Genetic diversity of African lungfish from Great Lakes of Uganda

John Walakira

Scientist-Research Fellow,

NaFIRRI-NARO

Uganda

Morris Agaba

BecA-ILRI, Nairobi, Kenya

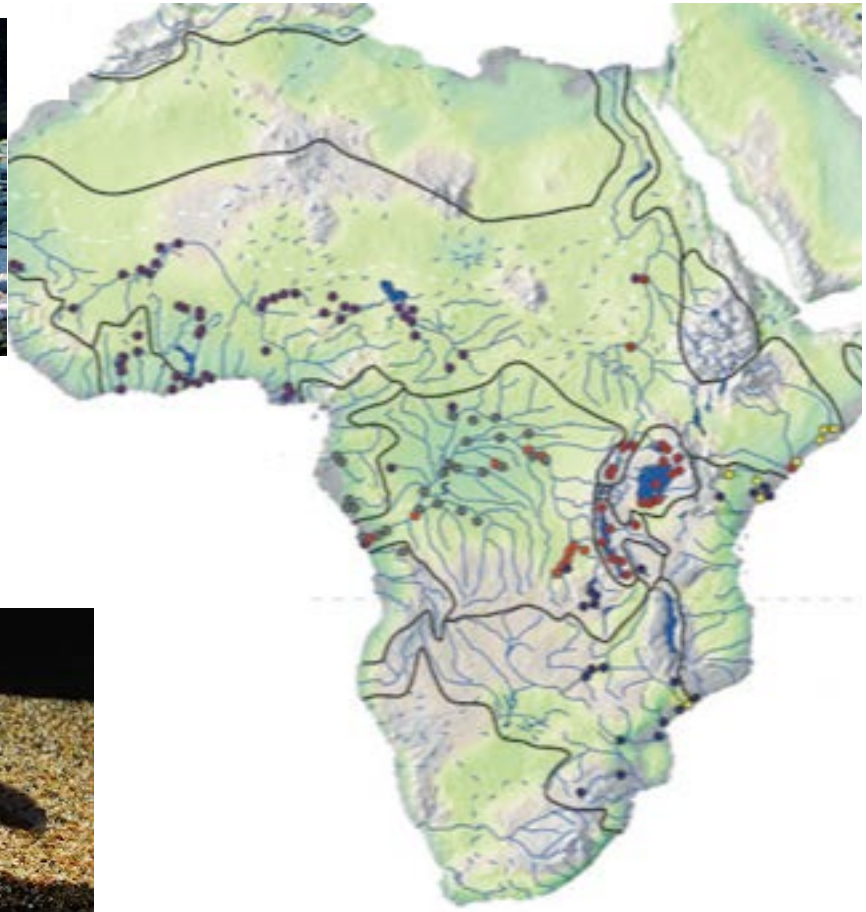
Taxonomy of African lungfish

- Species: *Protopterus aethiopicus* (Haeckel 1851)
- Genus: Protopterus
  - African lungfish with 4 living species
- Class: Sarcopterygii
  - Lobe finned fishes
- Order: Ceratodontiformes
  - Australian, S. American and African species

Source: http://www.fishlinkworldwide.com

Distribution of African lungfish

Protopterus annectans



Protopterus aethiopicus

Protopterus dolloi

Protopterus amphibius

Source: Paugy et al. (2008); Froese and Pauly (2009); Otero, (2011). http://www.britannica.com/animal/lungfish

Biology & ecology

1. Shallow/swampy/turbid freshwaters,

2. obligate air breathers,

3. Aestivation: able to survive the dry season (4-6 months),

4. Seasonal breeders?

Source: Ngugi et al.

Why lungfish?

Improves food nutrition & Income

Natural stocks are declining

Aquaculture potential; co-cultured with farmed tilapia + catfish

- Bio-control agent against shistomiasis (Daffalla et al. 1985)
- Evolutionary Research

Problem Statement

- Climate change + anthropogenic factors: decline in stocks/low production
- Intervention: Aquaculture production
- Problem: Lack of breeding technologies

Objectives

Overall objective: Genetic diversity of African lungfish for use in aquaculture in Uganda-East Africa.

Specific objectives:

Objective 1: Understand the phylogeny of African lungfish from six lakes in Uganda

Objective 2: Develop a panel of SNP to guide the domestication of African lungfish

Objective 3: Develop a tool that identifies sex of lungfish

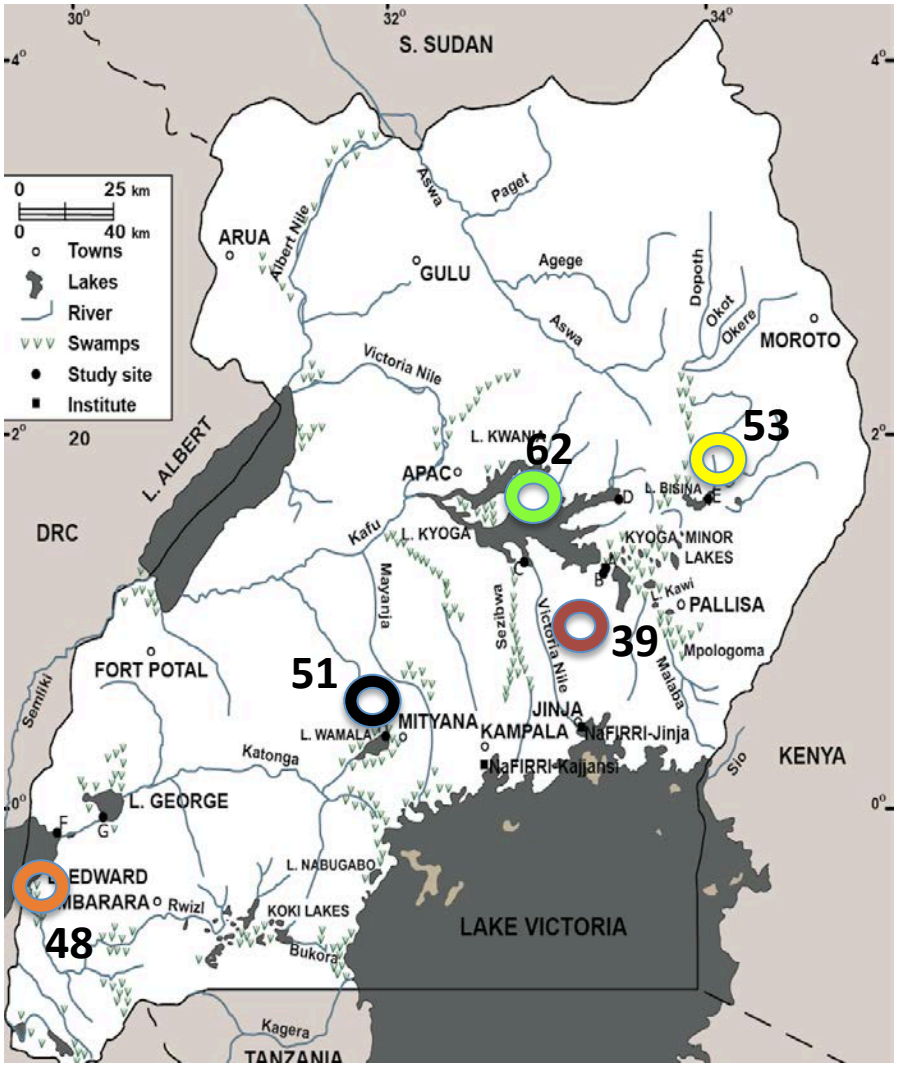
Population biology of African lungfish using mtDNA

Outputs:

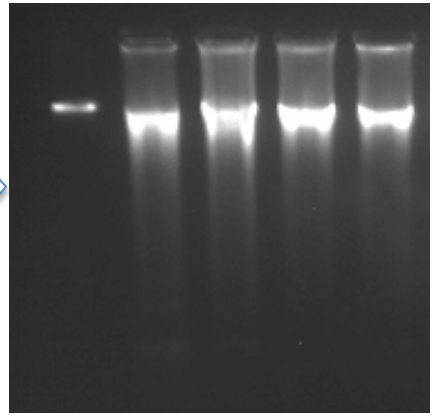
1. Diversity lungfish in populations

2. Phylogeny of lungfish

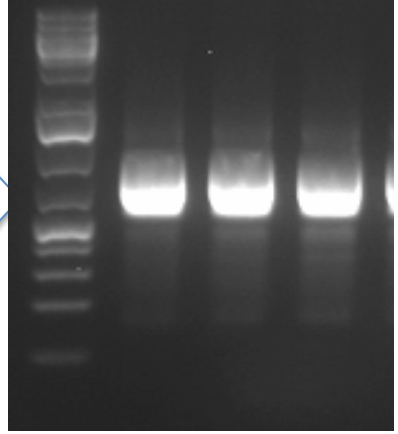
3. Gene-flow within lungfish population



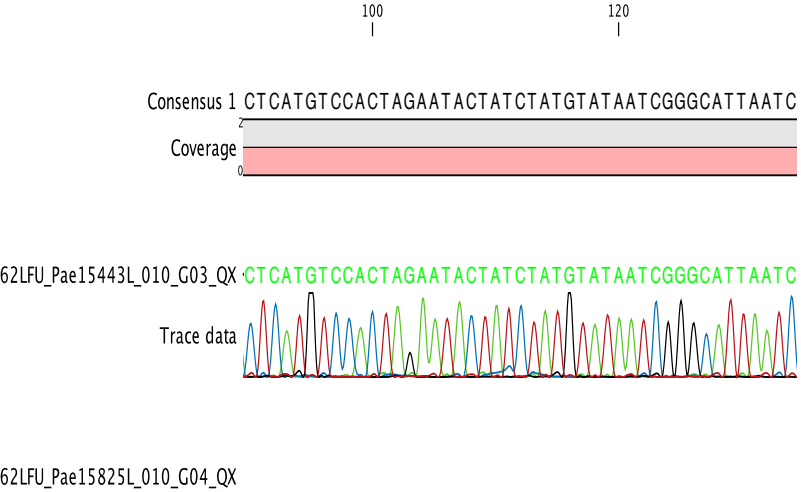
gDNA from 254 samples



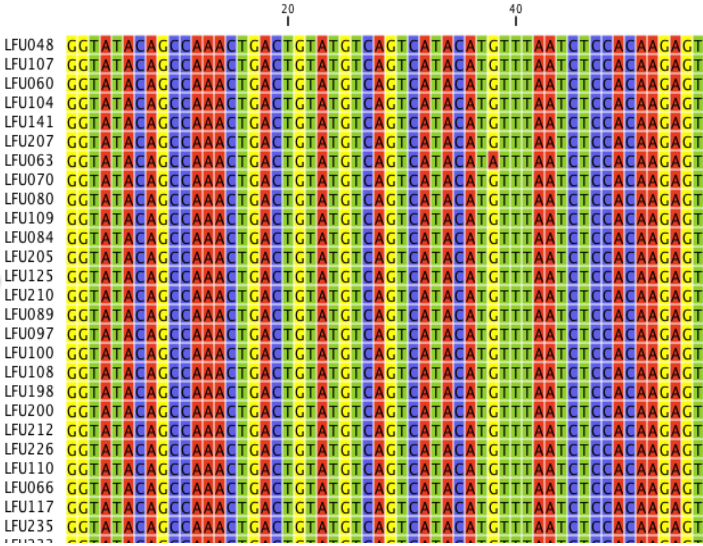
253 samples for PCR



253 products purified using Qiagen kit



244 seq. from Sanger platform



205 good sequences after QC, 843bp

AMOVA

Source of Variation	d.f.	Sum of Squares	Variance components	Percentage of variation
Among populations	5	182.885	1.009 V_a	20.91
Within populations	197	751.608	3.815 V_b	79.09
Total	202	934.493	4.824	

Fixation Index F_{ST} : 0.209

All values significant $P < 0.001$

Diversity of African lungfish

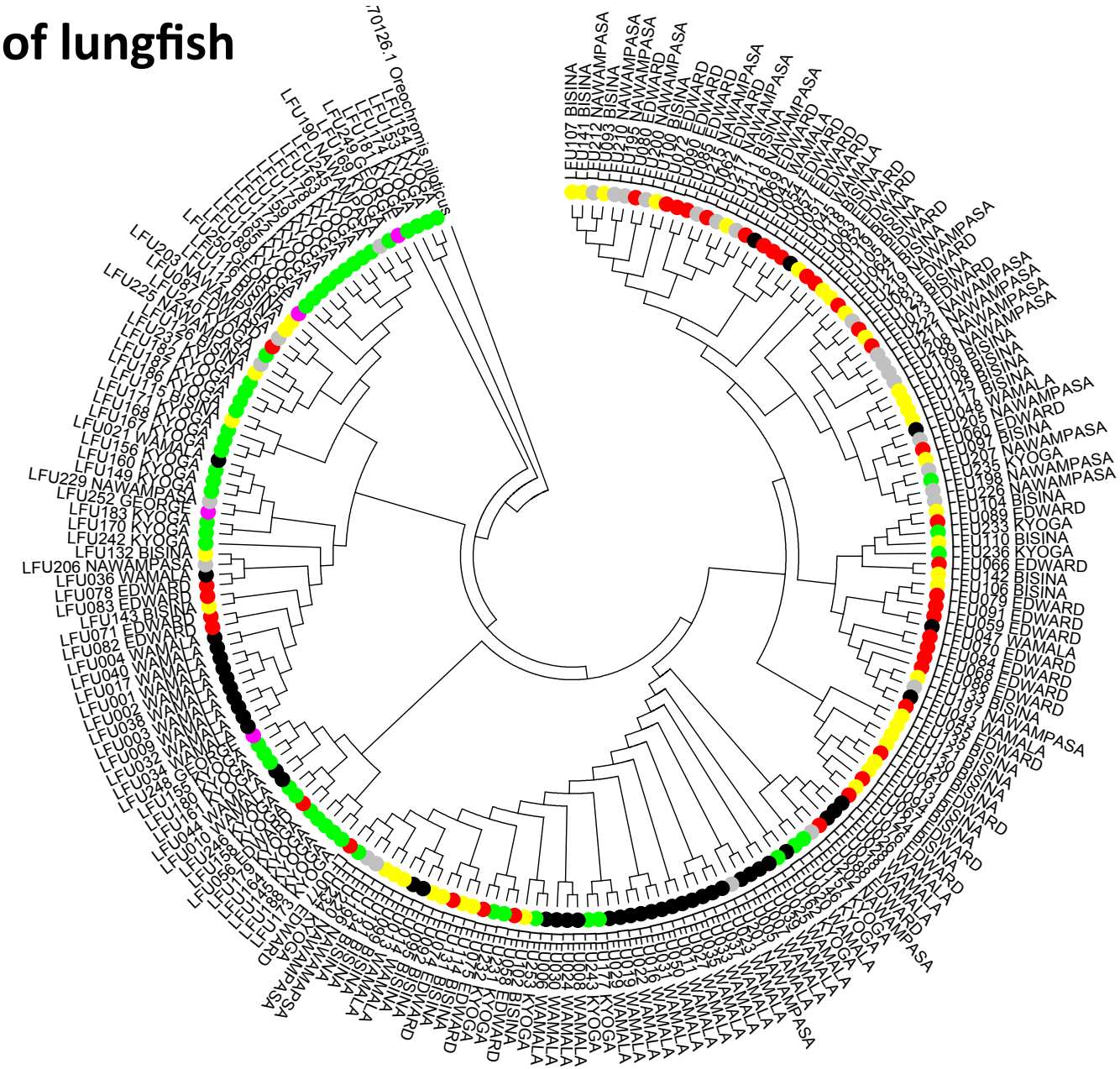
Population	Sample size	Number of Haplotype, H	Polymorphic sites	Haplotype Diversity, Hd, (±SD)	Fu, F
Overall	205	71	843	0.964 ± 0.005	- 35.059
Bisina	41	28	35	0.968 ± 0.015	- 11.185
Edward	38	27	33	0.976 ± 0.012	- 12.409
Kyoga	50	22	33	0.923 ± 0.020	- 4.073
Nawampasa	26	14	27	0.929 ±0.028	- 1.280
Wamala	45	13	27	0.775 ± 0.051	2.001

1) Tajima's, D: (-) low frequency polymorphisms relative to expectation (p> 0.100)

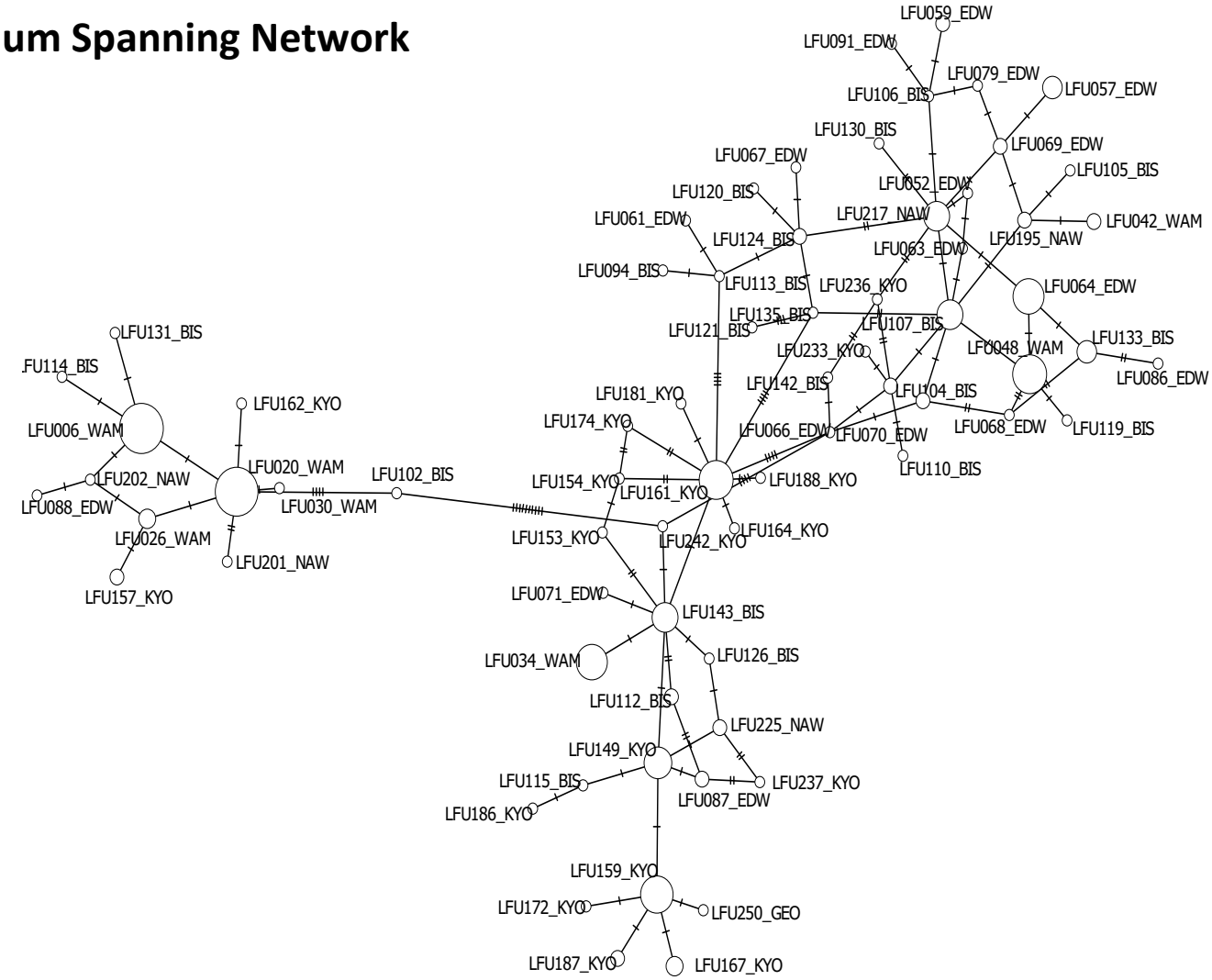
2) Fu, F: (-) excess number of alleles, as expected from a recent population expansion

3) Nucleotide (Pi): Less diversity = 0.010 ± 0.001

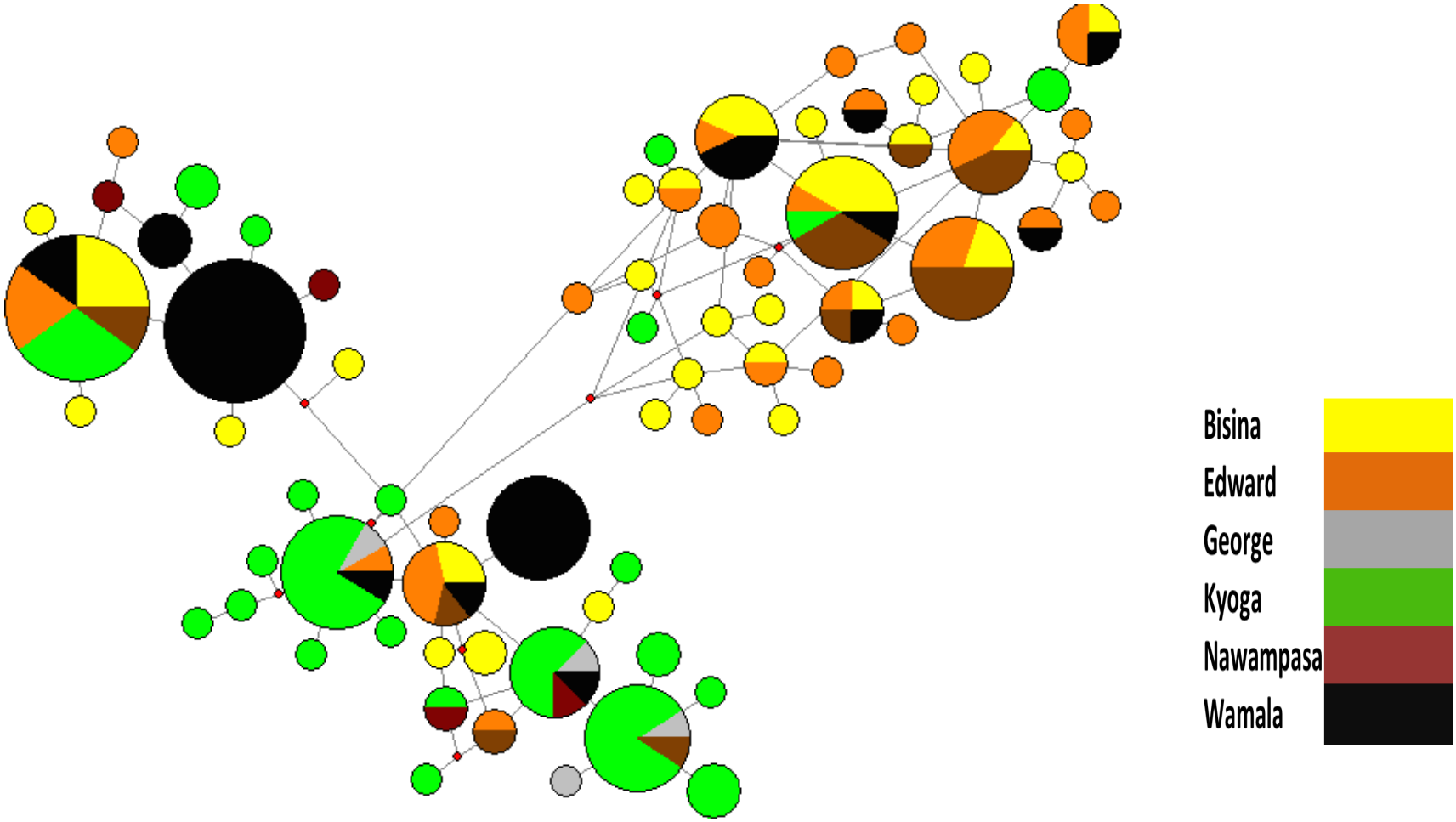
Phylogeny of lungfish



Minimum Spanning Network



ADMIXTURE, RECURENT AND HISTORICAL GENEFLOW



Pairwise Distance(F_{st})

	Wamala	Edward	Bisina	Kyoga	Nawampasa
Wamala	0.000				
Edward	0.295	0.000			
Bisina	0.237	0.011	0.000		
Kyoga	0.233	0.256	0.209	0.000	
N'pasa	0.312	0.014	0.010	0.247	0.000

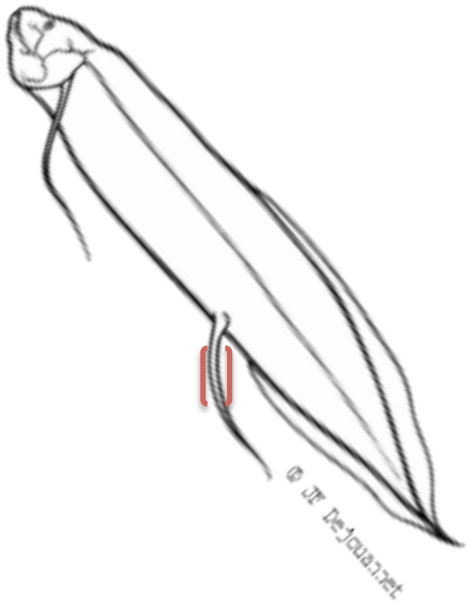
Nuclear diversity of lungfish using RNA Seq.

1. Largest genome; 133 billion bp of (Thomson, 1972),
2. Diploid 2n = 28 chromosomes (Omer and Abukashawa, 2012)

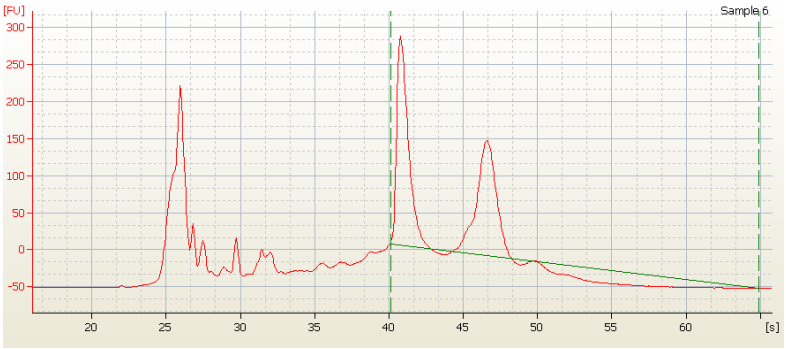
Challenges;

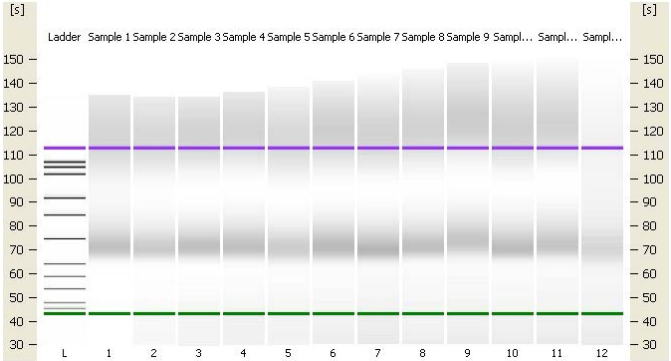
1. SNPs diversity using GBS; high throughput
2. lack of microsatellite/SSR markers

Nucleotide diversity: SNP panel using RNA Seq



Total RNA from 30/60 : Trizol protocol

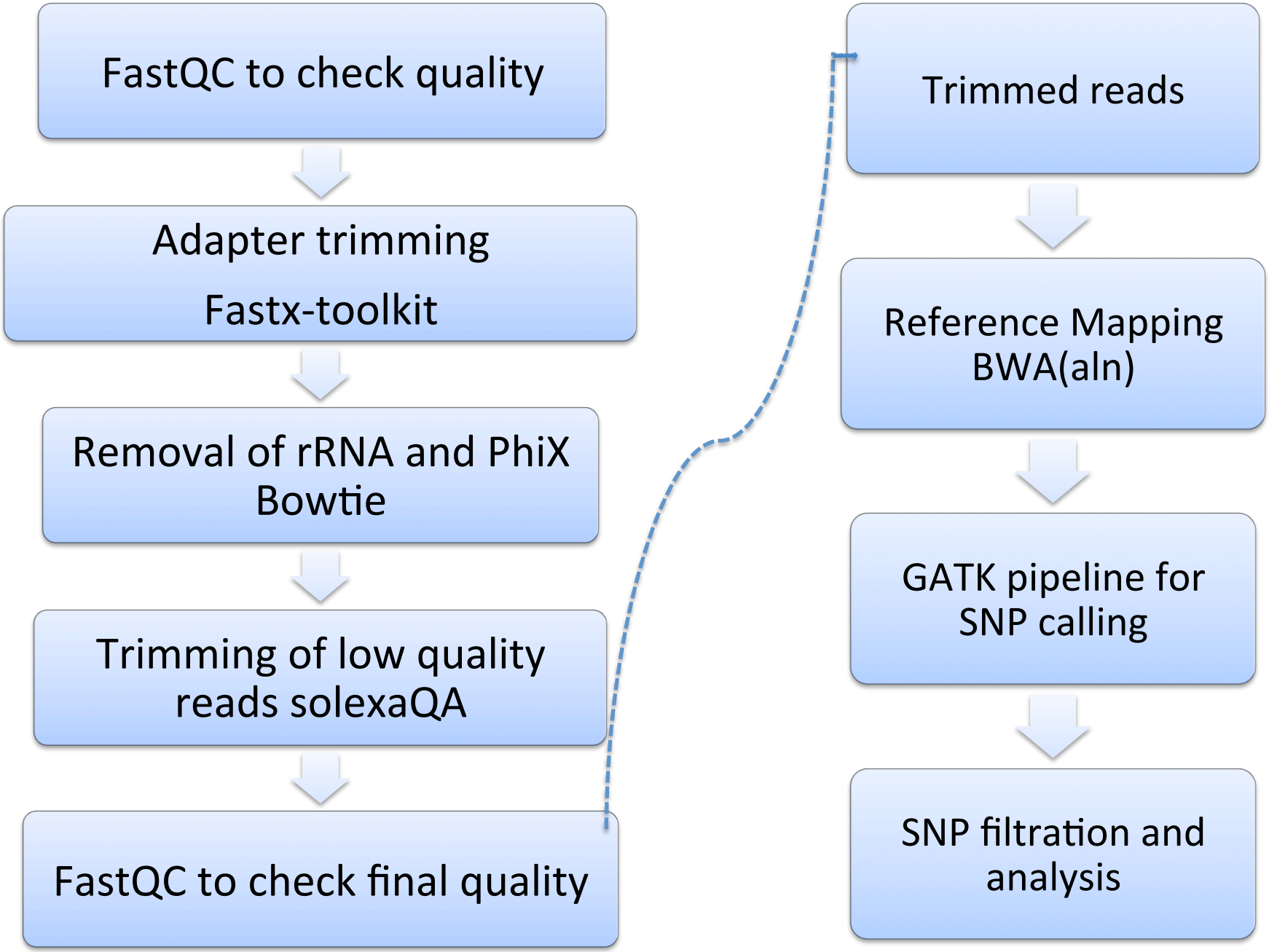




cDNA, MiSeq platform

mRNA from 18 samples (3/lake);
QubIt: 10.4 - 46.7, BIOANALYSER

44,375,286 reads; Paired end 2X 300;
35 – 301 bp


FastQC to check quality
Adapter trimming
Fastx-toolkit
Removal of rRNA and PhiX
Bowtie
Trimming of low quality
reads solexaQA
FastQC to check final quality
Trimmed reads
Reference Mapping
BWA(aln)
GATK pipeline for
SNP calling
SNP filtration and
analysis

Summary

	Reads
Total number of reads: Raw	80,358,264
Total number of reads: clean	75,647,110
Mapped to Reference	64,666,158
Mapped read Ratio	85.5%
Average GC content after Trimming	43.4%

Dataset	Mapped reads	Raw reads	Mapped reads ratio
Transcriptome		75,647,110	

Summary of SNPs discovered

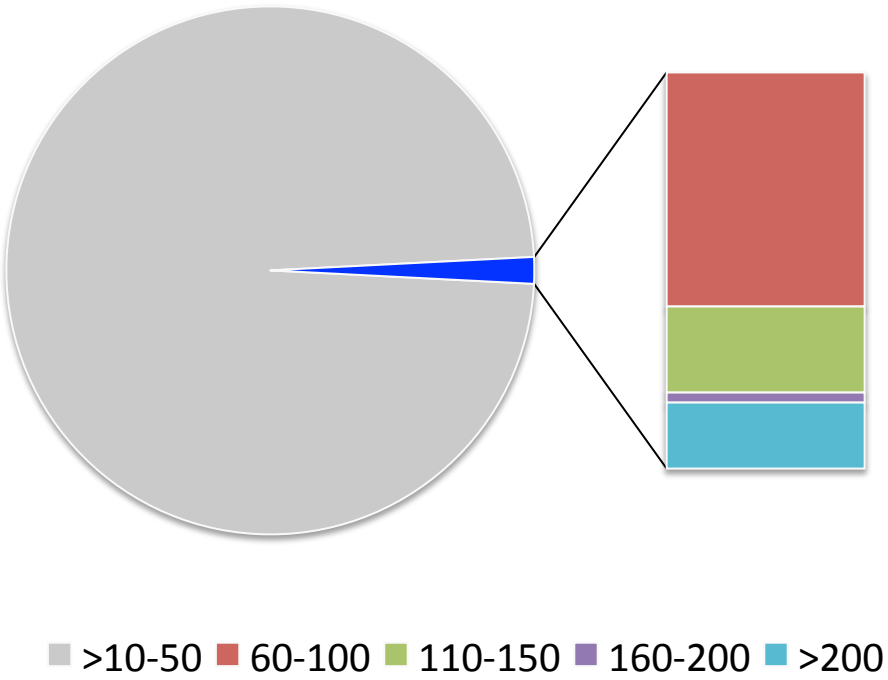
Type	**Transition**		**Transvertion**		
	GA	CT	GC	GT	AC
Number	2433	2614	1374	1613	1912
Percentage (%)	20.1	21.6	11.4	13.3	15.8
Total	5047		7038		

Summary SNPs

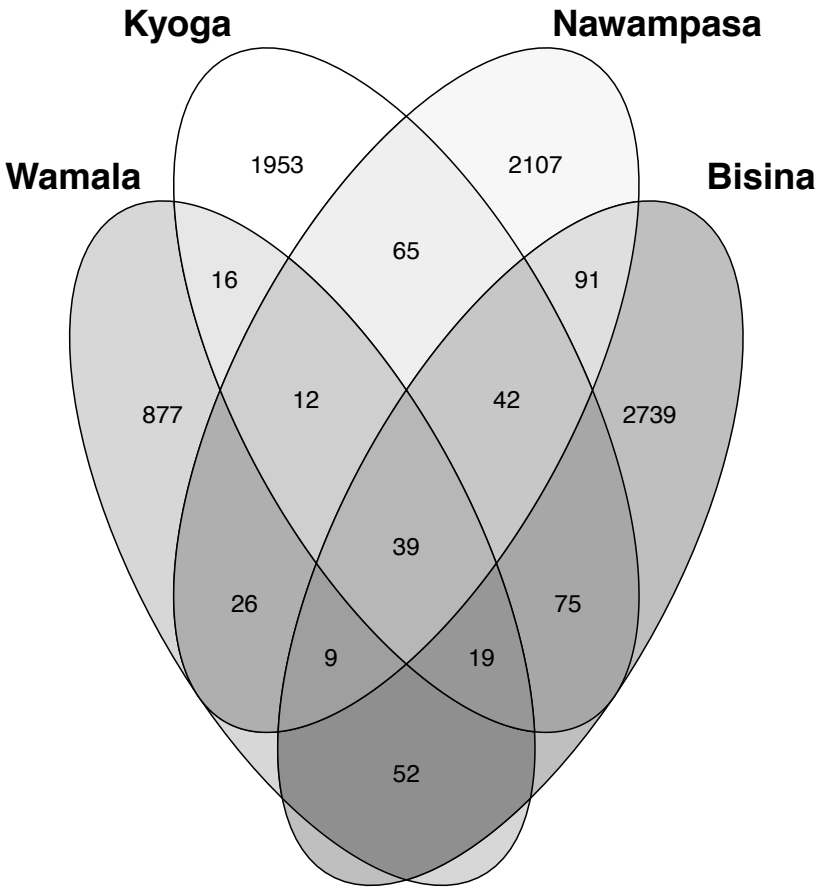
Lake	No. of SNPs
Wamala	831
Edward	1430
kyoga	1811
Nawampasa	2009
George	2070
Bisina	2602

Statistics of Read Depth in SNPs

SNP Read Depth	Number of SNPs
>10-50	11,984
60-100	117
110-150	43
160-200	5
>200	33



SNP panel


Kyoga
Nawampasa
Wamala
Bisina
1953
2107
65
16
91
877
12
42
2739
39
26
75
9
19
52

Conclusion

- Lungfish highly diversified within population based on Haplotype diversity
- Lake Bisina is more diversified & is located in Ramsar protected area; Policy implication
- Lake Bisina ideal source for rejuvenating parent fish stocks
- A total of **12085 SNPs detected**

What next?

- Fully understand the recurrent & historical geneflow of lungfish from great lakes of Uganda
- Annotation of SNPs detected
- Discovery more SNPs through *Denovo* pipeline
- Validate these markers
- Sex determination through GBS plate-form
- Lungfish breeding project into national Research program
- Publications

Acknowledgements

BecA-ILRI-TEAM;

- Josephine Birungi (analysis of mt DNA)
- Joyce Njuguna (Bioinformatics),
- Mercy Macharia (cDNA preparation)
- Jackline Kipkoech (Intern_RNA, CO1 data),
- Gloria Konyino (mtDNA extraction),
- Moses Ogugo (RNA extraction),
- Collins Mutai (mtDNA preparation)

AQUAFISH INNOVATION LAB-USAID

NaFIRRI-NARO-UGANDA

THANK YOU

The AquaFish Innovation Lab is supported in part by United States Agency for International Development (USAID) Cooperative Agreement No. EPP-A-00-06-00012-00 and by contributions from participating institutions.