovis, S. intercalatum*, and *S. matheei*, which all have terminally-spined ova (43).

Identification of repetitive DNA sequences were possible by the aid of bioinformatics searching tools in which advanced computer technology is applied to the manage biological information. Specifically, bioinformatics is the science of developing computer databases and algorithms to facilitate and expedite biological research (64). Identification of the repetitive sequence was performed by the aid of Gene-Bank information, Blast search, and nucleotide-nucleotide similarities.

The general aims of bioinformatics is the ability to organize data in a way that allows researchers to access existing information, to submit new entries as they are produced, and to develop tools and resources that aid in the analysis of data especially the biological data such as DNA analysis. This analysis can be done through different algorithm software such as NCBI that involve different sets of search programs designed for the Windows platform and is used to perform fast similarity searches regardless of whether the query is for protein or DNA. Also comparison of nucleotide sequences in a database can be performed. Algorithms were developed to help researchers rapidly identify similar gene or protein sequences. Such tools are extremely useful for determining whether a newly sequenced piece of DNA was at all similar to sequences already entered in a database. To determine how multiple sequences align and to view their similarities, multiple alignment programs were developed (BCM Luncher, multiple sequence alignment). Such programs helped scientists compare the

sequences of closely related genes or compare the sequence of a particular gene or protein as it appears in several species.

In the present study many newly identified repetitive clones were identified. Some of these repeats are specific for *S. haematobium* and others are specific for *S. haematobium* groups. Many of these clones were genus specific and others were only identified among *S. haematobium* related shcistosomes. Based on the identified repetitive clones it was obvious that it is not possible to identify *S. haematobium* form other terminal spined schistosoma groups using a simple PCR assay. For this reason a new approach of studying the flanking regions around the identified repeats was followed in order to find specific region that can be used specifically to identify *S. haematobium*. In this regard new primers were designed from two different repeats in order to be used in a PCR assay for differential identification of *S. haematobium*. The newly target repeats by themselves are not specific to *S. haematobium*; but the around flanking regions can be used for differential identification of *S. haematobium* since the amplified region has different length when compared with *S. bovis* amplified region. These primers were designed based on newly identified repeated sequence called (S.h.repeat 24 and S.h. repeat 25) obtained from *S. haematobium* genomic library.

The identification of new repetitive DNA clones provided a data that was used to design the new PCR primers, depending on the fact that the repetitive sequence can be used for the diagnostic purpose and species identification especially when these sequences do not code for proteins. Using reverse and forward primers that were designed from single repeat; each in independent PCR assay (77r/77f and 73r/73f) had shown to be very sensitive for the detection of *S. haematobium* DNA but they were not specific since they amplified *S. bovis*

and *S mattaeei* DNA under the same PCR assay. This result reflects the fact that the both identified repetitive clones (S.h.repeat 24 and S.h. repeat 25) were also found in the genome of other related groups of *S. haematobium*. Because these primer pairs were failed by themselves to differentiate *S. haematobium*, an inter-repeat amplification approach was used, which is depending on DNA amplification of sequences that are located between the two repetitive DNA clones. This DNA amplification approach requires testing all possible combinations between the designed reverse and forward primers that were obtained from these repetitive DNA sequences.

The first PCR combination reactions included all the direct and reverse primers that were designed from the newly identified repetitive DNA sequence (S.h.repeat 24 and S.h. repeat 25), while the second PCR combination reactions included all reverse and direct primers that were designed from the newly identified repetitive DNA clones and the already identified *S. haematobium* Dra1 primers (43). The end result of these PCR combination reactions produced (73f /4297) as the most potential primer pair candidates to be used for differential identification of *S. haematobium* (Figure3.8). The tests that were carried out in this study were mainly concentrated on the examination of *S. bovis* and S. *mattaheei*; both are *schistosoma* species that belong to *S. haematobium* related groups. *S. bovis* and *S. mattaheei* are the most widespread and prevalent species compared to the other *S. haematobium* related groups, thus they have the widest potential geographical overlap with *S. haematobium*. The other *S. haematobium* related species that were examined, are less widely distributed, they include *S. intercalatum* which is found in southern Africa, *S. margebowiei* in a relatively small areas in west and southern central Africa, and *S. curassoni* found primarily in western Africa.

The present described PCR assay used for differential identification of *S. haematobium* from other related schistosomes, amplifies genomic DNA extracted from *S. bovis* using 73f/4297 primer pairs in PCR reaction. It is very important to know that there is a major difference in the size of the amplified DNA segment between *S.haematobium* and *S. bovis* and this difference can be detected by simple PCR reaction, followed by analysis over gel electrophoresis. Also, these primers were not sensitive for *S. bovis* genomic DNA amplification (only 5ng can be detected which is 1000 less sensitive compared to *S. haematobium* genomic DNA amplification).

Based on the obtained results *S. haematobium* infected snails can be clearly differentiated from other snails infected by the most prevalent *S. haematobium* related animal schistosomes in Africa, i.e: *S. bovis* and *S. mattheei*, in areas where it is found alongside with *S. haematobium*. This result also will allow an easier monitoring of snail transmission. In addition several other repetitive clones were identified that could be a potential for specific identification of other *S. haematobium* related groups (Table 3.1). So for *S.matthee, S.curassoni, S.intercalatum, S.margebowiei* and *S.bovis*, different repetitive clones were identified by the aid of bioinformatics tools which can be used to design primers for each gruops to carry out a simple PCR reaction for the thier differentiation.

Conclusion

The view of the social and medical importance of schistosomiasis is changing with the greater appreciation of its day to day impact on public health and using them to screen pools of wild snails and water from potential transmission sites, such methods can provide useful information on active transmission in a particular setting, enabling remaining "hot spot" sites to be identified and transmission in areas where infection is rare disability. The well known of the infected cercariae , will avoid humans from being infected or it can be able to detect the infection in its early stage , thus will provide from reaching the advance stage that's really associated with patient death, such as liver fibrosis , urinary tract obstruction, intestinal bleeding , anemia and cancer.

The new molecular amplification technologies offer a positive step towards community monitoring of local water contamination by schistosome eggs and transmission to snails. Use of these sensitive and specific techniques will provide essential data for the development of the next generation of schistosomiasis control, based on reduction and ultimate elimination of local transmission. In addition to that, the discovery of the new inexpensive molecular diagnosis and identification tools for field programs in endemic regions is an issue relevant to large scale monitoring of control programs for schistosomiais. Such development in molecular technology will lead to get better detection sensitivity, operational simplicity and low cost, thus will extend the period of control program for long time.

REFERENCES

1- Viroj Wiwanitkit1 (2005) Clinical report on urinary schistosomiasis in the tropical Asia: an overview of Continuing Medical Education, 21(4):499-501.

2- Handzel T. S., Karanja D., Addiss D., Hightower A., Hrosen D., Colley D., Andove J., Slutsker L. (2002) Geographical distribution of schistosomiasis and soil transmitted helminthes in western Kenya: implications for anthelminthic mass treatment. *Am. J. Trop. Med. Hyg.*, 69(3), 318-323.

3- WHO, 2002. Prevention and control of schistosomiasis and soil-transmitted helminthiasis: Report of a WHO expert committee. *World Health Organization Tech Rep Ser* 912: 2-3.

4- WHO, 2007.Division of control of tropical diseases(the control of schistosomiasis). Report of a world health organization expert committee.

5-Hamburger J.,Hoffman O.,Kariuki H.,Muchiri E.,Ouma J., Koech D.,Sturrock R., King C. (2004) large –scale ,polymerase chain reaction-based surveillance of schistosoma haematobium DNA in snails from transmission sites in coastal Kenya:A new tool for studying the dynamics of snail infection . *The American Journal of Tropical Medicine and Hygiene* . 71(6),765-773.

6- Barbour AD.(1985) The importance of age and water contact patterns in relation to *Schistosoma haematobium* infection. *Trans R Soc Trop Med Hyg. 79,151*–153.

7- Chandiwana SK., Woolhouse ME. (1991) Heterogeneities in water contact patterns and the epidemiology of *Schistosoma haematobium*. *Parasitology .103,* 363–370.

8- WHO,1998 .Report of the WHO informal concsulation on schistosomiasis control. Geneva, World Health Organization Technical Report.

9- Brookeri S., Hays S., Issae W.,Hall A., Kihamia C., Lwambos N., Wint W,Rogers D., Bundy D.(2001) Predicting the distribution of urinary schistosomiasis in Tanzania using satellite sensor data. *Tropical Medicine and International Health*,6 (12), 998-1007.

10- Laamrani H., Mahjour J., Madsen H., Khallaayoune K., Gryseels b.(200) Schistosoma haematobium in Morocco: moving from control to elimination. *Parasitology Today*, 16 (6).

11- Chang W., Gang Pan Z., Hong Sun Y.(2006) Sigmoid colonic carcinoma associated with deposited ova of Schistosoma japonicum : A case report. *World Journal of Gastroenterology*, 12(37): 6077-6079

12- Montes M., White C., Kontoyiannis D. (2004) Symptoms of intestinal schistosomiasis presenting during treatment of large B cellymphoma. Am. J. Trop. Med. Hyg., 71(5), 552-553.

13- Satayathum S., Muchiri E., Ouma J., Whalen C., King C.(2006)Factor affecting infection or reinfection with Schistosoma haematobium in coastal Kenya: Survival analysis during a nine – year, school –based treatment program. Center for Global Health and Diseases.

14- Durand F.,Brion J.,Terrier N., Pinel C.,Pelloux H.(2004)Funiculitis due to schistosoma haematobium diagnosis using parasitologic analysis of semen. Am. J. Trop. Med. Hyg., 70(1), 46-47.

15-Neal P.(2004) Clinical overview of Shistosomiasis– An Unusual Cause of Ureteral Obstruction . Clinical Medicine & Research , 2(4), 216-227.

16 Lehman J.,Zoheir Farid J.,Bassily S.,kent D.(2006) Renal function in urinary schistosomiasis. The American Society of Tropical Medicine and Hygiene, 19(6), 1001-1006.

17- Cesmeli E.,Vogelaers D.,Voet D.,Peleman R.,Kunnen M., Afschrift M.(1997) Ultrasound and CT changes of liver parenchyma in acute schistosomiasis.The British Journal of Radiology ,70,758-760.

18- *Nguyen L.,Estrella J., Jett E., Grunvald E.,Nicholson L., Levin D.(2000) Acute* Schistosomiasis in Nonimmune Travelers: Chest CT Findings in 10 Patients.A merican Journal of R oentgenology,186,1300-1303.

19- Blute R., Oliva E.(2000) Case 31-2000-A32-year –old man with a lesion of the urinary bladder.the new England journal of medicine , 343,1105-1111.

20- Liping C.,Viana L., MeloGarcia T.,Rabello A.,Katz N.(1996) TheSerological differentiation of acute and chronic schistosomiasis japonica using IgA Antibody to egg Antigen. Hubei Institute of Schistosomiasis Control, *91*(6),751-754.

21- Nguyen L., Estrella J., Jett E., L.Grunvald E., Nicholson L.,Levin D.(2006) Acute schistosomiasis in nonimmune travelers:chest CT findings in 10 patients . *AJR*, 186, 1300-1303.

22- Kahama A.,Nibbeling H., Vanzeyl R., Vennervald B., Ouma J., Deelder A.(1998) Detection and quantification of soluble egg antigen in urine of Schistosoma haematobium – infected children from Kenya. *Am. J. Trop. Med. Hyg.*, 59(5), 769–774.

23- Salah á F., Demerdash á Z., EL Bassiouny G., Ismail á B., El Din S., Mansour M.(2000) A monoclonal antibody against Schistosoma *haematobium soluble egg antigen: efficacy for diagnosis and monitoring of cure of S. haematobium infection*. Parasitol Res ,86,74-80.

24- - Kock K., Wolmarans C.(2005). Distribution and habitats of the *Bulinus africanus* species group, snail intermediate hosts of *Schistosoma haematobium*and *S. mattheei* in South Africa. Water SA .,31(1).

.

25 – Kock K., Wolmarans C.(2005). Distribution and habitats of the *Bulinus africanus* species group, snail intermediate hosts of *Schistosoma haematobium*and *S. mattheei* in South Africa. Water SA, 31(1).

26 - Kock K., Wolmarans C., Bornman M.(2004) Distribution and habitats of *Biomphalaria pfeifferi*, snail intermediate host of *Schistosoma mansoni*, in South Africa. Water SA, 30(1).

27- Willingham L.,Hurst M.,O.Bogh h.,Johansen M.,Lindberg R.,Christensen N., Nansen p.(1998) *Schistosoma japonicum* in the pig: the host parasite relationship as influenced by the intensity and duration of experimental infection. *Am. J. Trop. Med. Hyg.*, 58(2):248-256.

28- Rollinson D., Stothard J., Jones C., Lockyer A., Souza C., Noble L.(1998) Molecular charecterisation of intermediate snail hosts and the search for resistance genes.Memorias do instituto oswaldo cruz .on-line, *93*, 111-116.

29- Khallaayoune K., Madsen H., Laamrani H.(1998) Evaluation of three methods to control *Bulinus truncatus,* the intermediate host of *Schistosoma haematobium* in an irrigation scheme,Tessaout-Amont, Morocco. Acta Tropica, 69, 51–63.

30- Kariuki H., Clennon J.,Brady M., Kitron U., Sturrock R., Ouma J., Mungal P., Hoffman O.,Hamburger J., Pellegrini C., Muchiri E., King C.(2004)Distribution patterns and cercarial shedding of Bulius nasutus and other snails in the mesambeni area ,coast proving , Kenya.Am. J. Trop. Med. Hyg., 449-456.

31- Mccutchan T., Simpson A., Mullins J., Sher A., Nash T., Lewis F., Richards C.(1984) Differentiation of Schistosmes by species , strain ,and sex by using cloned DNA markers.*Proc.Natl,acad.SCI.*usa., 81,889-893.

32- Garcia L. (2001) Diagnostic Medical Parasitology .4th edition, .Lsg and Associates Santa Monica in California.

33- Loker ES.(1983) A comparative study of the life-histories of mammalian schistosomes. *Parasitology, 87,* 343–369.

34- Rollinson D., Stothard JR., Southgate VR.(2001) Interactions between intermediate snail hosts of the genus *Bulinus* and schistosomes of the *Schistosoma haematobium* group. *Parasitology, 123* , S245–S260.

35- Webster B., Southgate V.(2003) Compatibility of *Schistosoma haematobium*, *S. intercalatum* and their hybrids with Bulinus truncatus and *B. forskalii*. Parasitology, 127, 231–242.

36- Barber K., Mkoji G., Loker E.(2000) PCR –RFLP analysis of the ITS2 region to identify *Schistosoma haematobium* and S.bovis from Kenya.*Am. J. Trop. Med. Hyg.*, 62(4), 434–440.

37- Jourdane J., Southgate V., Pages J., Durand P., Tchuente L.(2001) Recent studies on Schistosoma intercalatum: taxonomic status, puzzling distribution and transmission foci revisited. Mem Inst Oswaldo Cruz, Rio de Janeiro, 96, 45-48.

38- Mouchet F., Theron A., Bremond P., Sellin B.(1998) Pattern of cercarial emergence of *schistosoma curassoni* from Niger and comparision with three sympatric species of schistosomes. Parasitology, 78, 61-63.

39- Tchuente L.,Southgate V.,Jourdane J.,Kaukas A.,Vercruysse J.(1997) Hybridization between the diegenease *Schistosoma haematobium* and *S. mattheei*: viability of hybrids and their development in sheep.Systematic Parasitology 36,123-131.

40- Navarrete S., Rollinson D., Agnew A.(1994) Cross-protection between species of the *Schistosoma haematobium* group induced by vaccination with irradiated parasites. Parasite Immunology 16 (1), 19–25.

41- Walker T., Rollinson D.,Simpson A.(1986) Differentiation of *schistosoma haematobium* from related species using cloned ribosomal RNA gene probes.Molecular and biochemical parasitology 20, 123-131.

42- Kaukas A., Neto E., Simpson A., Souuthgate V., Rollinson D.(1994) A phylogenetic analysis of schistosoma haematobium group species based on randomly amplified polymorphic DNA.I nternational journal for parasitology , 24(2), 285-290.

43- Abbasi I., King C.,Sturrock R., Kariuki C., Muchiri E., Hamburger J.(2007) Differentiation of *schistosoma haematobium* from related schistosomes by PCR amplifying an inter-repeat sequence. *The American society of tropical medicine and hygiene*.

44- Mafe M.,Stamm T., Utzinger J.,Ngoran E.(2000) Control of Urinary Schistosomiasis :an investigation into the effective use of questionnaires to identify high – risk

communities and individuals in Niger State , Nigeria.Tropical Medicine and I nternational Health, 5, 53-63.

45- Doenhoff M., Chiodini L. (2004)Specific and sensitive diagnosis of Schistosome infection : can it be done with antibodies. Trends in Parasitology, 20, 35-39.

46- Kariuki C.,Clennon J., Brady M., Kitron U.,Sturrock R., Ouma J., Ndzovu S., Mungai P., Humburger J., Pellegrini C., Muchiri E., King C.(2004) Distribution patterns and cercarial shedding of Bulinus Nasutus and other snails in the Msambweni area ,coast province ,Kenya. Am. J. Trop. Med. Hyg., 70(4), 449-456.

47- Hamburger J., Abbasi I., Ramzy R., Jourdan J., Ruppel A. (2001) Polymerase chain reaction assay based on highly repeated sequence of *Schistosoma Haematobium*: A potential tool for monitoring Schistosome- water. *The American society of tropical medicine and hygiene*, 65(6), 907-911.

48- Brindley P, (2005) The molecular biology of schistosomes.Trends in Parasitology; vol.21 No.11.

49- Bowles J., Blair D., Mcmanus D.(1995) A Molecular phylogeny of The human schistosomes .*Molecular phylogenetics and evolution* , 4 (2) , 103-109.

50- Lodish H., Baltimore D., Berk A., Zipuosky S., Matsudaira P.,Darnell J.(1995)Molecular Cell Biology, 3rd ed.

51- Indiana Institute for Molecular and Cellular Biology (1998) Bioinformatics science.

52- WHO, 2002.Schistosomiasis burden and epidemiological trends: Report of a WHO expert committee. World Health Organization.

53- Sturrock R, 2001.Schistosomiasis epidemiology and control: How did we get here and where should we go? Mem Inst Oswaldo Cruz, Rio de Janeiro, 96, 17-27.

54- Da Silva I., Thiengo R., Conceicao M.,Rey L., Lenzi H.,Filho E., Riberio P.(2005) Therapeutic failure of praziquantel in the treatment of *Schistosoma haematobium* infection in Brazilians returning from Africa . Mem. Inst. Oswaldo Cruz vol.100 no.4 .

55- Chan M.,Montresor A., Savioli L., Bundy D.(1999) Planning chemotherapy based schistosomiasis control:validation of a mathematical model using data on

Schistosoma haematobium from Pemba, Tanzania. *Epidemiol*. *Infec*, 123, 487-497.

56- WHO, 2002.Questionnaires for rapid screening of schistosomiasis in sub-Saharan Africa: Report of a WHO expert committee. World Health Organization.

57- King C.,Sturrock R., Kariuki H.,Hamburger J.(2006)Transmission control for schistosomiasis -Why it matters now.Trends in Parasitology, 22(12).

58- Souza J., Morias C., Aroucha M., Miranda D., Barbosa C.Domingues A.,Jonior L.,Montenegro S.(2007) Treatment of human acute schistosomiasis with oxamniquine induces an increase in interferon-γ response to
Schistosoma mansoni antigens. Mem Inst Oswaldo Cruz, Rio de Janeiro, 102(2): 225-228.

59- WHO.2005. Report of the Scientific Working Group meeting on Schistosomiasis world health organization technical report.

60- WHO .2001.Report of the WHO informal consultation on Schistosomiasis in low transmission areas: Control strategies and criteria for elimination .World health organization.

61- Chipaux JP., Boulanger D., Bremond P., Campagne G.,Vera C., Sellin B.(1997) The WHO Collaborating Centre for Research and Control of schistosomiasis at Niamey, Niger. *Mem Inst Oswaldo Cruz*, 92: 725-728.

62- Noya B.,Ruiz R.,Losada S.,Colmenares C.,Contreras R., Cesari R., Noya O.(2007) Detection of schistosomiasis cases in low-transmission areas based on coprologic and serologic criteria .The Venezuelan experience. Acta Tropica, 103, 41–49

63-Abath F.,Gomes A., Melo F., Barbosa C.,Werkhauser R.(2006) Molecular approaches for the detection of *Schistosoma mansoni*: possible applications in the detection of snail infection, monitoring of transmission sites, and diagnosis of human infection. *Mem Inst Oswaldo Cruz*, Rio de Janeiro,. *101*(I): 145-148.

64- Fox J.(2002)What is Bioinformatics science . UBIC Bioinformatic center.

Appendix A: Newly identified repetitive clones in *Schistosoma haematobium* and other related species

A1:-*Schistosoma haematobium*.

***S.h* repeat 1:**

CTAAATGCCT AGCAACGGGG CTCCTGACCA CGTGACCTTC AGGTGCCGAG
TCCACAAACG AGACGTGCAC ACCCAGCTGA AACTTAAAGG AATTGACGGA
AGGGCACCAC **CAGGAGTGGA GCCTGCGGTT TAATTCGACT CAACACGGGA**
AAACTCACCC GGCCCGGACA CTGTGAGGAT TGACAGATTG ATA

***S.h* repeat.2**

ACTCTAAAAG AGCGACGAGA TATTCTTACA CAACCACGTN ATCGGTAGCT
GATCTATGAT AGTGATCACA TCCGAGATCC TGGTTTGAGA TGCATAAGGG
AAGACACCAG ATAGCACCTT TGGATATGAC TGTGCCGATN GTTAGTGCTT
GCTCGAAACA GGTTGTAGTC TACGCAAAGT TGCAAGAGAC ANGTCATCGT
TATTTAGACC TT**CGACTAAC AAGTCCCCAT CGGCCTCCTA**
TTCTGTGCCT AGACGACCTA CCTGTGCATT CAAGTNTNCA AAAGACTCTC
GCTACTATTA CAAAATNTGT CGT

***S.h* repeat.3**

ACTCTAAAAG AGCGACAGTT TCT**TACACAA CCACGTCATC GGTAGCTGAT**
CTTGATGTGA TCAATCCGAG TCCTGGTTTG AGATGCAAAG GGAAGACACC
AGATAGCACT TTGGATATGA CTGTGCCGAT AGTTAGTGCT TGCTCGAAAC
AGGTTGTAGT CTACGCAAAG TTGCAAGAGA CGGTCATCGT TATTTGACCT
TCGACTAACA AGTCCCCATC GGCCTCCTAA AAGACTCTCT TCTGTGCCTA
GACGACCTAC CTGTGCATTC AAGTCTCCAG CTACTATTAC AAAATCTGTC

GT

***S.h* repeat.4**

ATTGGGTACC GGGCCCCCCC TCGagGTCGA CGGTATCGAT AAGCTTGATA
TCGAATTCCT **GCAGCCCACT TTCATTCAAG TTGATCTTCT CTCCGAAGCT**
CAAACCCAGT GGGGGATCCA CTAGTTCTAG AGCGGCCGCC ACCGCGGTGG
AGCTCCAGCT

***S.h* repeat.5**

AAGCTGGAGC TCCACCGCGG TGGCGGCCGC TCTAGAACTA GTGGATCATC
GGTTGGTAAT CAGTGTCATG TATATCAGTT CCTCCAACTA CCATCTTATC
TCAACATAGT ACACTCAATG TTGAATCACC TAAACTGGTA ACCGTATTAC
AACTTGATCA TCATCATCAT CATCATCATC ATCCAAATCA TCATCATCAT

CATCATCATC CAAATCATCA TCTTCATCAT CATCATCATC ATCATCATCA
TCATCATCAT CATCATCATC ATCATCATCA TCATCATTAT CATCATCATC
ATCATCATCA TCATCATGAT CCCCCGGGCT GCAGGAATTC GATATCAAGC
TTATCGAT

***S.h* repeat.6**

TTGGGTACCG GGCCCCCCT CGaGGTCGAC GGTATCGATA AGCTTGATAT
CGAATTCCTG CAGCCCGGGG GATCATGATG ATGATGATGA TGATGATGAT
GATAATGATG ATGATGATGA TGATGATGAT GATGATGATG ATGATGATGA
TGATGATGAT GATGATGAAG ATGATGATTT GGATGATGAT GATGATGATG
ATGATTTGGA TGATGATGAT GATGATGATG ATGATCAAGT TGTAATACGG
TTACCAGTTT AGGTGATTCA ACATTGAGTG TACTATGTTG AGATAAGATG
GTAGTTGGAG GAACTGATAT ACATGACACT GATTACCAAC CGATGATCCA
CTAGTTCTAG AGCGGCCGCC ACCGCGGTGG AGCTCCAGCT TTTG

***S.h* repeat.7**

CTGTGAATCT GGCAATCCTC ACAGCGNCCG GGCCGGGTGA AGTTTTCCC
GTGAGTCGAA TTAAACCGCA GGCTCCACTC CTGGNGGTGC CCTTCCGTCA
ATTCCTTTAA GTTTCAGCTG GGTGTGCACG TCTCGTTTGT GGACTCGGCA
CCTGAAGGTC ACGTGGTCAG GAGCCCCGTT GCTAGGCATT TAG

***S.h* repeat.8**

CTCACCATAG GACACCCGCG TTACCATTTG ACGGATGTAC CCGCCCCAGT
CAACTCCCCG CCTGACATGG TCTTCAGAAC GGGTCGCACA TCACATCAGG
ACTCAAACAC CGCAATTTGC ACCACGACTC ATCTGACATG ACCAAGATAC
CGAAGCACGC AGCGATAGAC CACGCACCCC GGAATGTGCT TTGAACCAAA
ATACAATTTC GTTCCCGTTT CACTGAATAA GTAAAGAAAC AATGATAGCA
GTGGTATTTC ATTAGCGCTG GACACCGAAG CACCCAG

***S.h* repeat.9**

CTCACCATAG GACACCCGCG TTACCATTTG ACGGATGTAC CGCCCCAGTC
ACTCCCCGCC TGACATGGTC TTCAGAACGG GTCGCACATC ACATCAGGAC
TCAAACACCG CAATTTGCAC CACGACTCAT CTGACATGAC CAAGATACCG
AAGCACGCAG CGATAGACCA CGCACCCCGG AATGTGCTTT GAACCAAAAT
ACAATTTCGT TCCCGTTTCA CTGAATAAGT AAAGAAACAA TGATAGCAGT
GGTATTTCAT TAGCGCTGGA CACCGAAGCA CCCAG

***S.h* repeat.10**

CCTGNTNACT GTCACTCNNC TCACACCATA ACCAACACGT GACACTTCTG
CTTTCTCACA GCGTTTCGTG TCACTTGAAC ACTTGGNGCA ATGCACAACA
CCGACCAGTG GCACTGCCTT ACCTAGTCCC ACCCACATCA TAG

***S.h* repeat.11**

CCTGGGTGTG CACGTCTCGT TTGTGGACTC GGCACCTGAA GGTGACGTGG TAGGAGCCCC GTTGCTAGGC ATTTAG

***S.h* repeat.12**

CCTCTCTAAA AGTTTCCGAT GCACCCGGTC TACTGAGAAC GATTCTGAGC
TGTCTAAACA CTGATTACGA TAATTACTCA AACCCAAATC AGGTAATCAA
AACCGGCGAA CATGGCGCAG TGTAACTGCC AATAAATTTA TAGATGGATG
CCACGTTTAG GTTGTCATTC CTTTGGCAAT ATGACAAGTA GGCTAGCAGA
GGTATTCTAA AACAATAAAC AAAATACTTT ACCAATTAAA GCCAACATTT
TTCTTGCAAT TTCGATTTAC TGTCTAGTGC TTGTGACTCA TCACTAAC**AG**
CTTGTGAATT CAGGCAATAT CGAGGCATAC GCACAGTATG CACATATGCC
AATAACAGAC TGATCAATTG CAGTGTTAAA CCTCAATGGG AAGATACAAG
TAAAAGACAA TATCAAGTGA ATTCGAATTC ACCACATTGC ACAAGCAAGT
GGCTATGAGA ACTCAGTAGT AAATGAATGT GTTTGACTGA CGGTAGTAGT
ACTTTTTGGA GGAGGAG

***S.h* repeat.13**

CTGCAGTGTT TGCACTTGAA ATAAAGGTAG CATACTTCGA CATGATTTTT
ACACATTTGA AGATTCATCT GA**TATTGTTT GCTTGNATCT TCCCATTGAT**
GTTTAGGATT TCAATTGATC AGACCCTTAA TGGCATATTT ACATCCTGTG
CAGATTGCCT CGATATTGTT TTAAGTCACA AGCATTATAA GCAGAGATGA
ATAGTGGCTA GCAGTGGAAT ACAGGACGCG TGTTTCGTCT TATTTTGGAA
CTCGTTAG

***S.h* repeat.14**

CCTACTGAGT TCTCATAGCC ACTTGCTTGT GCAATGTGGT GAATTCGAAT
TCACTTGATA TTGTCTTTTA CTTGT**ATCTT CCCATTGAGG TTTAACACTG**
CAATTGATCA GTCTGTTATT GGCATATGTG CATACTGTGC GTATGCCTCG
ATATTGCCTG AATTCACAAG

***S.h* repeat.15**

C**ACCTAGCCT GTCGTATGCC CTGATGGTGT TCTCGTAACT TTCGGGTTGC**
TGATCTGCCA AGGGCGATGG GACAGTGCAT GACGCTATTG TTGTGTGCTA
GGTTCAAAGA GAATTGTATG CTATATGCAT GCAAATCCGC CCCGTTATTG
TTCCTATTTC AAACTTTTAC ACTGTTGAAG CGATCCGGTT TGGCTTGCCA

TTCACGGGTT TGCTGCCTGG CATGCACCTG GCTTCGTGCT GGACTGCATG

***S.h* repeat.16**

C**ACATCCAGC TGACGAGTCC CGAACAGGAC GAAACGCGCG TCCTGGATTC CACTGCTAGC CACTACCCAT CTTTGCTT**GT TTAACAGTTT ACTGTTTCAA
TATTTATTCA AATCCCAAAT TTGGCAAGAG TATAATTTAT GGACGAACCA
ATCAGGTTTA CGATAATAGG TCTTGGATTT TGTCACGAAA TTCATTCATC
CATTTGGTTA AAGGGCATTT TGGGATTTTA GCTGATGTCA GTTCGTAATG
AAAACCATCA ACTGTAAATC ATAATTCTGA TTTCTCACAC TAATTTACAA
CCTTCATTTA GCCCTAGTAA GTTGGTGGAT ACTCTCAAGG CCATTTTAG
GTTACCCAAA TGTCGTCCGT AAATTAGAGT CTGGCACCGA ATTTTCTATT
TTATCTCTAT TTATGTATGT

***S.h* repeat.17**

CC**ACCTAGCC TGTCGTATGC CCTGATGGTG TTCTCGTAAC TTTCCGGGGT TGCTGATCTG CCAAGGGCGA TGGGACAGTG CATGACGCTA TTGTTGTGTG CTAGGTTCAA AGAGAATTGT ATGCTATATG CATGCAAATC CGCCCCGTTA TTGTTCCTAT TTCAAACTTT TACACTGTTG AAGCGATCCG GTTTGGCTTG CCATTCACGG GTTTGCTGCC TGGCATGCAC CTGGCTTCGT GCTGGACTGC ATGTAC**AATC AGACACGACC GAGGAAACGT TAGCTGGCAA GCAATGGCAA
TTTAAGCTAT TGCATGCAGT CACTACGACG TGTGACAAAC ATCACGAGCT

***S.h* repeat.18**

CCACCAGGGA CTGTCTTTAA CAAAGTGTTG CTGAACCGGA TGAAATGGAT
TCATAGTCGC CCAACTTCAA GGTCG **ACAGG CTGGATTCCG TAAAGATCGGTCATGCACAG ATCAAGTCAC GACACTACGG ATCATCATCG TGAGTGGAAC TCGTCACTGT AC**GTTGAGAT TCATTACCCC **AACAGTTAAT**
AACTGCGTTC GAAAGGACAC TTCGAAGTCC GTCACTGCAG AGCTACTGAA
ACATAAGTCC TGTATCGTAG GTAAAAACAA TACTAACGAT AACCCCAGAA
TGGGAAGTCT
TAAGGGT

***S.h* repeat.19**

CACGTAAGTA AATAAGTAAG TAAGTAAGTA AGCACATAAA TAAGTTAAGT
ATTAATAAAT TATGTAAGTA AGTAAATAAG TAAGTAAGTA AGTAAGTAAG
TAAATAAGTA AGTAAGTAAG TAAGTTAGTA AGTGAGTAAT TAATTAATTA
AGTAAGTAAG TAAGTAAGTA AGTAAGAAAG TAAGTAAGTA AGTGAGTA**CC CCGGTCAAGA AAATATCAAC CACTGAAGCC TAGGGTAATA ACCACCAAGC CAGGGTAC**GG CAAAATGTCG TGGTTCATAG TAATCCACGA TAAATAATAA
TTTATTATAT CTTACAGTAT TTCATTCACA CTACAGATAG AGT

***S.h* repeat.20**

CACTTCTACG AGAATGGCAG GGCCCTGAAT CTGCACTCAC TTCTGGGGTA
GTAGCAGATA GAAGTAGTCG AGAAGTTCGT GTATCTGGGT AGCTGTGTAA
GTGCTGGT

***S.h* repeat.21**

CACATCATTC AATAAATACA TTGATCTAAA TTTGGAATGA CCCTGGNATT
ATTATTGTTA ATATTTTGAC TGAGATTTGA TCATTCTTGA TTGACATATG
TATATCCTAT GCGGATTGTC TTGCTATAGC CGCAGGTTAG TATTGATTGT
TACAAGTAGT AATTAAAATG TTTCGTTAAT TTAAGTTATA AAGTCAATTT
CAATCCAGTA TAGTTATCTG CATAAAACAG ATAGGATATA ACTAGCAGTG
GAATCCATGA CGCACGCTTC ATTCTATTTA GGTCTCGTCA GATGGATGT

***S.h* repeat.22**

CACCTTTATA ATTTTTATTC CTAATATTGT TTATCATCAT TCCTTTGAGT
TCATTTTGTT AATTGTACCT GCATCTCAGA GTTGATGCTT AATCTGCGAC
TCGAACCCAG T

***S.h* repeat.23**

CACATCAAGG GGGTTCTGAC ATGCCTTAAA ATGTAGGCAG ACAAACCAAT
AGATCGTACC GATCGATAGC TGGGGCGGNG TTCAGGCTAA TTTCCTCTAT
ATCTCACTCC AATAAATGAT ATCCCGTTAA CTCAGTATTT AAGGTAGGCT
AAATCTCCTA GT

***S.h* repeat.24**

CCTGGGTGTG CACGTCTTGT TTGTGCACTC CAAACCTCAA GGGTCACCGT
GGTGGAGCCC CGTTGCTAGG CATTTAG

***S.h* repeat25**

CCTTGGTCAC GTGATTTTCA GTTTGCCCCA CCCTGATGCT GGCTGCCCCA

CCTCGACCGG CATAAGGTGG AG

***S.h* repeat.26**

CCTGNAGNGT NTGTATTTAA AAAATGATTN TCATCANCGA ATTGATATCT ATTGGAGATC TATTGATAAC AGTAAACAAC TAGTTTGGGA CTATTTAGCT GAATAGGTCC TATTTCGTAT GTNATTTCCC CTTTCTGGGA ATGGCTGTAC TATAGGAAAA TTGGAGATTG CAAG

A2:Schistosoma bovis

***s.b* repeat.1**

ATAAAAAAAC ANCCAAATGC AAATATATAT GGATTAGAAT AAAGCTAATT
AATAAAATGA AAACATGGAT ACAGGTTTTT **TAAGCAAAGA TGGATAGTGG**
CTAGTAGTGG AATCCAGGAC GCGCGTTTCG TCCTATTTGG GACTCGTCAN
CTGGATGTAC CTGAATCTCA NAGAGTTGAT GTTCACACCT GGACTCAAAG
TTGTNGTTTG CATAACAAAA AATGGTATCA GGTTATCTAT CCCTTTCCAA
TATTCAATAA AAGTAAATNA TCTANTATTC NCGANCATTN TACCACCTCA
TTNAATGANA CANCGCTGCG TNNAACACCA CCACNTTTCA CTNTGNCGCT
CCNTCNACGT NCTTAATCAN NCCTCCACAT TNCCNCCCNT NCNTCCCACT
CNCCNTNTC

***s.b* repeat.2**

GAAAAAATAA GACAGATATT GCTATTAGTA TCGTTTAAAT GTCGACTTTG
CTATAAAAAA AACAACACAT TTGTATGCCG ATTAGTTTCT GTGCTTGTAT
ATGTGTTTTC ATGTACCATT TTCGCTGATC GGTCACAGAC ACTTTTGTGA
GTAGTTTTCT TGTTAGCACC GTTCTTATTG **GGACCTTAGC GCCAGAGACG**
GATTTCAGTG GGATAAGGTG AGGTGTGCAT GNNGANCGTC ACCTNTCATN
TANCCTNTAN CNTTTCTNCG NNTNAAGGTT GTNTCCNCNC NCCCCCNCNC
NATCTCACCC NNTCNTACCC CACNTCCCCN CGCTACACGC CCCCTTATNN
CCTNAACCCA NTNNCCCCTT ATCTTCACAT NTNCCTTNCC NNCTCTNCNT
CACTNNCTCC NNTTCCCACG TCCCTNNCNT CCCCTNCNNC GNNC

***s.b* repeat.3**

ATGGCAACAG TCAACACCAC TAAGAGACTT TTTCACANGT GTGTCGGTTT
ACTGAAACTA TCTTTTTAGA ATCTCCTAGC TAGGCTATAC ATGTGTTGTC
CATAAAAATG GAAACATGAG CTATCGTGCA ATAACGT**AGC TGGTTATGAG**
TACGATGAAG TAAACGTTAA AATATGAGAT CGAATCCGCC AGGGAGCATC
AGTTCCCTCA AGATTATAGG TACACCTTGC TGACGAGTGC CAAGTAGCAT
GAAACCCGGG TCCAGGGTTT CCTGTTGACT ACCTCCAACC ACCATCTAAA
TCTCAATTCA TAGTGCCCGC AGTGTCGAGT CACCTAGATT GGTGGCCATA

TTGGAACATG ATC

***s.b* repeat.4**

TCACCTATCA AGACGAAACA AANAAAATTT TCAAATCGTT GGTGGAAGTG
CCTGTTTCGC AATNTCTCCG GAATGGTTGG TCGTATCGTC GTGAAAATTG
TTTCATATTA TTGGNGGCGA TCCCACCTAT CAANACGAAA CAAAGAAAAT
TTCCAAATTG TTGGTGTAAG TGCCTGNTNT GCAATATCTC CGGAATGGTC
GGTGGNATCG TTNTGAAAAT NGTTTTATAT TATTGGTGGC GACTC

A3:Schistosoma curassoni

***s.c* repeat.1**

ACCTAAGGAT TTGTTAAAAC TACCAGAGCT GAACACTTAT TATGTACATG
GGAAACGAGA CTAATTTAAG GACGAACCAA AGAGGTATAA GATAATAGAC
TCAAAGATCT TGGCAACCAA CTTCTATTCA TATTCAATGG CACTCACTCA
ATACAATCTG TTGTAAATTT CTACTCCAAT GAATTTGATT GCCTTACATT
CCTCCCACCA TATTATATTC CACTTTGTCT CAGTAATCGT TAGACGAATG
CATCTTGAAA AACCGAAATA ATTCCATACA CCGGGATAGG AATTATGACT
GTTGTTCGAG GGTTGGTTAG TTGAGTTAAT AAAGGTATGG AGATATGAAG
TGACCAACTT CTGATCCGGT TAAAGCGCCG GACATTCGCT TTTCGTCCTC
TCATTTTCGT AAACAACACC CCCGCCACGA GAAGGCACTG AGTAGGACTT
CCCTGGCAGA GGCTGTATAC GCGTGGCTGT GTGAGAGTAT TNCGAGAGGG
AGAGCGGACT CACCACACTC TCGGCCGTAC CAGGGCATTT GGGGACGAAC
AATACGTTTG AATTCTATGA AACATGACTT GATACTGTNT CACTCGATCA
TGAGTTAGAT AATTAGCTTG AATACTAGAC TGGAGGCCAC TAGATTGGGT
ANCCTCAGCT ACTTANAGTC TGGCCTAAAN ANGAGAACCA GCTTANATGA
GGCNAGNNNT TGGTTGACGC AANTANTANC CNGNTACCNN ANAGTANCAA
CCACGGNTAC TTATAANNAA TTCNCTTTAN TCATTCTGGC CTCACNCTNG
CNCACTGNCC NANACCTANC N

***s.c* repeat.2**

ATCAGTCAAT TGAAGTTAGA CATCAAGAAG TATTTCATGC AAGTTCCAGT
GAGAAGCAGT GACCAGTGGA GTTCAACCAG GCTTGTTGTG AGATAGTAAC
TCACTGAAGA CAATAGTGGA TTTATCGCTG AATTTCGTGA ATTAATTTAA
GTTAAACATG AACACCATTG GATGCTGGCT CAATGGTTTA GGGGTTAAGG
GTTCGTGAGC GANATCATTT ATCTTAATCA ACGATTTCAA AAGACAGTGC
ACAGTCATAT GGCAGTNCAC AGTTAAATGA AATCCCGAAT NCTTGNGANA
CAACTGTTCC TAATCAACCG ATTTATCAGA ATTCTCCATG CTNTTGCCCC
AAGTNTGACT TTCCCTNATG ATCCATATAT NTCTAATAAA NCTNCTTGCN
TTCCTGANGA AACTATNTCN ANTNCNCG

***s.c* repeat.3**

ATTTAATCAA CTTATTTCAG TTTAAACATT ACTGGGNCAT AAAAAACAAC
ACAATGAATA CATTTTTTGA TTATGTTCCT ATGAAGAAAA AAAAAGATTA
GAGAATTATA TGTATCAAAA CTATTTTACA TATTTAAGCC AATATTACGA
TTTCATTTTA ATGTATCCCT TGAGTACATG TATTTTCATG ATAACTTGTT
TGAATAAAAT GAATATAACG AAAATAAATG GGAATTAGAA AACGATAGAA
TATTGTATAG AATGAATAAA ATATATTATT CTTTATACAA ATTAAAAAAA
ACTATTCAAA TGTTACACAC AAAGAACATA ATATTGTTTC ATCTAC**GATC**
TCACCTATCA GACGAAACAA AGAAAATTTT AAAATTGTTG GTGGAAGTGC
CTGTTTCGCA ATATCTCCGG AATGGTTGGT GGTATCGTTG TGAAAGTTGT
TTCATGTTAT TGGTGACGAT C

***s.c* repeat.4**

AATA**CATCCT CGACCAGAAC GAAACCCAGC CTGCTCCTCG CGAGTCAATC**
TTTCTCGGGT TTTGAACAAC CTACGAAGTA TGACGGAAGC CTATAGCTTG
GACGCAATAG GAAGTAGACT TATCCTCCGA TAGTTGTTAC AGGAACGACG
TGAACCCTTT TTAAAGATAG GGACAACTAT CGACTCATTC CATGACGTTG
GTACACTCTC TAGTTCCCAA ACCTTTGTAA ACAACGTCGT CAATTCCTTA
GTCAAAAAGT CACCACCATC TTTAAAAAGA GCCGGAGGTA AGTCATTTGG
GCCAGGTGAT TTGTAACGCT TTAAGAGTTG GAGTTCCTTG CGGACTTTCG
CCTCATTTGG TGGATTGGTC GTCACCGGCC ATGGAGGGCA GGACAGTCTG
ACCNATGTTG CCCGGAGCAG CAGGCCCCTT GAACTGCCCT CCGAAAAATC
CTGCCCCAAC GCCCCANGAC GCNCGNTGGA TGNCACTGAT NGNCCNNTCC
CATCNTCCCT CCCCCGATTG TNCCGCCCAC ACCNNGACTT TNNTGCTTGC
CCANTGNNCT CNGGANGAGT TCCTGCNCCN CCCCNCNNNT CCNNCCCAAC
CCCATTNT

***s.c* repeat.5**

TTTGGGTCAT ATAAAATAAC CATAATTCAA ACAGTAATAC TAGTACCCAA
CGTCAAAAGT ATCAATTTGC ACTTACTCTT GATAATCGAT TTTGATTTTA
AAAAATGTTT TTAAACAATA TGCTTTCGTC AACAGGAAAT CCGAAAACAA
TTATGTTAGC GCTTTCATTA TGGAAAAGCA GTGATACGTT TTGTTGAATG
GTACGAAAAC ATATAAAATT TCTTATTGAT TTATTTAATT CATTTTAGAA
TAAACGTTTA CTATGATATC ATGTTTCTGG AGTTATTTGA TAAACAGACG
TCGGTCAGCT CTCTGACAGG AATGGAACAC TGGTGACCGC TCGAAACCAA
ATCATTCAGT CTTACTACAG TATCACCCAA ATATTTACAT GAATATCAGT
AGATTGTCAC AAATATTTGT CAATTATTCT GTGATATAAA GACTAAGTTT
CCATAATATT GAGATTCTTT CTAACTAGAC TCAAAATCTT CTTCAGTTAC
ATTAAAAGCA TGAATGCCAT CCAAACGTCT CTTCTAGTGT AATTCGTATA
TTCATAAAAG GTTTTAATTT TACACTTTGA ACTAAGAAAA CCTTATCCAA
TATATTCACT TATTTATTNT AAAAAAAGAG ACANAGAAGA ACTTTTAGCT
GACTAATATT CTTGGTTTTT ATGACAGAAG AGTGGGAAGA TGCANGAGAC

TACTTGGGGA TAGGATAACA NAAGTTTGGG GAAGGAAAAG GAAACCTTGT
AATGCCCCTA AGAGAGAAAG GAAATTATGA CTTAATACTT CCCATCTCTT
CCCCGGGCTG NAGGAATTCG ATATCAAGCT TATCGATCCG TCGACCTCG

A4:Schistosoma intercalatum

***s.i* repeat.1**

AAACTATTTC TTGGTACTTT ATCGTTCTCA TAAGCGTACT TCTTGTATAT
TTTGTGTTGT TTTGTATAAG TAGTATGTTT GGGTACACAA TTCACCCGAG
TACAATTCTT CATAATCCTT GATATCACTA ATATGAGATT CTGGGTAATG
ACTTAGGTTG TCAAGTTTGA ACCTGATTTC ATTATCTTCA TTTCTACACT
GTATTGCACC AACGCTTCTA CATAATGAAT GTCCGGTAAA GTAGTCCAAT
ATAGATCAAT AGTTTCTAGA TAATAGAATT AAGAATAATT ACCGAATCAC
GTCTGATGAC GTCTTATTGG TCTAAAGGTT GGATATTCGC GTATGAAACT
ATAGGTCCTG GGTTCGAATA CTCCAGTCAA CGACACCTGG ATCATGAGTC
AATTGAAGCT AGACCACCAT GGAAAACCTG GAAGCACTGG ACGGCCGTTT
CGTTCTATTG TGAGACTCCT CAACAGCGTG CATTCACGAC CCCGCCTTCG
CGGGATTCAA **ACCCAGTACC TATCAGTCTC GCGCCAGGCC CTTAACCCGC**
GAGGCGGGAT CATAATTTCA ATACTTCCAC GAGCCAATAA AACATAAAAA
ATTCAATACA TCACCTTTAT GAACAAGTGT ATCACCAGGT GTTNGGTTTT
GCATCTTTAA **AAATTTT**AAT GCAAGGAGCT CGGAAGACAA CCTTGACTTG
AAATTTTTTT AAAAAATTTT NAAAAAAAAN GGGANTTTCC CATTTTNGGA
AAANTNCCTT TTTGAATTAT GGAAGGGAAA AAAGTTATGT TNNCCATTTA
AAAAAATTTT TTANGGAACT ATTGTTTGGT AAATGAANG

***s.i* repeat.3**

TGTGTTATTC CTTTTCTAGT ATGCTTCTCT GAGTGTCTAG TCTACTCTTG
AATTAGTCAG TTGAACGTGT GGAAACCAGT GCTTCGGGTG GAAGGTTCCA
TGAATTAATG ACTCATTGTA AAAACTTGTA TGTCACGGGT GTTTTGTTCG
TTCTACTCGC CTCCTATGTT CTCTTAGATG TCCCAATTTA GTGAGTTTAT
TCAGTATGTA TTATCATCTT AAGTTGATTG ACTGACGTTT GAGTCAATTT
GGATAGTGGC TAACAGTGGA ATCTACGAGG CATATTTTAT TATATCTAGC
ACATGTACAC ACGATCAAAA TTATTCTTTT GAAAATGTGT TAATTCAAGA
AATTGCAGTC TGAATGGTAC TGAGCTGCCG AATT**CCCTGG TACGGCCGAG**
AGTGGGCAGA GTCCGCTCTT CCTCTCGAAA TGCTCCCACA TGGTCACGCG
TATACAGCCA CTACCAGAGA AATCCTACTC ATTGCCTTCT CGTGGCAGGG
GTGTTGNCTA CGAAATTGAG AGGACGAAAA GCGAATGTCC GGCGCTTTAA
CCGGATTGGT GGACACAGAG GTCCACCTAG GGGGAGGTTG GGAAAACCTT
GGTTCCCAAA CCANTGGTGC ACAATGGGCT ACAATACCCT GGAAGGGAAN
CAAATGGGCG TATGGAACCA NATCGTTGGG NGATCCCANT GCCCNCTGCT
ACTGTCGNAA CCGTTCAATT CATCNGATGG CCTAAGGTTC TACAGNTCTA
AAATGNGGCN NNTTGNGCCC AAAAANTTCN TAACGGGNCC CCNGTACCCC
CCNATCTAAC TANCCCAAAA CCTTGNNNGT CCTCAAAAAN GGATCCCACT
TA

***s.i* repeat.4**

CTTTAATCTG CAATGACATT CTTAATGAAT TCGAGGAAAC TATTCCANAA
GAGTCAAGTC TTGATGTCAT AGCAAATATT ATTTGTCCCC ATAATTCATT
TGTTTCTTGT GGGAAGCTTG TTCAATGCGA AGCACAAGTA TTAAATGAGC
TCCATTCTGA TTACAATTCA GATTATTTCA CATCAATTGC TGTTAATCCT
TATCACAAAT TCACTTCCAA TGTATACTCC AATCAATGTG AGAAATATGT
TTGAAACGAA GCCACATTAT TCATAACTTG GGGATATAAA GATCGTCTAG
TTGCATCCAA GATGTCTCTT GTATTCCTTG CTTTCCGTCA GGTGTGGAGG
TTTTTATAAT TCAGTCAAAT ACTAGCCCCC AAATGCCCTG GGTATGGCCG
AGAGTGGGGT GGGCCCGCCC TCCTCGAAAT GCTCTCATAT **GGGGGTGTTG**
TTTACGAAAT TGAGAGAACG AAAAGCGAAT GTCCGGCGCT TTAACCGGAT
CAGATAATTA TCCTTTTGAT TTTTGTATAT GAACTAACCT TGAAAGTAAA
GTTTTCAGTT TAAATTCGGA TTGAAGTTTC ATAGATTGTG GAAGTCGATA
AATGGGCGTT CAATTTCAAT ATCGAATTAT TTGTTCAAAG TTTATATATA
TTAACAAAGT CTAGTCAAAA TTTAGTTCTT GATATCAACA ATGGGATATT
GNGAACTCTT ATCAATATAC ATATATGTAT ATTTTAAACA GTTAAGATCT
AACAGAAACG AGTGATTTTC TTCTGTCCTA TATTTGATAT CAAGTTATAC
AACACTGACC CGTTGAACTG ACAAATGGAT ATACAGTAAG CCCTTCNCGN
ACCCTTGATA C

***s.i* repeat.5**

TCACCTATCA GACGAAACAA AGAAAATTTT AAAATTGTTG GTGGAAGTGC
CTGTTTCGCA ATATCTACGG AATGGTTGGT CGTATCGTTG TGAAAATTGT
TTTATATTAT TGGTGACGAT CGAGGTCTAC AGGACAAGCT GTGAGCCATA
AGTGGATAAA TTCGAATACT TCACTTTTAC CTAAAGTTTG TGCTGATTAT
GTCGTTCAGA AGAACGATGA AAGCTCCACG **ACCAAAGGAT CCACTAGTTC**
TAGAGCGGCC GCCACCGCGG TGGAGCTCCA GCTTTT

***s.i* repeat.6**

AGCATTAAAT TTNAAAAATA TTCAAGTTTG ATGGNTACCC AGNTCTCAAA
ACTTAAGTCA TTNCATCAGA TAACACGCCT ACTTCTAAAA CTACTTTAAA
TTACACGTTA TAGACAAACA CTTTATTAAA CTGAAATAAC AAACATATCA
CATAAAGCAC TTGTATATCA TAACA**GTATG TTAATTTTCT TTATTTGATT**
TGATCATCAA CAAAATTACA TCATTATTAG GAACAGCAGG TGAATTATTA
ACGTATTTTA TTCATTATCA CTTTTCATTT CNATAATTCC ATAGGTTCTA
GTGTTAGGTA TGGGATTTGA ACCTACGCCT ACANAATGGA CTGTGATTTG
GACGCAGCGC CATTGACCCC TNAGCCATAG GTTAAGAAAA GTATGATC

A5:Schistosoma margrebowiei

***s.marg* repeat.1**

TCTGACACAT ANCAATGTTT TAACATTAAT AGTCCACACG TTAACATAAA
AGACGGGATT ACCAGAATAA CTTCATATCC CTAAAGAAGT CAAAAGTTGA
AATGGGACAA TTATACATAC ATAGTAAAAT CCAATGAGCT TAGGGTAAAT
TACTTAGTTA ATAAACCAAA AACAATCAGT CAGATAAGAA CAAAACACAT
TGAGGAAATA AATAT **GATGC GCACTGCTGA AGAGTTCCAC AATAGGACGA**
AACGGCCGTT CAGTGCTTCT TGNTTTTCCA TGNTGGCTAG CTNTAATTGC
CTCATGNTTC NANCCATTGN ANTTCTC

***s.marg* repeat.2**

NGCTCNACTT TATAGTNGGG CTNANGANAC CGGACTNACC TTNATTACTG
ANATGTGNNA AGTCATGCAT NTGCACCATG TCGCANACTA CAGTGTCAGA
CNTGTGAAGA CTNNTGTCTA TCTAGTATCC CTAGNNTCNA AGAAN**GATTT**
ANAGAGTTCT GGGGCCCCAT GATNTAAAGG GNTNNNCTNA TNGCNACAAA
AATGCCTTCC GAGCAAACCT TGCACTGGTN NCGNNNAAGC GCANAACTGG
CCANNTGGAN NGANNANNTT TCCATACAAN NTNNAATAGT TNTATCCGNT
CCCATTTACA GNACNGAANC ATATNNCTAC CNNNTCACTC CAAAAGGACA
NTGGNCACTT TGGAGCCGTA TNCGACGTGN GAANCACGGA NANCACCNTN
AGGACTCAAN TNTAACGCTN TACGANAGTA ACGGNCTTCA TTCACTAGAC
CTTTACCCAT TAGANAATAN GCGNCTTACA GGTGATTTAT NAATGGNTCA
TNGCTCTTCT TAACACTTCT GGACATTCTC TTAGAACACC TACTTAAGCT
TAGGCTTCG AATNATAACC TACGGGGGTA ACACCCCANA AACNGGAAAC
TCAACATAGC TTGACGGGGA TGNCGACACA ACTTTTACTN CTTAAGAGNT
GCNAATGCGN GGAATCCGCT GCCGGCCGAG CTATNCCAAG CGACTTNTNA
GGAGTCCTTG ANGAGGCATC TTTCACCTAN TCTTACGGAC CAAGGCAGTT
TCNTACTATG ATTACCANTT TCTCTCCTCT NTTTTGGNAA CATACCTANG
TTCCTGCCNG NAAGATTTGG GCTCCCCNGC

A6:Schistosoma mattheei

s.mat repeat.1

GATCGAAGTC ATTAAAAAAA CGAAACGACC TTCACCTAAT GTGCAAATCA
AACAAGTCCA AGTTAATCAA AAAATCAATA AAACCTGATC ACAAGAACGA
ATTAATGACT TAATAATTTG ATAATAAAGA CAAGACCACT AACCAGGAGA
ATAAGTTTAC TTATTTAATG AAACCACATG CTTATATAGA AAATNCCAA**C**
ATAAAATAAA CATGATAGTG TTATAGGCTA TTGGAGGAAC TAAGTTTGTA
GGTCACTCAT TCTCCAATCA TTTAAGACTT AATNTCNGAN GGGGGTTTTG
TGGAGATTTT AGTTATTCCC CNTTGTTGNG ANCGCCCCCA ATCATATANA
ACNNTTCCCN cc

***s.mat* repeat.2**

ATTTATAATA CATNATGATG ATATGATATG ATTAGGCGNN GGCAGAAAAC
ATTTAATATT TACATCTGTA GCAGTCCCAT TTACATTTGT ACCATAACAA

TCAATATTAT TACCTGTTAA GGATAATTTT TGATTTGNTC TAGAATCTGA
ATACAATTTT AATAGTGTAT CCTTATTATT AATGTCTGTA GATGTTATAC
AATTAATGTT AGGATTTGTT GNTGCTGTTG NTGTACTAAT CTGTACATCA
GTAGTAATTG TTTTATCAAT TTTAGTACTG AGATTAATAG AAACTATACC
ATTAACATAA TTATCACTTG TTGTAGGTGA NCTTAAATGA ATTATTGCTT
CATTNGGGAG CCCTCGTCGN GCATTATTTG CCGACTCTGG GACTGGTTGN
AAAAAGCGGA GAGGTGCTCA GCGNATGCCA TGGTGTCGTG GTATGAAAGA

***s.mat* repeat.3**

GATCATGAGT CACTNNNACC TACCACCACC ATGGCANAAC CTGNAAGCAC
TGNACGGCCG TTNCGCCCTN TTGTGGGACT CCTCC

***s.mat* repeat.4**

GATCCCATAC TGATTGNTAC TCAACGGTGA ATTATATTGT GATTGTAGGA
GAACCAATGC TCAATGTGGA GATCAATTGC AGTCCTAAAT AACAAAGGGA
AGATACAAGT AAAACAACAC CAACTAAATT TAAACTTCAC CCTATTGCAC
AAGCAGACGG CTCTTAGGAC TCAGTGGCTA AGTGGATAAC ACGATAGCAC
TTGAAGCGAA CGGTTCTGAG TTCGAGTCCC AGAGTGGACA TCATTTCTGG
GATTCAGGTA CATCCAAGTG ACANGTNTGA AATAGGACGA NACGCCCGTC
AAACTGGATT ACACTGNCAC CCACTATCTA TCTTTGNTTA TNTAATCCTC
CTNAAAAATG ANAACCATTG AATCNTACGT AACCCAGACT CGTNNCCTAC
NATTNNCTGC CCTGCCTNTA ATCCTCCNTT NTAACCCNCC NCNCTCCCCC
CNCTGCTGNN TCNCTTNNTC NTNTGNCCTG CCTTCCNCCC CCCCTCNTCT
TCTCCCCCTC CCCCCCCTCC CCTCCTCCTC ACCCNCNTCC C

***s.mat* repeat.5**

AAGANGTTGC TANGTTGCGT AATGATTATT TTATAAAATG GATCTTCGNC
ACCAGGCTTC ACAGCTGCTT GATCCAAGTT ACCATACCTC ATTAGCACAA
CAAGATGGAC ACCGGATTCA TAGAAGTGGT TAATTCAGTG TTGGTAATAT
ATCAAAGAAA GATTGCATAT AAGGACATAG TACAGGAAGG AAGAACTAGT
TCGTAGAAAG AAAGATATGG AGCGATTTTA ATCTCATTGT TTAAGAGGGG
GGACAAAGAG TGTATAAACC GACGCCACTG TGATC

***s.mat* repeat.6**

GATCCACCAC TTTAAGAAAT TTAGGACAAA TAATCCTACT TTTAGATAGA
TGTGTCTCTC CCAACTAAAT ATGTCACAGA TTTAATCTAA TGACTGTTGA
AATGCTTATC CGTTTGAAGA TTGTTCTTTA CTAGTTACAG TATTTTCAAG
AGAATTAGGA GTAATCTATA TTTATGAAAC ATGTGAACTT TCTCTAGGTG
CTCTTCTCCG TTTGGCCCAG TCCAATAAAA AACCAATTGA GTCAGGTAAT
ATTAACGCTG ACGAAGAAAT TCATCTAAAC AGTCTGAATA AGTTAATTGT
GAAGGAACTA TCGGATATCA AATTTGATGA GGACAAATAT GAATCAGATG
ATTATAAGTC CAGTTAGCGG GATTGCATTA AAATGTACTT GTAAATTTCA
TAAAGCAGAG TATTTCCAAA CCATCTTTCT TTTTTTGAAT GGTATTATGA
GAAAACTGAG GTACTATACA CATAGCTCAC TTTATGTAAC CAACAGATTT

TATAAGTCGG TTATAAGATT GAAAACATTC AACTCAACAG TTTGGTGCCA
TCGTAGAAGG GTCTACCTAA TGCGTGAAAT ACCTGAAGGC TATAAGTTGC
GTTTAAAATC TTTAATACTG GCACATATAT CGTGTCTCTA GTTATAAATA
ATTGTCGCTT GGTATTAGAA CTGCCTTTTA CTTGGCGCAG AAAACTGAGT
ATATGGACTA AGGAATACCC ACCAAAAATC TATGGAAATT ACGAATGCAC
TATAAAAAAA CATTTTTGAC GTCACCCTTT TCTTCCATAG CACCAAGCTG
CGTGGTCAGC ACAAGTGTAG ACATGGTCTT TGC**GGGATCC ACTAGTTCTA**
GAGCGGCCGC CACCGCGGTG GAGCTCCAGC TTTTGTTCCC TTTAGTGAGG
GTTAATTTCG AG

***s.mat* repeat.7**

GTAAAACGAC GGCCAGTGAA TTGTAATACG ACTCACTATA GGGCGAATTG
GGTACCGGGC CCCCCCTCGA GGTCGACGGT ATCGATAAGC TTGATATCGA
ATTCCTGCAG CCCGGGGGAT CGGTAAATTA ATCTGTCGAA TCGTCCGTCA
AAAAATTTTA ATCAATATTT TACAACATAC AGTAAAATTA AAGAACACCT
AAATTAGAGA AAGATATTGT AGCGTGGCGT CGAGTTGAGG TTAACTATGA
ACACTGCTAC CAAGAAACAC GTGCCAAGTA AATCAGAAGG CGAAAATAGA
ACTTAACCAG GTTTATTTCG TTGGCGATCG TAGATGCACA CTTCTGAGGA
GTCCCACAAT AAGACGAAAC GGCCATCCAG TGCCTCCAGG TTTTCCATGG
TGGTCTAACA TCAATCAGTT TACGATCTCA ATCAAAAACT TAATAATCTC
CACAACCCTA GACTGATAGT GAAACTCCAA TTGAGGTGCT GATGGTAAAA
AGATAAGTTC AAAAAATCTC AAACAAG**TGT TGTTCCACAT ACTCATCGAT**
CCACTAGTTC TAGAGCGGCC GCCACCGCGG TGGAGCTCCA GCTTTTGTTC
CCTTTAGTGA GGGTTAATTT CGAGCTTGGC GTAATCATGG TCATAGCTGT
TTCCTG

smat repeat.8

ACTTGATTGA TGGATAAAGT ACTGTAGTAA TAGTTAAACT ATGCTAAACT
GATAATGTAT TAACGGTACT AATATGGTTC TGAAACCCCT ACTAATCAAT
CGAATAGTTT TCGTAAGACA CTTTTAACCA ACTCCTAGAT GAATAATTAT
GCATTCAACA ATTACAGCTA TTATTTGTGT TTGTAAAAAC TCGAATCAAA
ATAAGGTATG ATGAAAAGAA CAGTCACAAT AGCAGTGAAT AATTATTCAC
TATCAAGTGT TCAATGTGTA AATTATGGTC CTTTGTCGGT TATGTTGATC
ACTGACGAGG TTAATAGTTT GATAGTAATA TTGATATTGA TAAGTGCCAA
GATTGTCTGA GCACTACTTA AAACAAAAAT CCTTACTTCA AAAACTATCA
AATGATATAC ATCAAATATT AATATGTGCT AAAGCGTCAA TTACTTCTCA
ATGTTCAATA TACATTGTGT ACAAGTTGAA CTACAATATA ATGCAGATTG
TATTTTAAAA GTTTAGGCTG TTTGAATGCT AGTCGAGAAA TAAATTGTTT
ATTATCTGTA ATTAACTTTA GTTGATAAAA AACATTTAAG GTGGTCCCGA
TATGTACTTT TTTTTATTCG ACCAGATATC CTAAAACAAA TACTCAACTA
TTATTTGCAT TTACTAAGTA TCCTCATAAA GGAAAACATA NTTTCGGATA
TTTTAAAAAT ATTAAAACCA AATTTGTANT TACCGAACCA TGGAACTTAA
ACCTTTTGGG GGTTGGCATT CNAAAGTTTT NGGAAATGGG GGTTTTAANN
CGNGGGTTTN GNACTTTTAN NANGAGCCTN ATTTTCNANN GGGGCCCANA
ATTTCCGNNC AATTNNNANG GNCCNCN