

## Short Research Paper

# Whole-Genome Shotgun Sequencing and Assembly of *Anopheles gambiae* G3, the Host of Malaria Parasite *Plasmodium* sp.

Eva English<sup>\*1</sup>, Clara Gulick<sup>\*1</sup>, Ashley Jane<sup>\*1</sup>, Minseo Kim<sup>\*1</sup>, Nat Kpodonu<sup>\*1</sup>, Aden Lee<sup>\*1</sup>, Ramon Kodi Suzuki Lopez<sup>\*1</sup>, Krish Patel<sup>\*1</sup>, Achyuta Rajaram<sup>\*1</sup>, Leyla Unver<sup>\*1</sup>, Anne E. Rankin<sup>1</sup>, Shimaa M. Ghazal<sup>1</sup>✉

1. Science Department, Phillips Exeter Academy, Exeter, NH 03833, USA.

\*These authors participated equally in this research and the names were arranged alphabetically.

✉ Corresponding author: Shimaa M. Ghazal, PhD, Faculty at Phillips Exeter Academy, Exeter, NH. 03823. Phone: +1-603-203-7348, email: sghazal@exeter.edu.

© The author(s). This is an open access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>). See <https://ivyspring.com/terms> for full terms and conditions.

Received: 2025.08.05; Accepted: 2025.10.17; Published: 2025.10.27

## Abstract

*Anopheles gambiae* or the African malaria mosquito is the main vector of human malaria. G3, is an *Anopheles gambiae* strain, that was isolated from Gambiae in 1975, We report here a 216.3 Mbp draft genome sequence and assembly for *Anopheles gambiae* strain G3, with almost 82 thousand scaffold, 53X coverage and a G+C content of 44.5%.

Keywords: *Anopheles gambiae* G3, Malaria, Next Generation Sequence, Illumina HiSeq, MASURCA, GenBank

## Introduction

Despite being considered eradicated from North America since 1951, Malaria remains one of the most fatal diseases in the world with over 240 million recorded cases per year according to the World Health Organization, About 2,000 cases of malaria are diagnosed in the United States annually, mostly in returned travelers, and an estimated of more than 600 thousand people died from Malaria around world in 2020 according to CDC (1). The disease is caused by a protozoon from the *Plasmodium* genus that is carried by the mosquito *Anopheles*, the human infection takes place during the blood meal of the female mosquito *Anopheles* sp.

*Anopheles* were first described in 1818 as a genus of mosquitoes. Many *anopheles* are vectors for *plasmodium* parasites which cause malaria in birds, reptiles, and mammals; however, *anopheles gambiae* is the only mosquito genus known to be a vector for human malaria transmitting *plasmodium falciparum*. (2)

The clinical symptoms of malaria are largely a result of the replication of asexual states in human

blood but transmission to mosquitoes is only achieved through the development of gametocytes in the bloodstream (3). Developed female mosquitos carrying malaria, which are exclusively capable of its transmission, are known to be more persistent in feeding attempts, feed on a more diverse selection of humans, feed more frequently (4,5,6) and suffer greater feeding-associated fatalities than uninfected females (7 and 8). An infected female mosquito first takes a blood meal. This process involves the mosquito numbing the site of the bite with its saliva. Saliva harbors sporozoites which infect humans. In the human the sporozoites proliferate in the liver forming a schizont which will eventually erupt and be exposed to the human blood stream. In the blood stream there are two outcomes. Firstly, the Erythrocytic cycle where the more infectious particles are formed; this stage is responsible for the clinical manifestations of the disease. The second outcome is the formation of gametocytes which exist in male and female forms. Once a second infected female

mosquito has a blood meal, male and female gametocytes are joined to form a macrogametocyte through the sporogonic cycle. This process creates a oocyst in the mosquito which will burst forming more sporozoites allowing the mosquito to infect another human in the next blood meal. (1,7)

Many species of *Anopheles* mosquitoes exist as part of species complexes; that is groups of very closely related species. *Anopheles gambiae* complex was recognized as a species complex in the 1960s, The individual species of the complex are morphologically difficult to distinguish from each other, although it is possible for larvae and adult females. The species exhibits different behavioral traits and includes the most important vectors of malaria in sub-Saharan Africa, particularly of the most dangerous malaria parasite.

## Materials and Methods

The *Anopheles gambiae*, strain G3 was isolated in 1975 in The Gambia, Africa. (1,2) DNA extraction was carried out from three adult female mosquitos that didn't have a blood meal, using the Quick-DNA Tissue/Insect Miniprep Kit (D6016, ZYMO Research, Irvine, California, USA) and following the manufacturer's instructions. This step immediately was followed by library preparation, using Illumina DNA Prep, (M) Tagmentation, Accession number: 20060060), and IDT® for Illumina® DNA/RNA UD Indexes Set A, Tagmentation (Accession number: 20027213) were used. The library then sent to Hubbard Center for Genome Studies (University of New Hampshire, Durham, NH), where the draft genome of *Anopheles gambiae*, Strain G3 was generated using Illumina technology (9) techniques.

## Results

A standard Illumina shotgun library was constructed and sequenced using the Illumina HiSeq 2000 platform, which generated 216,250,770 bp. The Illumina sequence data were assembled using MaSuRCA v. 4.1.0 (10), which combines the efficiency of the de Bruijn graph and Overlap-Layout-Consensus (OLC) approaches. The final draft assembly contained 83,491 contigs, with an N50 of 7.2 kb. The total size of the genome is 216.3 Mbp, and the final assembly is based on 3,341 Mb of Illumina draft data, which provided an average 53.0 × coverage of the genome, with a G+C content of 44.5%.

## Discussion

This paper is a report for draft genome sequence for *Anopheles gambiae* G3, and this whole-genome shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession no. JAVFHU000000000.1. The version described in this paper is version JAVFHU000000000.1.

It is one of only nine genomes that were previously deposited in the Gene Bank for *Anopheles gambiae* G3 and one of four on the assembly level as scaffold, with 81,919 scaffolds for our strain. The list of all the nine genomes could be found here <https://www.ncbi.nlm.nih.gov/datasets/genome/?taxon=7165>. The data on the gene Bank showed the current gene set as 81.8% protein coding genes (12,518 genes), 8.8% non-coding genes (1,339 genes), 7% as small-RNAs (1,075 genes) and 2.4% pseudogenes (362 genes).

## Acknowledgments

*Anopheles gambiae* strain G3, the frozen adult female, kindly obtained from Povelones Laboratory, Department of Pathobiology, School of Veterinary Medicine, University of Pennsylvania, 3800 Spruce St., Rosenthal 304, Philadelphia, PA 19104.

The research was accomplished at Phillips Exeter Academy, Exeter, NH. is fully funded by Aileen and John Hessel, Class of 1952, Innovation Fund.

## Competing Interests

The authors have declared that no competing interest exists.

## References

1. World Health organization, WHO. Malaria. December 4, 2023.
2. Oaks SC Jr, Mitchell VS, Pearson GW, Carpenter CCJ. Institute of Medicine (US) Committee for the Study on Malaria Prevention and Control. Malaria: Obstacles and Opportunities. Washington (DC): National Academies Press (US); 1991. PMID: 25121285.
3. Meibalan E, Marti M. Biology of malaria transmission. Cold Spring Harbor perspectives in medicine. 2017; 7(3):a025452.
4. Anderson RA, Koellaf JC, Hurd H. The effect of *Plasmodium yoelii nigeriensis* infection on the feeding persistence of *Anopheles stephensi* Liston throughout the sporogonic cycle. Proceedings of the Royal Society of London. Series B: Biol Sci. 1999; 266(1430):1729-33.
5. Koella JC, Sörensen FL, Anderson RA. The malaria parasite, *Plasmodium falciparum*, increases the frequency of multiple feeding of its mosquito vector, *Anopheles gambiae*. Proc. R. Soc. B and P Roy Soc B. 1998; 265(1398):763-8.
6. Koella JC, Rieu L, Paul RE. Stage-specific manipulation of a mosquito's host-seeking behavior by the malaria parasite *Plasmodium gallinaceum*. Behav Ecol. 2002; 13(6):816-20.
7. Anderson RA, Knols BG, Koella JC. *Plasmodium falciparum* sporozoites increase feeding-associated mortality of their mosquito hosts *Anopheles gambiae* sl. Parasitol. 2000; 120 ( Pt 4):329-33.

8. Cator LJ, Lynch PA, Thomas MB, Read AF. Alterations in mosquito behavior by malaria parasites: potential impact on force of infection. *Malar J.* 2014; 13(1):164-174
9. Bennett S. 2004. Solexa Ltd. *Pharmacogenomics* 5:433-438.
10. Zimin AV, Marçais G, Puiu D, Roberts M, Salzberg SL, Yorke JA. The MaSuRCA genome assembler. *Bioinformatics.* 2013; 29(21):2669-77.