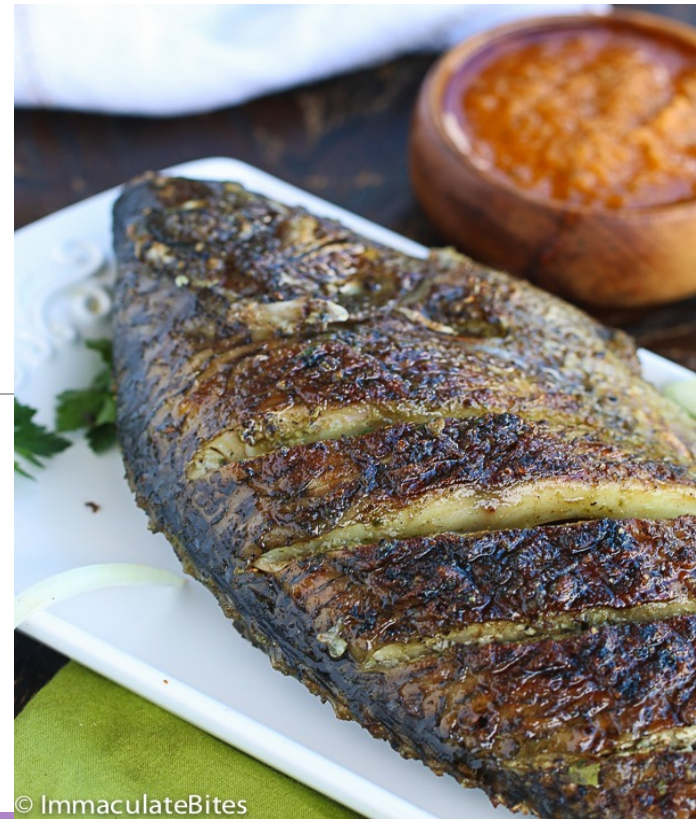# The delicious Nile Tilapia

Noah Mohamed

Introduction to Genomics 485
Fall 2020

© ImmaculateBites

Outline

- Fun Facts

- Studied genome paper

- General genomic information

- Sequencing strategy

- Sequencing methods

- Genome Assembly

- Genome outcome

\<Fun Facts\>

1- Among the most important farmed fishes globally.

<Fun Facts>

2- Some commercial names of Nile Tilapia:

- Mango fish

- Nilotica

- Boulti

<Fun Facts>

3- Scientific name of Nile Tilapia:

Oreochromis niloticus

- Oreo = Gold

- Chromis = Fish

- Niloticus = Nile river

<Fun Facts>

4- It is advantageous to grow-out only male Tilapias, why?

<Genome paper>

RESEARCH ARTICLE **Open Access**

CrossMark

A high quality assembly of the Nile Tilapia (*Oreochromis niloticus*) genome reveals the structure of two sex determination regions

Matthew A. Conte[1], William J. Gammerdinger[1], Kerry L. Bartie[2], David J. Penman[2] and Thomas D. Kocher[1*]

<General genomic information>

1- Number of chromosomes in Nile tilapia:
   Diploid 2n = 44
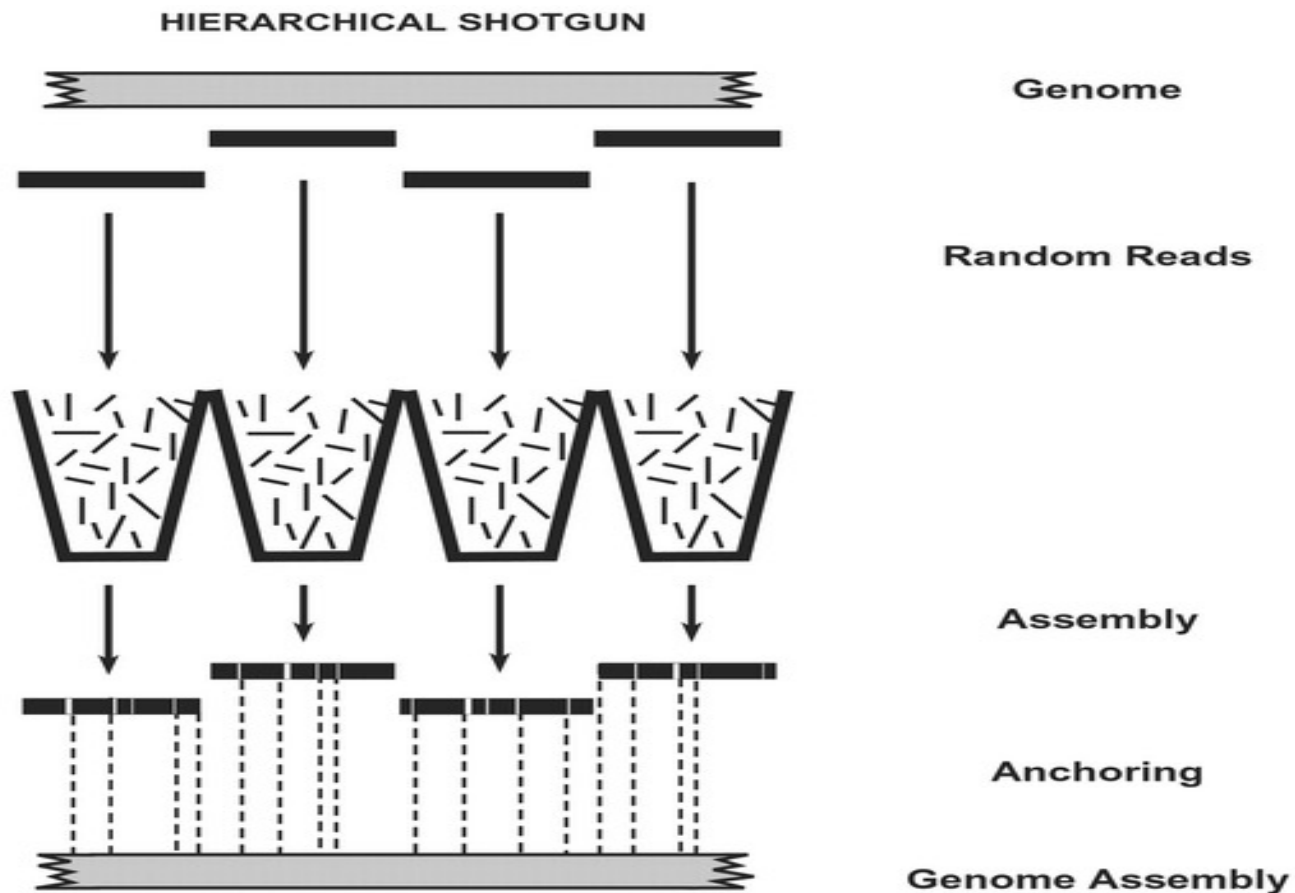

2- Source of DNA used to sequence Nile tilapia:
   Homozygous clonal XX female Nile tilapia

<Sequencing strategy>

- The used strategy was **BAC Clone Sequencing** .

<Sequencing methods>

- PacBio SMRT sequencing method.
  SMRT → Single Molecule, Real-Time sequencing

- Paired-end sequencing library was used.

LIBRARY PREP

SMRT SEQUENCING

DATA ANALYSIS

<Genome assembly>

- Using Bac clone sequencing, 37 de novo assemblies were generated based on both Canu and Falcon versions.

Canu assemblies	Genome size	N50 contig
Assembly #1	1,072,836,553 bp = **1Gb**	2,994,568 bp = **2.9Mb**
Assembly #3	985,204,113 bp = **1Gb**	875,570 bp = **875Kb**

Falcon assemblies	Genome size	N50 contig
Assembly #25	975,157,591 bp = **1Gb**	3,213,649 bp = **3.2Mb**
Assembly #28	982,875,404 bp = **1Gb**	4,333,907 bp = **4.3Mb**

<Genome assembly>

- The final chosen assembly was #14.

- Genome size → 1,003,343,259 = 1Gb
- # of reads → 5.1Mb
- Genome coverage → 44X
- Number of contigs → 2.9Kb
- Number of scaffolds → 5.9Kb
- N50 contigs → 3.3Mb
- N50 scaffolds → 2.8Mb
- Largest contig → 20Mb
- Largest scaffold → 13.6Mb

\<Genome outcome\>

Features	Orenil1.1	O. niloticus
Genome size	9.2Mb	1Gb
Number of contigs	5,677	2,960
Genes and pseudogenes	30,174	38,412
Protein coding	26,329	29,249
Non-coding	3,508	8,599

\<Genome outcome\>

- New assembly and annotation of the Nile Tilapia.
- The comparison between the candidate de novo assemblies.
- Two large sex-determination regions were characterized.
    - The first is ~9Mb region in O. niloticus
    - The second is ~50Mb region in the related
species O. aureus
- Further characterization of these regions will have important economic implications for farmed Tilapia.

Questions

- The better the assembly, the more no. of contigs we get.
  (True or False)

- Can Sanger sequencing method be applied for this genome project. Why if yes or no.