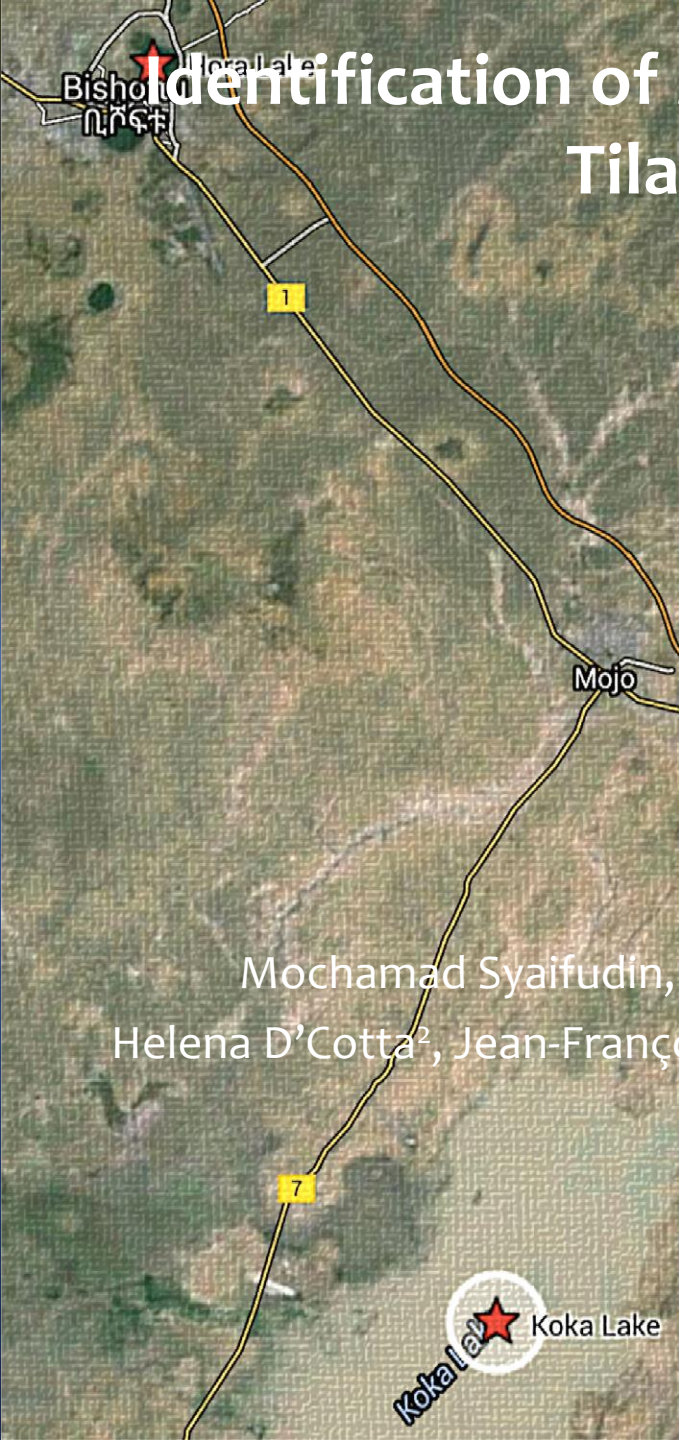


Identification of Species-Diagnostic SNP Markers in Tilapias Using ddRADseq



Mochamad Syaifudin, Michael Bekaert, John B Taggart, Gideon Hulata¹, Helena D'Cotta², Jean-François Baroiller², David J Penman and Brendan J McAndrew

¹ ARO, Volcani Center, Israel; ² CIRAD, Montpellier, France



Introduction

Tilapia Lineage

- a subdivision of the family Cichlidae, consisting of three genera with at least 70 species
- found throughout Africa and into the Middle East



Cichlidae

Haplochromine

Tilapiine

Tilapia

(substrate spawners,
biparental)

Sarotherodon

(mouthbrooding, biparental +
paternal care)

Oreochromis

(mouthbrooding, maternal)



T. zillii



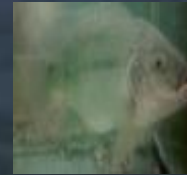
T. rendalli



S. melanotheron



S. galilaeus



O. aureus



O. niloticus



O. mossambicus

WILD

Introgression in Africa:

- *T. zillii* and *T. guineensis* in the Ivory Coast
- Introduced *O. niloticus* and native *O. esculentus* in L. Victoria, leading to loss of native species

Introgression elsewhere:

- High degree of mixing between *O. mossambicus* and *O. niloticus* in Southern Sri Lanka.

FARM

- In China, about 1/3 of the production is an F1 hybrid :
O. niloticus x *O. aureus*



- Red hybrids (multispecies, multigeneration)

Hybrid Occurrence

Meristic and morphometric characteristics show variation and broad interspecific overlaps

Genomic differences between species

- ❖ Species-diagnostic DNA markers
- ❖ Management applications

Objectives

- ❖ To look for species-specific markers, that distinguish between tilapia species.

A species-diagnostic SNP marker is one that has one allele unique for a given species (e.g. allele “A” for *O. niloticus*, “T” in all other species)

- ❖ To construct phylogenetic tree(s) among tilapia based on common SNP markers
- ❖ To verify tilapia species status of our samples using mtDNA COI gene (“Barcode of Life”)
- ❖ To analyse the distribution of diagnostic SNP markers in the reference genome

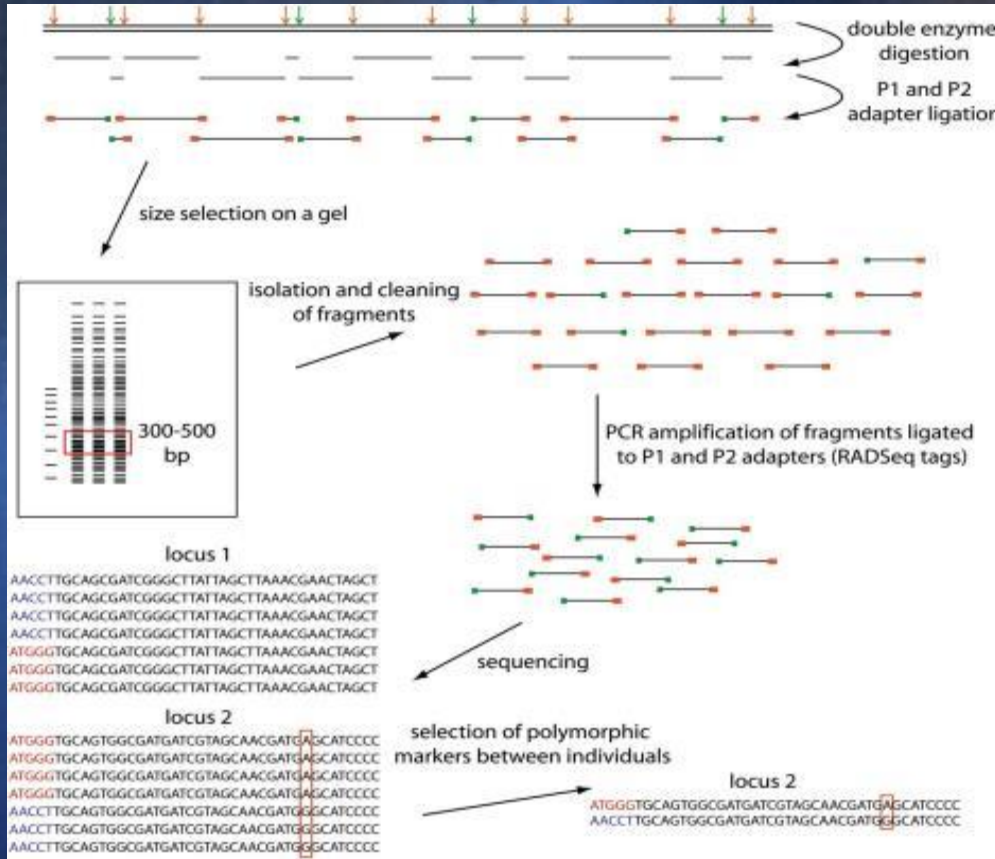
Materials



No	Species/sub species	Strain/ Population	Origin	n
1.	<i>O. niloticus</i>			
	a. niloticus	Stirling	L. Manzala, Egypt	6
		Kpandu & Nyinuto	Ghana	24
	b. cancellatus	Hora, Koka, Metahara	Ethiopia	33
2.	<i>O. mossambicus</i>	a. Stirling	Zimbabwe	5
		b. Natal	South Africa	10
3.	<i>O. aureus</i>	a. Stirling	L. Manzala, Egypt	5
		b. Ain Faskha	Israel	10
4.	<i>O. karongae</i>	Stirling	L. Malawi	5
5.	<i>O. u. hornorum</i>	Israel	Israel	5
6.	<i>O. andersonii</i>	Itezhi-tezhi	Zambia	6
7.	<i>O. macrochir</i>	Itezhi-tezhi	Zambia	4
8.	<i>S. galilaeus</i>	Israel	Israel	5
9.	<i>S. melanotheron</i>	Ghana		4
10.	<i>T. zillii</i>	a. Stirling	L. Manzala, Egypt	5
		b. Ghana	Ghana	5
Total samples				132

Double Digest RAD Sequencing (ddRADseq)

Quantified DNAs



Recknagel H et al. G3 2013;3:65-74

Phylogenetic Tree Construction

SNP Diagnostic Marker & Physical Map

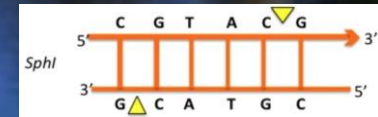
Green arrow: a rare-cutting enzyme

• *SbfI*



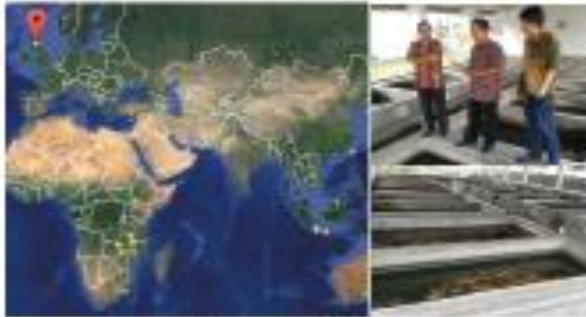
Orange arrow: a frequent-cutting enzyme

• *SphI*



METHOD

DNA BARCODE -COI



Samples Collection



DNA Extraction



FishF1-5'TCAAACCAACCACAAAGACATTGGCAC3'
FishF2-5'TCGACTAATCATAAAGATATCGGCAC3'
FishR1-5'TAGACTTCTGGGTGGCCAAAAGAATCA3'
FishR2-5'ACTCAGGGGTGACCGAAGAATCAGAA3'



Polymerase Chain Reaction



Purification and Sequencing



DNASTAR



Mega5.2



FigTree v1.4.0

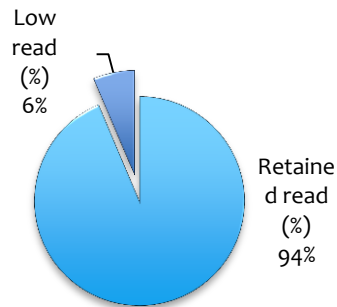
A screenshot of a software interface showing a list of DNA sequences and their translated amino acid sequences. The sequences are listed in a table with columns for 'DNA Sequences' and 'Translated'. The translated sequences are shown in a color-coded format (green, blue, red, yellow).

Data Analysis

Results – de novo Analysis

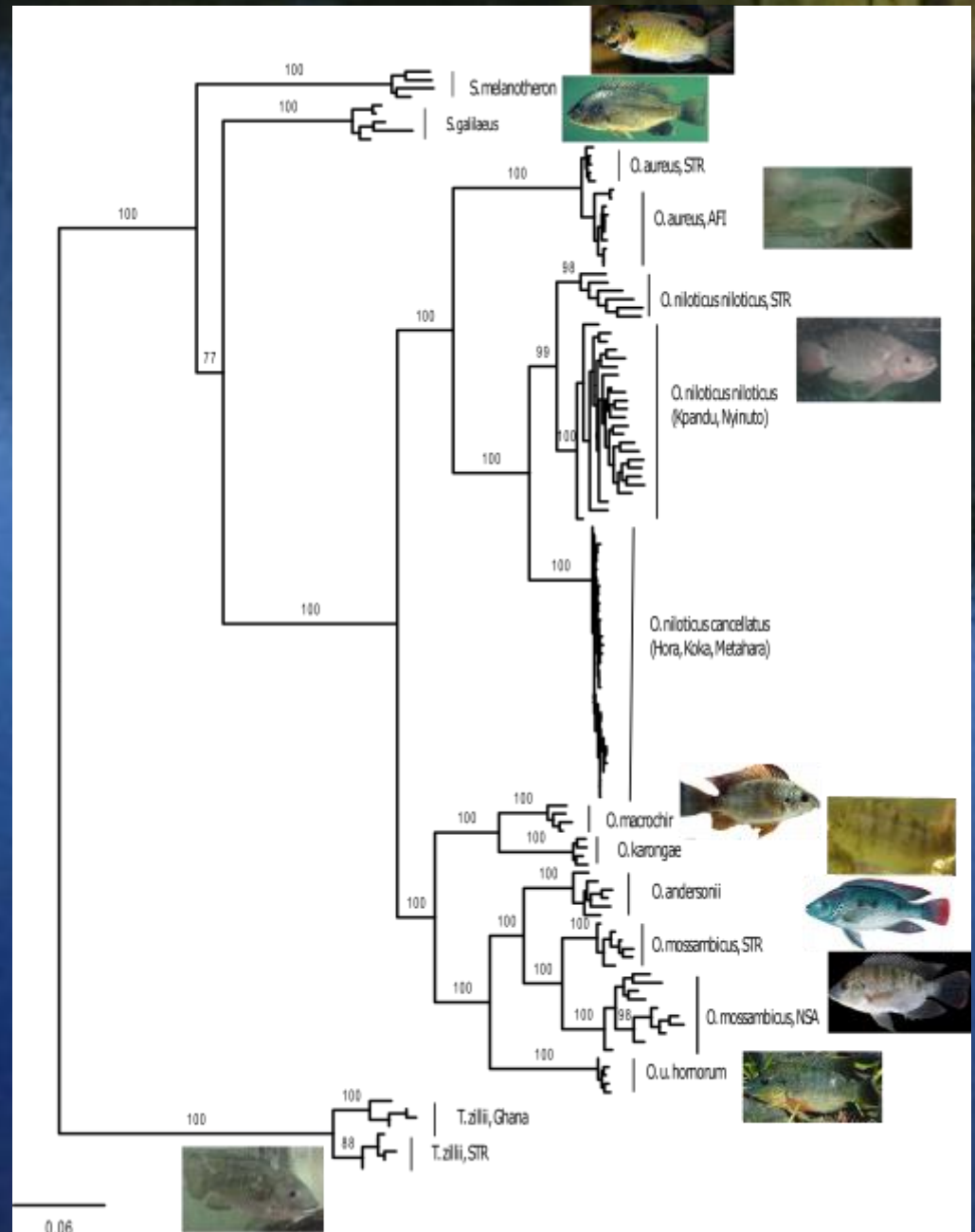
ddRADseq

109,287,766
raw reads



8,364 unique RAD-tags

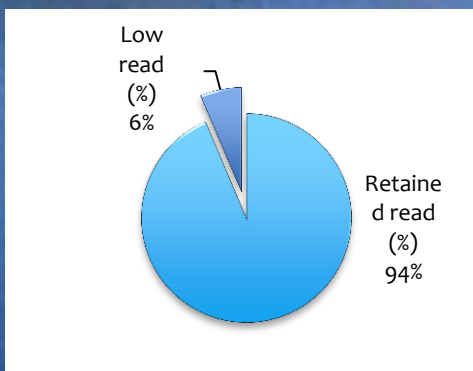
882 SNPs in 549 RAD loci



Results – Reference Genome-Based Analysis

ddRADseq

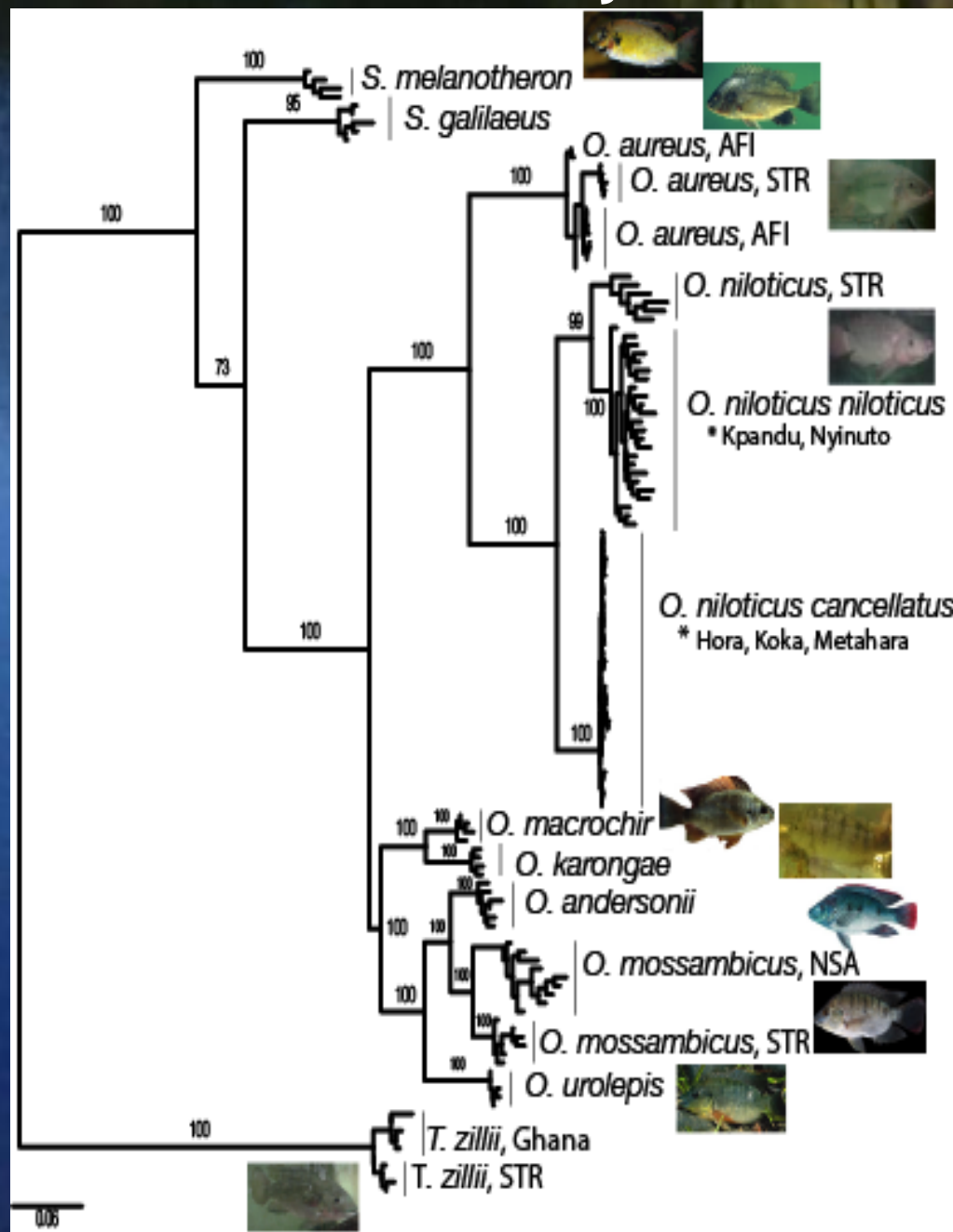
109,287,766
raw reads



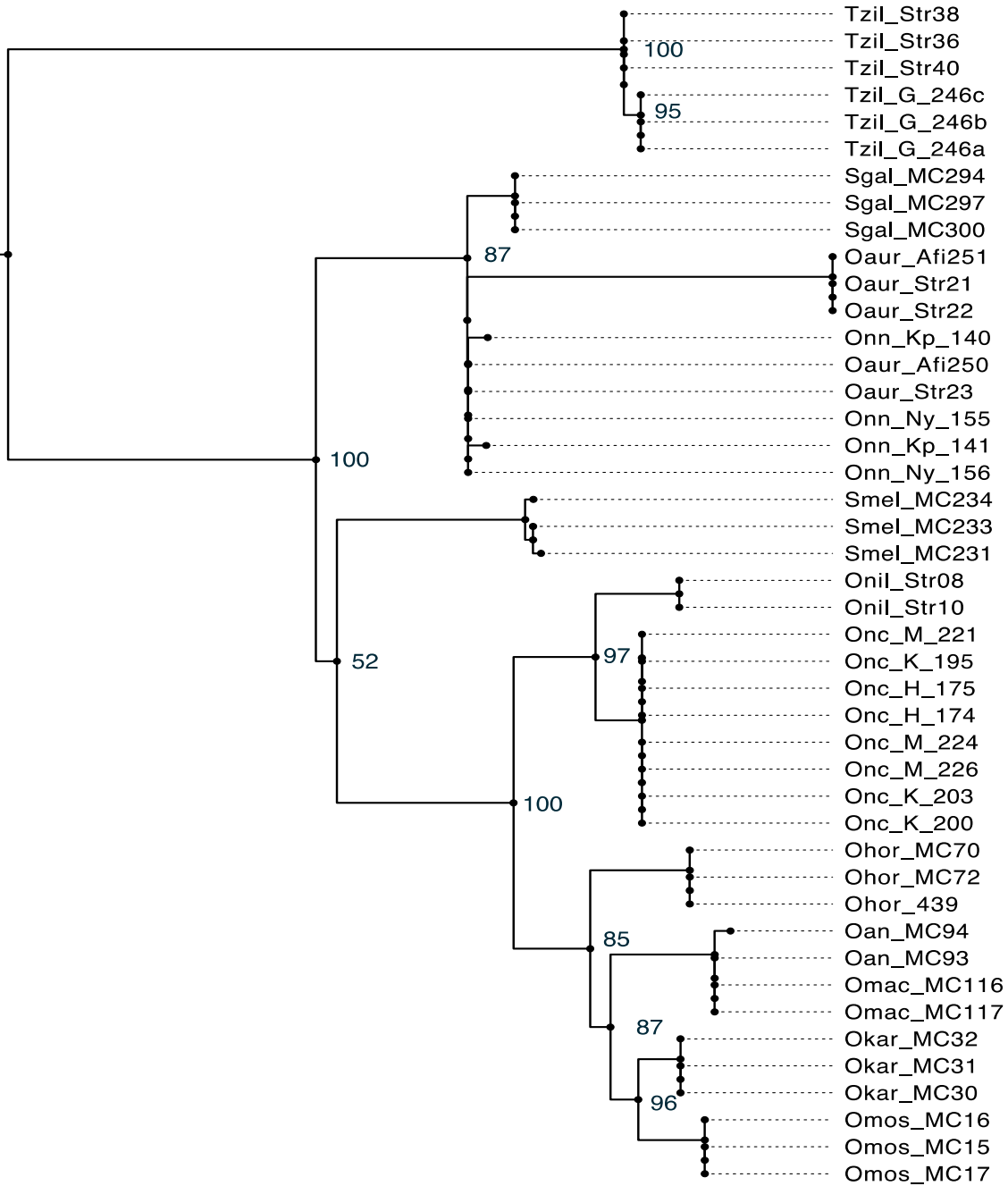
8,364 unique RAD-tags

Screen RAD loci against *O. niloticus* genome assembly

635 SNPs in 372 RAD loci

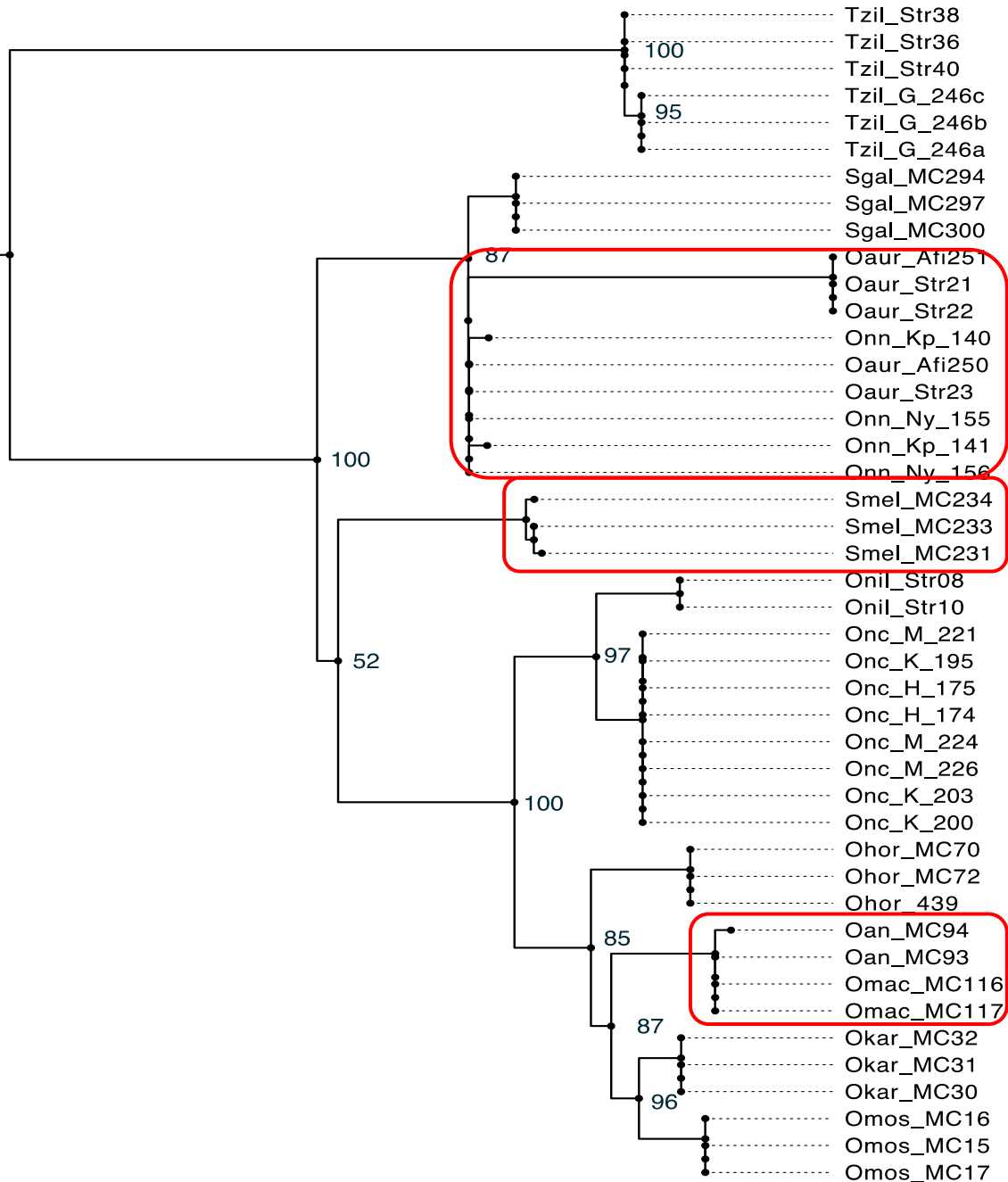


mtDNA COI Gene Tree



0.02

mtDNA COI Gene Tree



Largely agrees with previous publications based on allozymes and other mtDNA sequence, but:

- W. African *O. niloticus* group with *O. aureus*
- *Sarotherodon* clustered with *Oreochromis*
- Could not separate *O. macrochir* and *O. andersoni*
- only one marker

Genomic Distribution of (RBA) Species-Diagnostic Markers

- The markers were found across all linkage groups, apart from LG3
- LG3 is physically the largest chromosome, but is the smallest in the current genome assembly – contains lots of repetitive DNA

LG	Physical size of LG (bp)	<i>Oau</i>	<i>Oka</i>	<i>Omo</i>	<i>Oni</i>	<i>Our</i>	<i>Oan</i>	<i>Omac</i>	<i>Smel</i>	<i>Sga</i>	<i>Tzi</i>	Total	Number of SNP/Mb
1	31,194,787	0	0	0	2	0	0	0	3	0	2	7	0.22
2	25,048,291	0	0	0	1	0	0	0	3	1	0	5	0.20
3	19,325,363	0	0	0	0	0	0	0	0	0	0	0	0.00
4	28,679,955	0	0	0	0	0	0	0	0	1	2	3	0.10
5	37,389,089	3	0	0	0	0	0	1	4	1	3	12	0.32
6	36,725,243	1	2	0	0	0	0	1	4	3	2	13	0.35
7	51,042,256	0	0	0	0	1	0	1	1	2	2	7	0.14
8_24	29,447,820	1	2	1	1	0	0	1	2	3	2	13	0.44
9	20,956,653	0	1	0	0	0	0	0	1	0	1	3	0.14
10	25,048,291	0	0	0	0	0	0	0	1	0	2	3	0.12
11	33,447,472	0	0	0	0	0	0	0	4	1	1	6	0.18
12	34,679,706	1	0	0	0	0	1	1	0	2	0	5	0.14
13	32,787,261	4	0	1	1	0	0	0	2	1	5	14	0.43
14	34,191,023	0	1	0	0	0	0	0	2	0	3	6	0.18
15	26,684,556	1	0	0	2	0	0	0	2	1	1	7	0.26
16-21	34,890,008	1	1	0	1	0	1	0	2	3	3	12	0.34
17	31,749,960	0	1	0	2	0	0	0	3	2	0	8	0.25
18	26,198,306	0	0	0	0	3	0	0	1	0	0	4	0.15
19	27,159,252	0	1	0	0	0	0	1	1	1	2	6	0.22
20	31,470,686	1	0	2	0	0	1	0	2	0	0	6	0.19
22	26,410,405	2	1	0	1	0	0	0	1	2	0	7	0.27
23	20,779,993	3	0	0	1	1	0	0	1	1	1	8	0.38
Total	665,306,376	18	10	4	12	5	3	6	40	25	32	155	1.02

LG = Linkage Group, equivalent to a chromosome, here based on the current *O. niloticus* genome assembly

Species-Diagnostic SNP Markers (1)

- A species-diagnostic SNP marker is one that has one allele unique for a given species
- This analysis is based on all ten species and RBA
- 3-40 diagnostic markers per species

Species	n	Diagnostic SNP markers
<i>T. zillii</i>	10	32
a. Stirling	5	
b. Ghana	5	
<i>S. melanotheron</i>	4	40
<i>S. galilaeus</i>	5	25
<i>O. niloticus</i>	63	12
a. <i>O. n. niloticus</i>	30	
b. <i>O. n. cancellatus</i>	33	
<i>O. mossambicus</i>	16	4
a. Stirling	6	
b. Nathal, SA	10	
<i>O. aureus</i>	15	18
a. Stirling	5	
b. Ain Feskha, Israel	10	
<i>O. karongae</i>	4	10
<i>O. u. hornorum</i>	5	5
<i>O. macrochir</i>	4	6
<i>O. andersonii</i>	6	3
Total	132	155

Species-Diagnostic SNP Markers (2)

- This analysis is based on just the four species most commonly involved in hybridization in aquaculture
- 42-104 diagnostic markers per species

Species	n	Diagnostic SNP markers
<i>O. niloticus</i>	63	42
a. <i>O. n. niloticus</i>	30	
b. <i>O. n. cancellatus</i>	33	
<i>O. mossambicus</i>	16	104
a. Stirling	6	
b. Nathal, SA	10	
<i>O. aureus</i>	15	89
a. Stirling	5	
b. Ain Feskha, Israel	10	
<i>O. u. hornorum</i>	5	90
Total	99	325

Conclusions

- Analysis of ddRADseq data from tilapia species detected large numbers of species-diagnostic SNP markers among the species studied, distributed across the genome
- Species-diagnostic markers from this analysis should be robust across the species and thus of utility in analysing hybrids/introgressed populations
- Genotyping by sequencing is likely to be the most cost-effective and robust way of analysing species and potential hybrids (<£10 per individual)

A scenic view of a lake with a rainbow in the sky and a building in the background. The rainbow is the central focus, arching across the sky and reflecting in the water. The building, a large white structure with a prominent tower, is situated on a grassy hill behind the lake. The sky is overcast and grey, and the water is dark with some white swans visible. Bare tree branches frame the top and sides of the image.

Thank you for
your attention!!!