

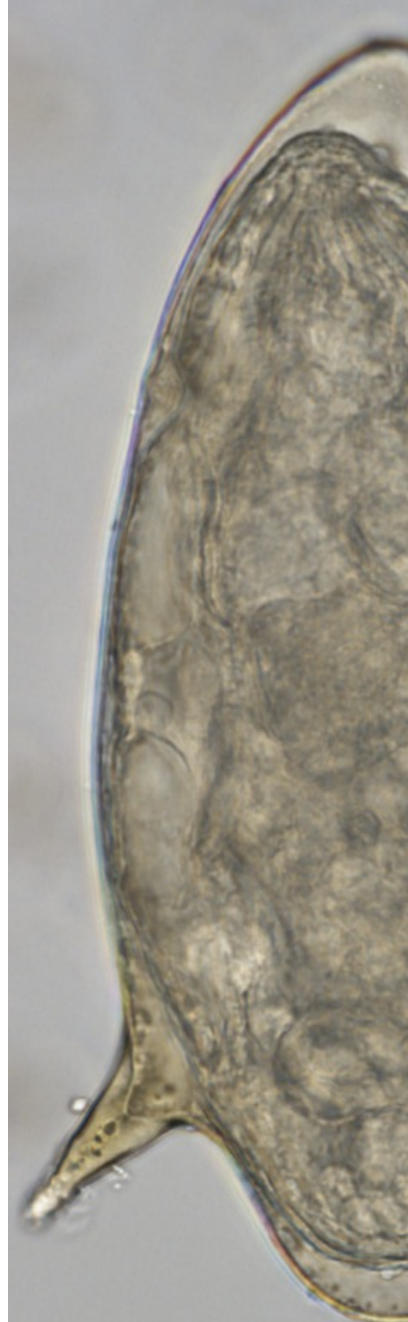


# A look into the genome of *S. mansoni*

**Arwa Khaled**

Intro. to Genomics (485)

Tuesday 18/1/2021



# Outline

- Selected paper
- Organism:
  - Taxsonomy
  - History of discovery
  - Physical charactstics
  - Parasitism
  - Geographic distribution
  - Life cycle
  - Significance
- Sequenced sample info
- Sequencing strategy
- Sequencing method
- Genome annotation
- Interesting genome outcome

## ARTICLES

# The genome of the blood fluke *Schistosoma mansoni*

Matthew Berriman<sup>1</sup>, Brian J. Haas<sup>3,†</sup>, Philip T. LoVerde<sup>4</sup>, R. Alan Wilson<sup>5</sup>, Gary P. Dillon<sup>5</sup>, Gustavo C. Cerqueira<sup>6,7,8</sup>, Susan T. Mashiyama<sup>9,10</sup>, Bissan Al-Lazikani<sup>11</sup>, Luiza F. Andrade<sup>12</sup>, Peter D. Ashton<sup>4</sup>, Martin A. Aslett<sup>1</sup>, Daniella C. Bartholomeu<sup>3,†</sup>, Gaëlle Blandin<sup>3</sup>, Conor R. Caffrey<sup>9</sup>, Avril Coghlan<sup>13</sup>, Richard Coulson<sup>2</sup>, Tim A. Day<sup>14</sup>, Art Delcher<sup>7</sup>, Ricardo DeMarco<sup>5,15,16</sup>, Appolinaire Djikeng<sup>3</sup>, Tina Eyre<sup>1</sup>, John A. Gamble<sup>1</sup>, Elodie Ghedin<sup>3,†</sup>, Yong Gu<sup>1</sup>, Christiane Hertz-Fowler<sup>1</sup>, Hirohisha Hirai<sup>17</sup>, Yuriko Hirai<sup>17</sup>, Robin Houston<sup>1</sup>, Alasdair Ivans<sup>1,†</sup>, David A. Johnston<sup>18,†</sup>, Daniela Lacerda<sup>3,†</sup>, Camila D. Macedo<sup>6,8</sup>, Paul McVeigh<sup>14</sup>, Zemin Ning<sup>1</sup>, Guilherme Oliveira<sup>12</sup>, John P. Overington<sup>2</sup>, Julian Parkhill<sup>1</sup>, Mihaela Pertea<sup>7</sup>, Raymond J. Pierce<sup>19</sup>, Anna V. Protasio<sup>1</sup>, Michael A. Quail<sup>1</sup>, Marie-Adèle Rajandream<sup>1</sup>, Jane Rogers<sup>1,†</sup>, Mohammed Sajid<sup>9,†</sup>, Steven L. Salzberg<sup>7,8</sup>, Mario Stanke<sup>20</sup>, Adrian R. Tivey<sup>1</sup>, Owen White<sup>3,†</sup>, David L. Williams<sup>21,†</sup>, Jennifer Wortman<sup>3,†</sup>, Wenjie Wu<sup>4,†</sup>, Mostafa Zamanian<sup>14</sup>, Adhemar Zerlotini<sup>11</sup>, Claire M. Fraser-Liggett<sup>3,†</sup>, Barclay G. Barrell<sup>1</sup> & Najib M. El-Sayed<sup>3,6,7,8</sup>

*Schistosoma mansoni* is responsible for the neglected tropical disease schistosomiasis that affects 210 million people in 76 countries. Here we present analysis of the 363 megabase nuclear genome of the blood fluke. It encodes at least 11,809 genes, with an unusual intron size distribution, and new families of micro-exon genes that undergo frequent alternative splicing. As the first sequenced flatworm, and a representative of the Lophotrochozoa, it offers insights into early events in the evolution of the animals, including the development of a body pattern with bilateral symmetry, and the development of tissues into organs. Our analysis has been informed by the need to find new drug targets. The deficits in lipid metabolism that make schistosomes dependent on the host are revealed, and the identification of membrane receptors, ion channels and more than 300 proteases provide new insights into the biology of the life cycle and new targets. Bioinformatics approaches have identified metabolic chokepoints, and a chemogenomic screen has pinpointed schistosome proteins for which existing drugs may be active. The information generated provides an invaluable resource for the research community to develop much needed new control tools for the treatment and eradication of this important and neglected disease.

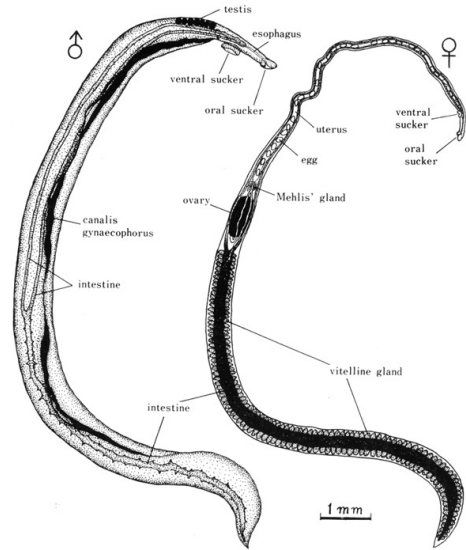
# My organism: Taxonomy

- SK: Eukaryota
  - K: Metazoa
    - P: Platyhelminthes
      - C: Trematoda
        - O: Strigeidida
          - F: Schistosomatidae
            - ***G: Schistosoma***
              - ***S: mansoni***

# History of discovery

- 1950 BC – Egyptian pharaohs record information about urinary/bladder infections
- 1200 BC – Mummy uncovered with cirrhosis of the liver and the presence of a schistosome egg.
- 1851 – Dr. Theodore Billharz, working in Egypt, identified the worms responsible for Schistosomiasis. (*Distomum haematobium*)
- The species distinction was first recognized by Patrick Manson
- Louis Westenra Sambon gave the new names *Schistosoma haematobium* and *Schistosoma mansoni*

# My organism: Physical characteristics



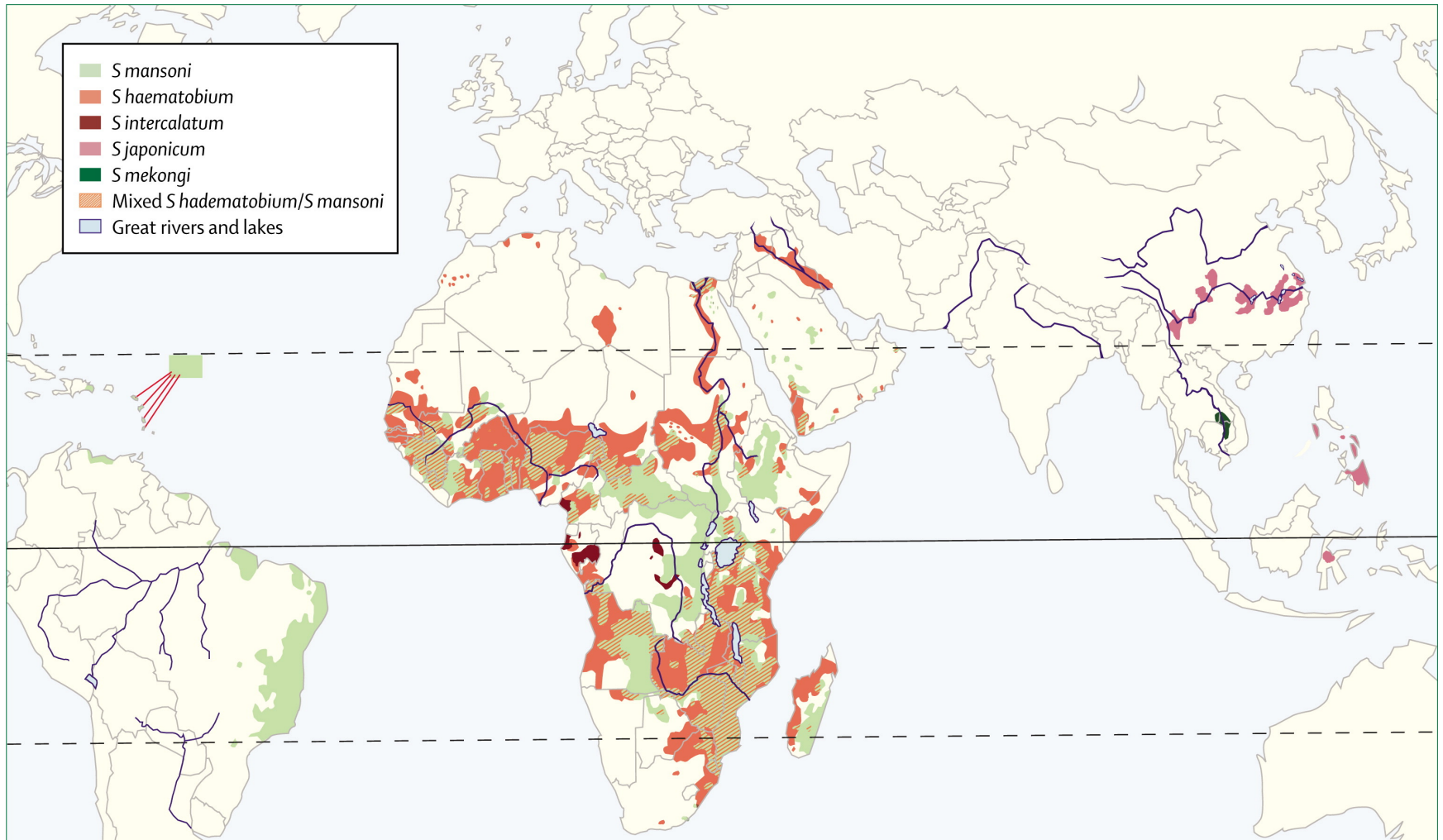
- Adult human schistosomes are dieocious
- Male length is (0.6-1.1 cm) and the width is 0.1 cm
- Female is thinner and longer (1.2-1.6 cm ) and the width is (0.016cm)
- Male possess a structure known as a gynecophoral canal.
- S. mansoni possess a nervous system that includes an anterior brain and longitudinal nerve cords
- S. mansoni egg shows a prominent lateral spine

# My organism: Parasitism

- Disease: schistosomiasis.
- Infection is caused by larval penetration during the exposure to the contaminated water.
- Intermediate host: *Biomphalaria alexandrina*.
- It causes abdominal pain, diarrhea, and blood in the stool and in urine.
- praziquantel, is almost exclusively used to treat the infection but this does not prevent reinfection

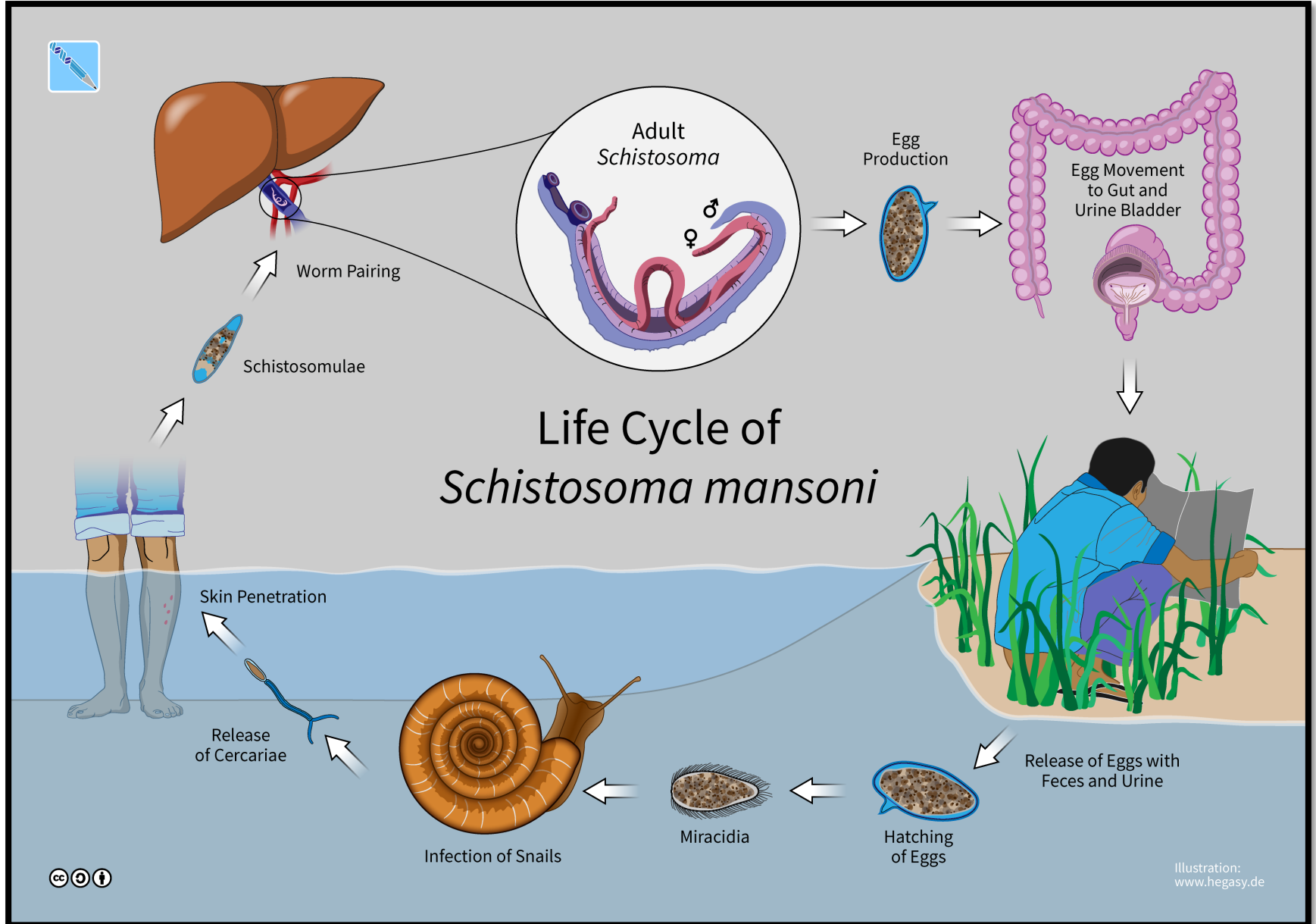
# Geographic distribution

- occurs across much of: sub-Saharan Africa, parts of the Middle East, Brazil, Venezuela and some West Indian islands





# Life cycle



# Significance

- WHO estimate cites over 700 million people as being at risk of infection, and approximately 200,000 annual deaths from Schistosomiasis alone
- After years of infection, the parasite can also damage the liver, intestine, lungs, and bladder.
- it offers insights into early events in the evolution
- of the animals, including the development of a body pattern with bilateral symmetry

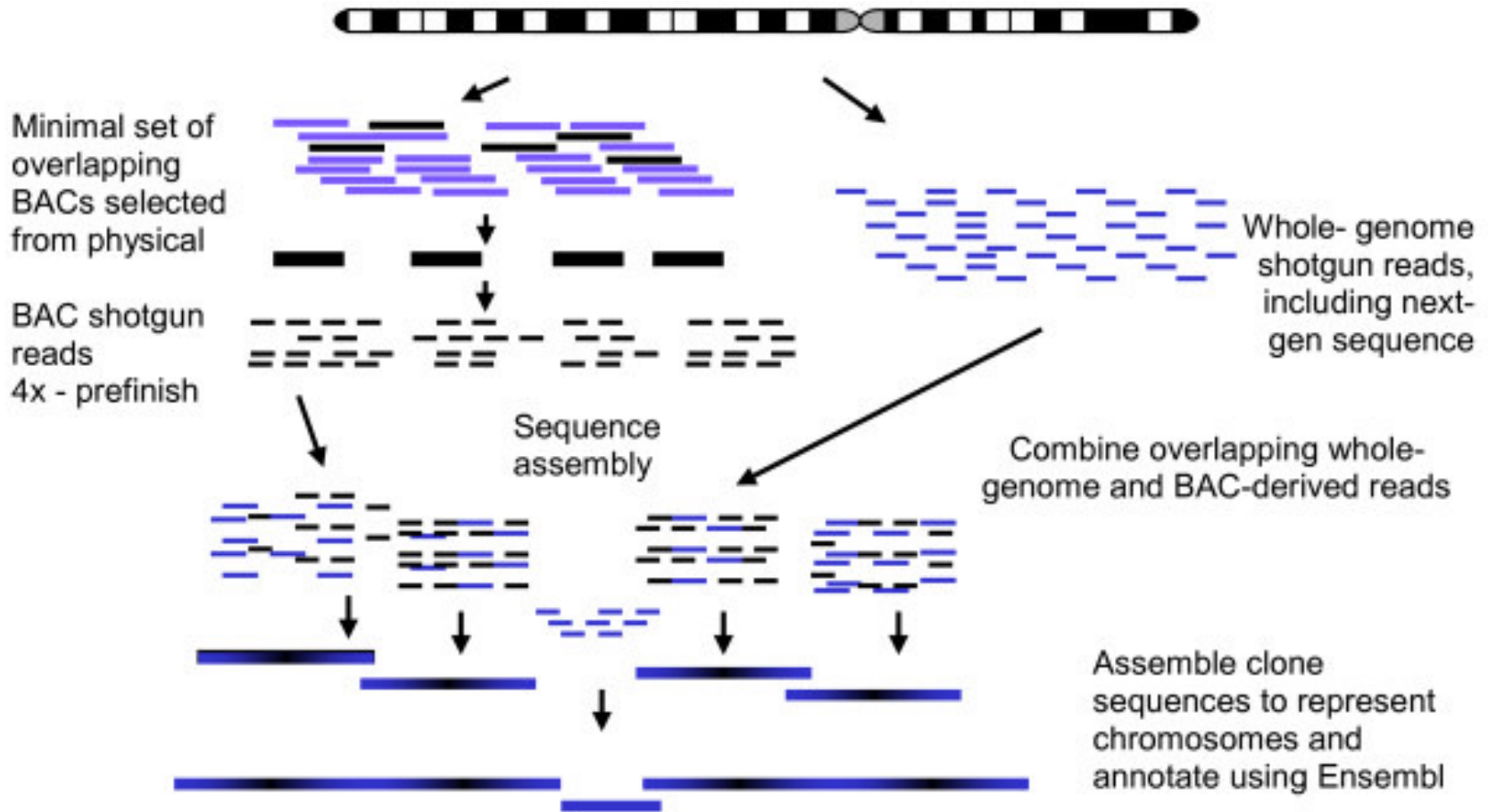
# Sequenced sample info

- Mixed-sex cercariae from the Puerto Rico isolate of *S. mansoni* released from infected *Biomphalaria glabrata* snails.



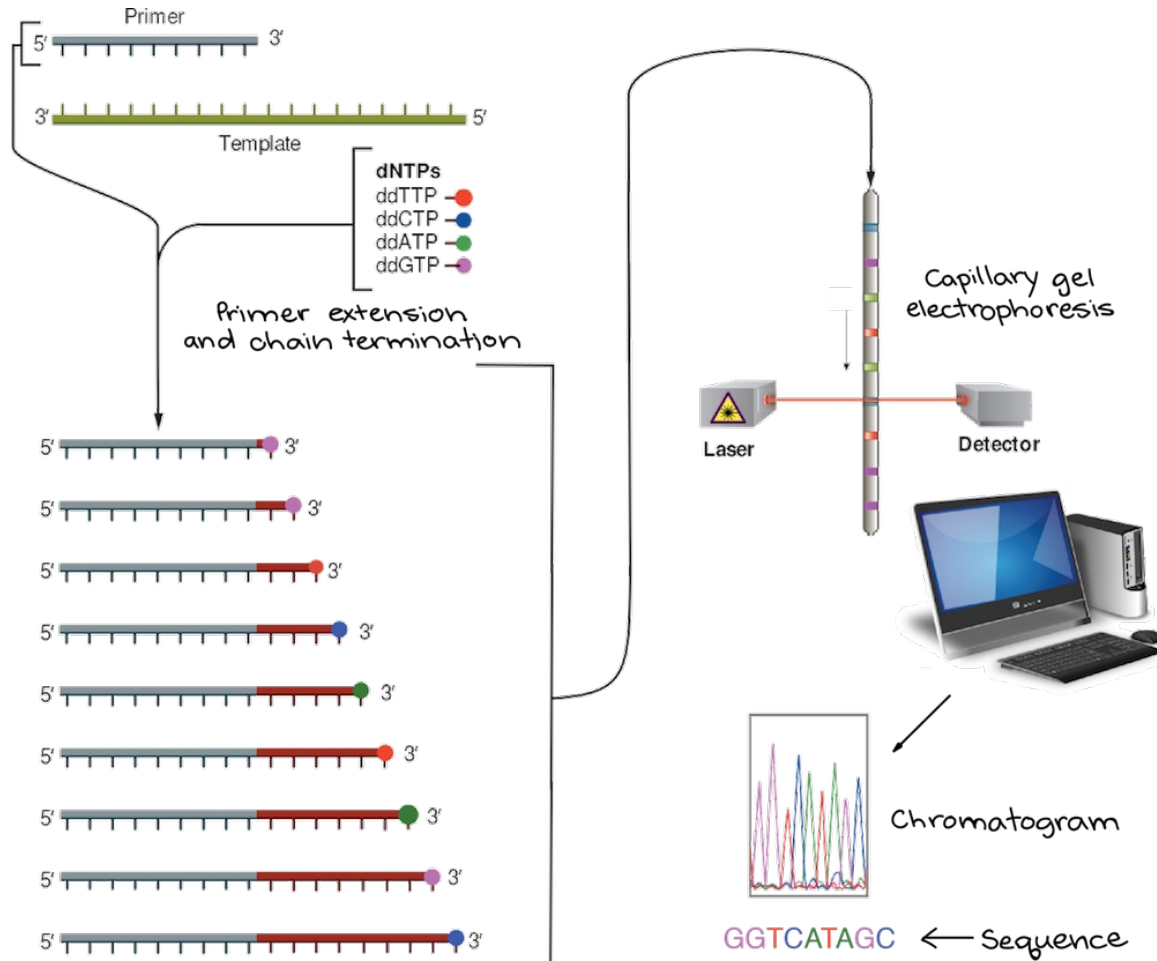
# Sequencing strategy

- Whole-genome shotgun+ Hierarchical (BAC clone) sequencing



# Sequencing method

- Sanger sequencing method



# Genome annotation

- Genome size: 363 Mb
- Coverage: Sixfold
- Number of chromosomes : 14+ZW or ZZ (2n)

## Assembly statistics (filtered, contigs < 2kb)

Total scaffold length, bp	362,875,701
Number of scaffolds	5,745
Scaffold N50, bp	879,876
Mean, bp	63,164
Largest, bp	4,179,495
<hr/>	
Total contig length, bp	350,053,037
Number of contigs	31,407
Contig N50, bp	17,677
Mean, bp	11,146
Largest, bp	139,372

# Genome structural annotation

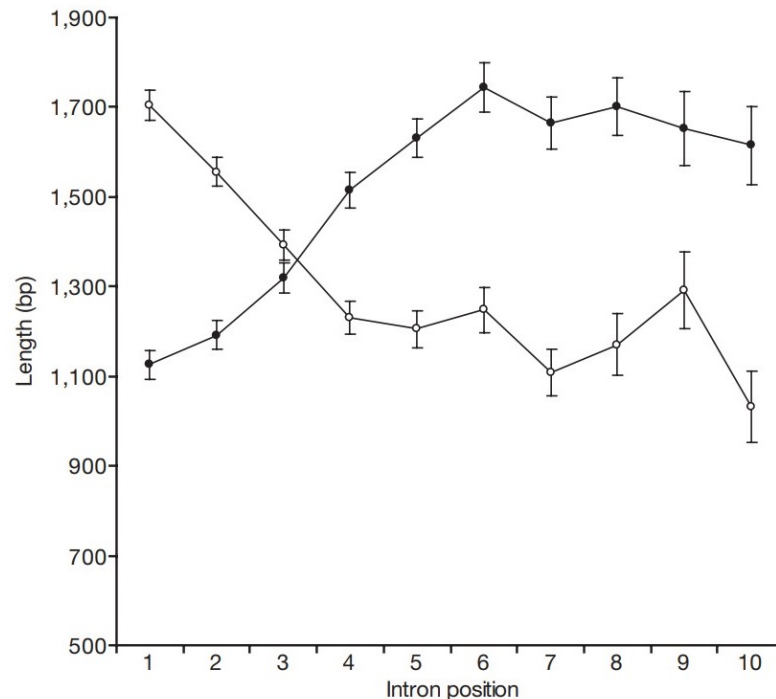
## Genome content

Average G + C content	35.3%
G + C (noncoding regions)	35.2%
G + C (coding regions)	36.3%
Repetitive sequence	45%
Coding bases	4.72%
Total number of protein coding genes	11,809
Total number of annotated transcripts	13,197
Number of genes with alternate splicing	980
Average exon number per gene*	7
Average intron size*, bp	1,692
Average exon size*, bp	217



# Interesting genome outcome

- Approximately 45% of the *S.mansoni* genome was found to be repetitive.
- The length of introns varies according to their position in a transcript, counting from the 5' end (solid circles) and the 3' end (open circles).





# Questions



Thank  
You!

- What do you think is the advantage of using both sequencing strategies?
- Why genome coverage is only 6x ?