



Are we underestimating the number of plant species in the tropics?

New insights from population genetics approaches applied on African forest trees

Olivier J. Hardy, Kasso Daïnou, Armel Donkpegan, Jérôme Duminil, Eben-Ezer Ewedjé,
Davy Ulrich Ikabanga, Brandet-Junior Lissambou, Rosalía Piñeiro



Evolutionary Biology and Ecology Unit

Funding:



Outline:

- Genetic signatures in the course of a speciation process
- Species delimitation in the tropical African flora
- Do we underestimate the number of tree species in Africa?

Estimating species richness

- ⇒ Often important for community ecology and conservation biology
- ⇒ Difficult to achieve in hyper-diverse communities (e.g. tropical forest flora)
 - because (1) incomplete sampling (missing rare species)
 - (2) lack of taxonomic revision
 - (3) problem of species delimitation / species concept

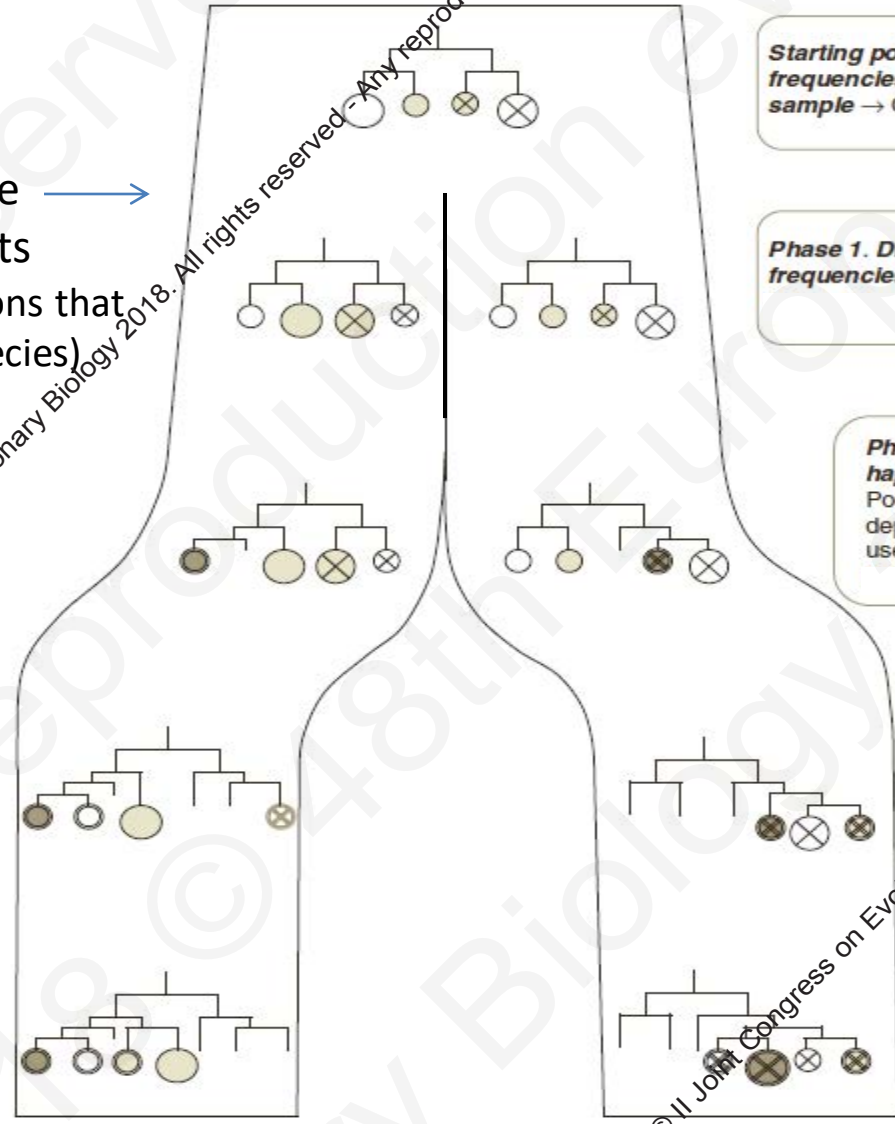
The large majority of plant species are delimited based only on morphological criteria

Genetic data can potentially confirm/infirm 'taxonomic' species, or reveal cryptic species

**Question: Have plant taxonomists been excessive splitters or lumpers
in the light of genetic data, in the case of tropical trees?**

Speciation process and gene tree

Reproductive isolation starts
(2 isolated populations that will become 2 species)



Starting point. One set of allele frequencies from any representative sample → One gene cluster.

Phase 1. Differentiation of allele frequencies → Different gene clusters.

Phase 2. Extinction of haplotypes + mutations → Polyphyly or monophyly depending on the sequence used.

Phase 3. Extinction of lineages + mutations → One monophyletic group with possibly mutual allelic exclusivity.

Phase 4. Extinction of lineages + mutations → Reciprocal monophyly

2018 © II Joint Congress on Evolutionary Biology 2018. All rights reserved - Any reproduction even in part is prohibited.

2018 © II Joint Congress on Evolutionary Biology 2018. All rights reserved - Any reproduction even in part is prohibited.

Reciprocal monophyly => criterion behind the genealogical species concept

Genetic approaches to infer speciation

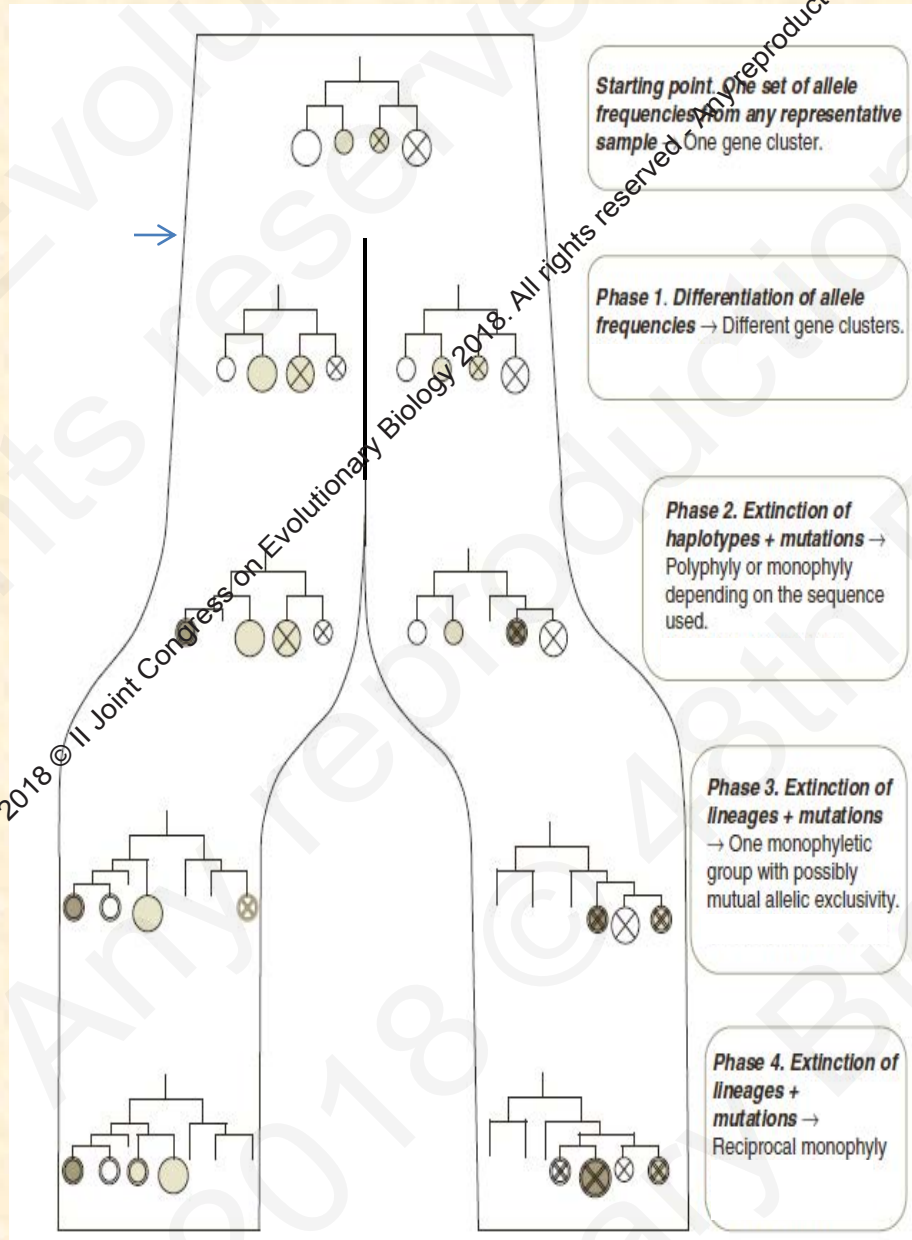
Many shared alleles -> **population genetics approach (e.g. clustering)** => can confound true species (reproductive barrier) from currently isolated populations (that could mix if they meet)

Many unique alleles and few remaining shared alleles

No shared alleles but incomplete lineage sorting -> gene sequencing (at several loci) and **haplotype approach**

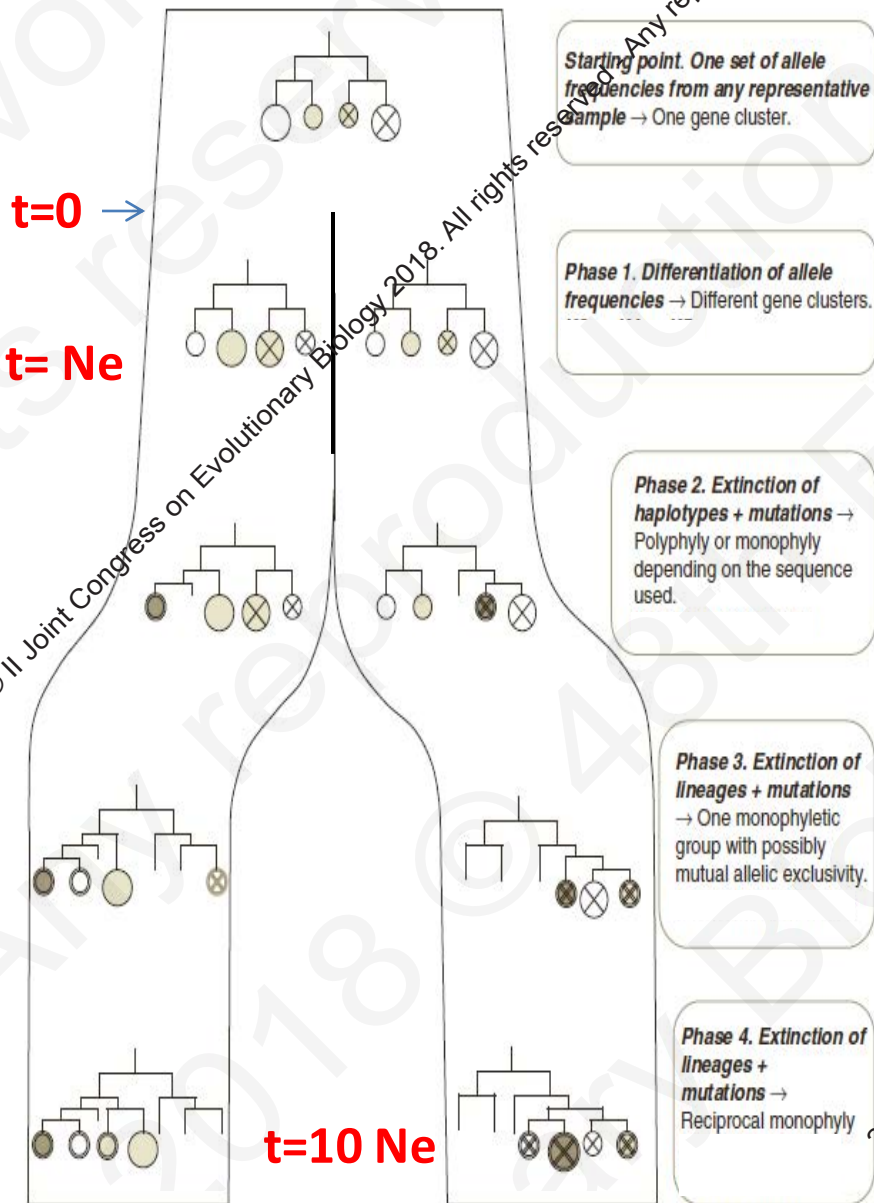
Flot et al. 2010 BMC Evol Biol

Reciprocal monophyly -> gene sequencing + **phylogenetic approach** sufficient

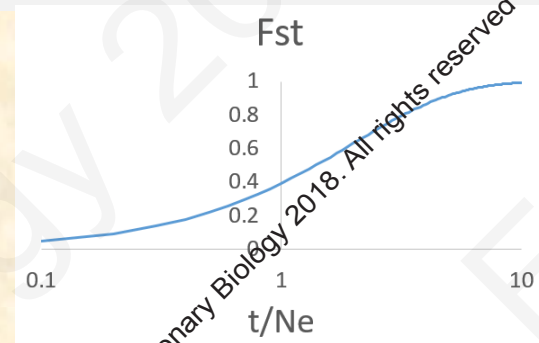


Time scales of speciation

t (number of generations since speciation)
depends on effective population sizes (N_e)



Allele frequencies differentiation
→ mostly drift effect (depends on N_e of each descending population).
Differentiation becomes high when $t > N_e$ generations.



Reciprocal monophyly
 $t > 10N_e$ generations ($t > 5N_e$) for 90% (50%) of loci showing recip monophyly
(Hudson & Coyne 2002, *Evolution* 56:1557–1565)

Time scales of speciation in long-living organisms (trees)

Reciprocal monophyly (=phase 4) sometimes taken as criterion to define species but, in some organisms, it may never occur:

Trees in mature tropical forests => mean **generation time = 100 - 200 years**

If large populations maintained through evolutionary time (say $N_e = 10\,000$)

=> time to reach **reciprocal monophyly**: $10 * 100 * 10\,000 = 10 \text{ millions years!}$

=> new speciation events could occur before RM is reached

⇒ Limit the criterion of reciprocal monophyly to define species

⇒ **Need to consider methods able to distinguish species at previous phases**

⇒ **haplowebs (for phase 3)**

⇒ **sympatric genetic cluster approach (for phase 1 or 2)**

The “genetic cluster” approach

1°) Identify genetic clusters by nuclear genotypes of 100's samples (e.g. STRUCTURE)

2°) Check geographical distribution of genetic clusters + % admixed genotypes:

- **Sympatry & no (rare) admixture** -> **distinct species**
(reproductive isolation despite direct contact)



- **Parapatry & no (rare) admixture** -> **distinct species or recent secondary contact between conspecific populations**



- **Parapatry & gradual admixture** -> **secondary contact between conspecific pop or hybrid zone between species**



- **Allopatry** -> **isolated conspecific populations or distinct species (inconclusive)**



3°) For putative new species: check if diagnostic morphological traits occur

⇒ **useful for phase 1 & 2 of speciation process**

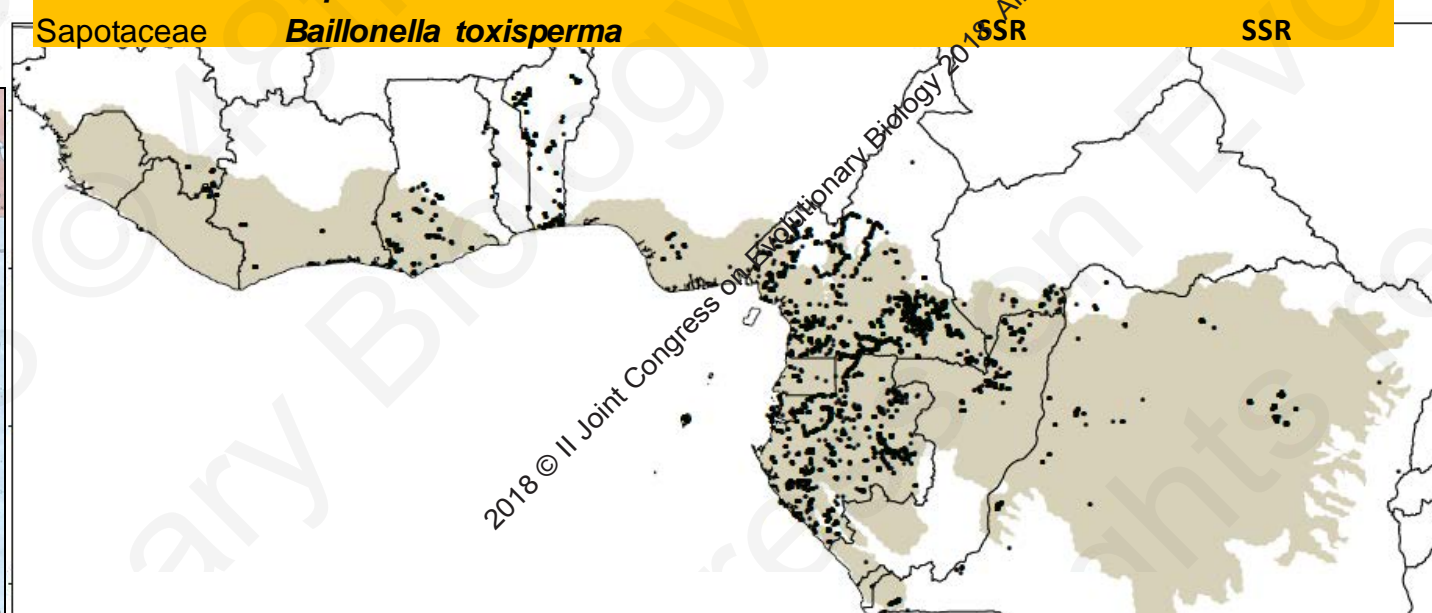
but risk of confounding isolated populations versus distinct species

2. Species delimitation in the tropical African flora

Tree species
with good
population
genetic data in
West and
Central Africa

Sampling

		Nuclear DNA	Plastid DNA
Annonaceae	<i>Greenwayedendron suaveolens</i>	SSR	SEQ
Burseraceae	<i>Santiria trimera</i>	SSR	SEQ
Caesalpinjiaceae	<i>Azelia spp</i>	SSR	SEQ
	<i>Distemonanthus benthamianus</i>	SSR	monomorph
	<i>Erythrophleum ivorense + suaveolens</i>	SSR + SEQ	SEQ + SSR
	<i>Guibourtia spp</i>	SSR	SEQ
	<i>Scorodophloeus zenkeri</i>	SSR	SEQ
Clusiaceae	<i>Symphonia globulifera</i>	SSR	SEQ + SSR
	<i>Pentadesma butyracea</i>	SSR	SEQ + SSR
Fabaceae	<i>Pericopsis elata</i>	SSR	monomorph
Meliaceae	<i>Entandrophragma spp</i>	SSR	SEQ
	<i>Khaya anthotheca</i>	SNP	SNP
Moraceae	<i>Milicia excelsa + regia</i>	SSR + SEQ + SNP	SEQ + SSR
Ochnaceae	<i>Lophira alata</i>	SSR	(SEQ)
Sapotaceae	<i>Baillonella toxisperma</i>	SSR	SSR



Are taxonomists excessive splitters ?

⇒ Compare congeneric species difficult to distinguish in the field or with very few diagnostic traits

	Plastid markers		Nuclear markers		
	Exclusive haplotypes	Reciprocal monophyly	Distinct gene clusters in sympatry	Exclusive haplotypes	Reciprocal monophyly
<i>Carapa</i> spp	No	No	Yes	?	?
<i>Haumania danckel./librechtsiana</i>	No	No	Yes	Yes	?
<i>Milicia excelsa/regia</i>	Yes	No	Yes	Yes	No
<i>Lophira alata/lanceolata</i>	Yes?	?	Yes	?	?
<i>Erythrophleum ivorense/suaveolens</i>	Yes	No	Yes	Yes	No

Apparently no

Are cryptic species frequent ?

Santiria trimera (Burseraceae)

Medium-size tree species
of mature forests

One or several species?

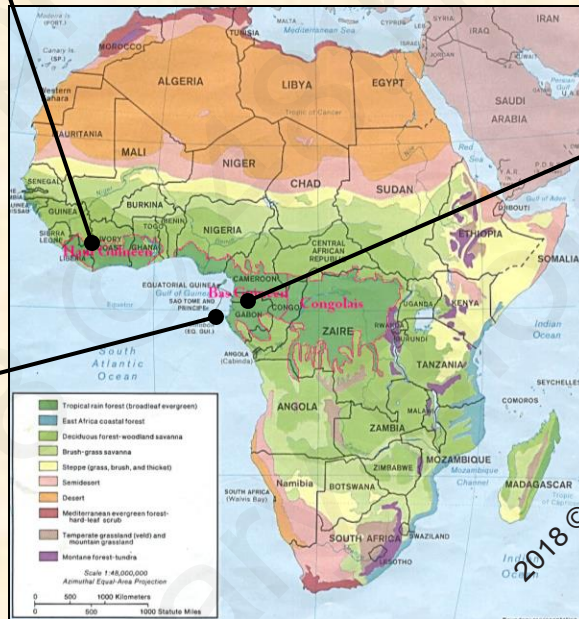
One species
recognized by Onana
(2009)



Upper Guinea



São Tomé

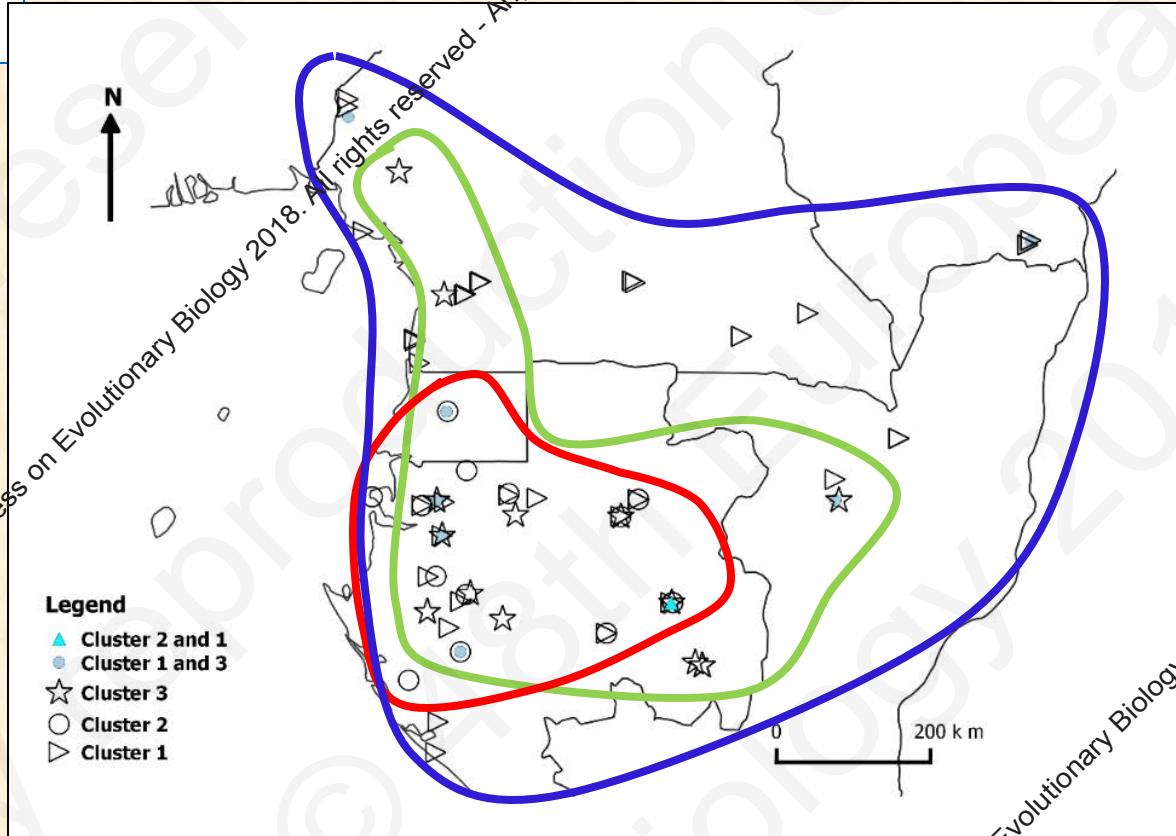


Gabon

Lower Guinea + Congolia
3 morphotypes in sympatry

Santiria trimera

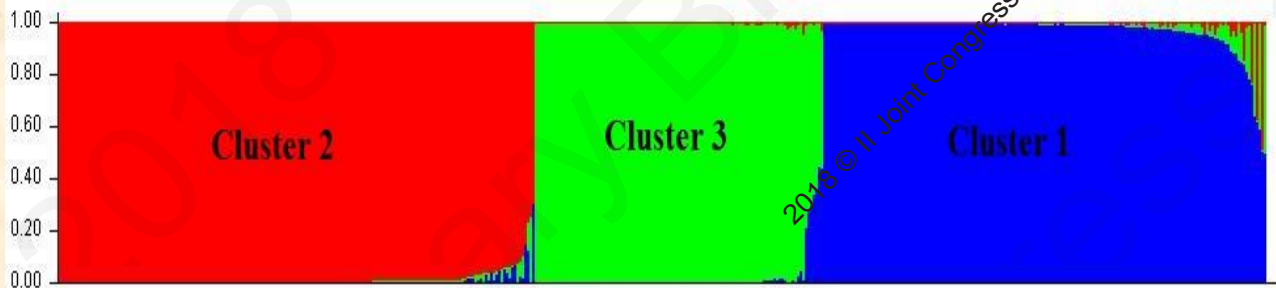
PhD thesis of Davy Ulrich Ikabanga



In Lower Guinea:
3 sympatric gene pools
=> 3 species

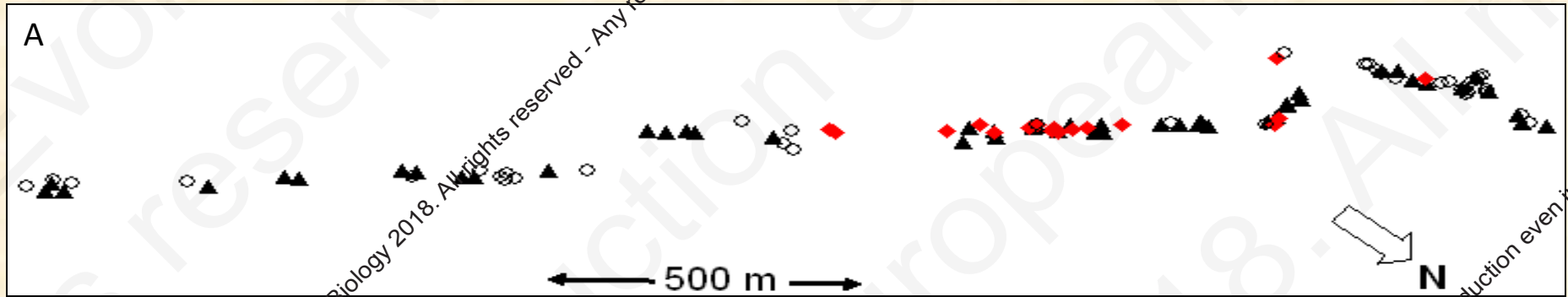
$F_{st} = 0.32 - 0.38$

SSR



Santiria trimera

PhD thesis of Guillaume Koffi



3 morphotypes:



Small leaflets
Stilt roots



Large leaflets
Stilt roots

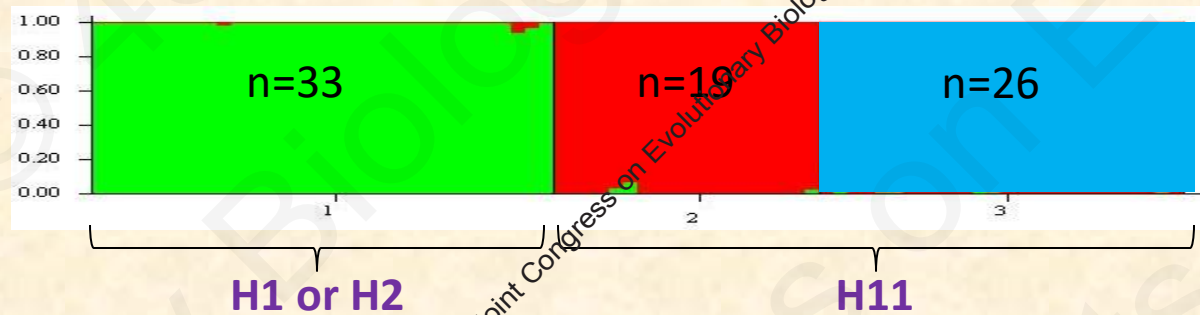


No stilt roots

Along a 3km long transect:

3 sympatric gene pools
> 3 species

Assignment to 3 genetic clusters (software STRUCTURE)



Chloroplast haplotype:

→ 3 species following the Biological Species Concept

↪ 2 of them not distinguished using chloroplast DNA

Santiria trimera – morphology

Trait	GC1 (N=33)	GC2 (N=43)	GC3 (N=10)
Flowers pubescent	no	yes	no
Color of immature fruit	Green	Red	Green
Color cotyledons	Green	Red	Red
Inflorescence	Axillary	Terminal	Terminal

=> Diagnostic combination of traits

The identification of “species” based on genetic markers allowed to identify distinctive morphological traits

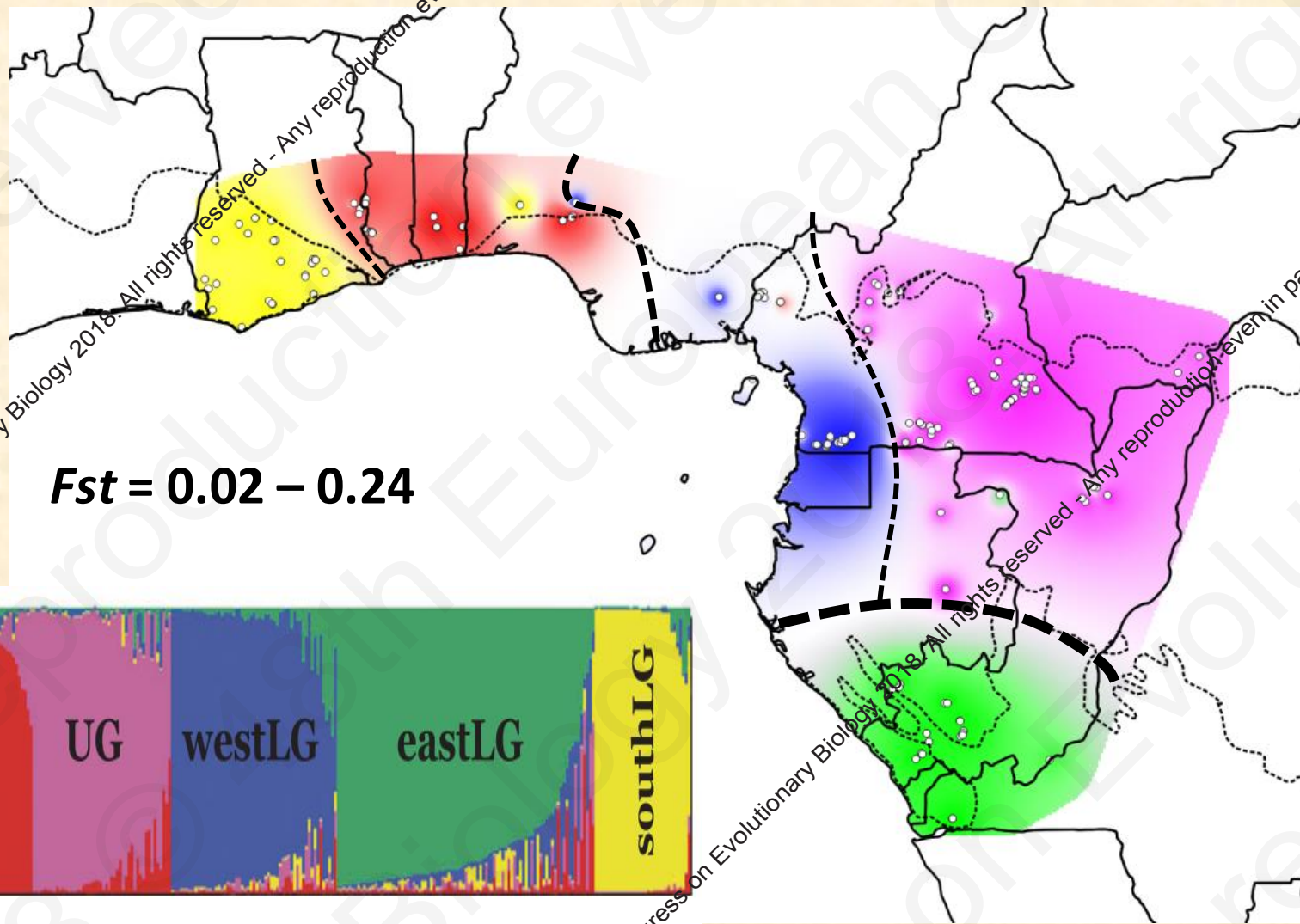
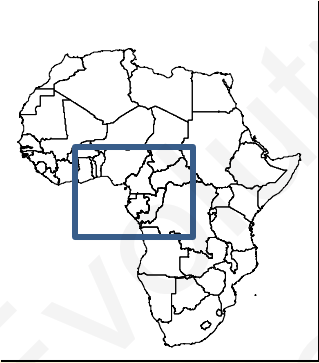
-> taxonomic delimitation supervised by nuclear genetic marker

**3 genetic clusters in sympatry
No or little admixture in contact zone
Diagnostic morphological traits**

3 species

Ikabanga et al. 2017, Phytotaxa

Terminalia superba (Combretaceae)



$F_{st} = 0.02 - 0.24$



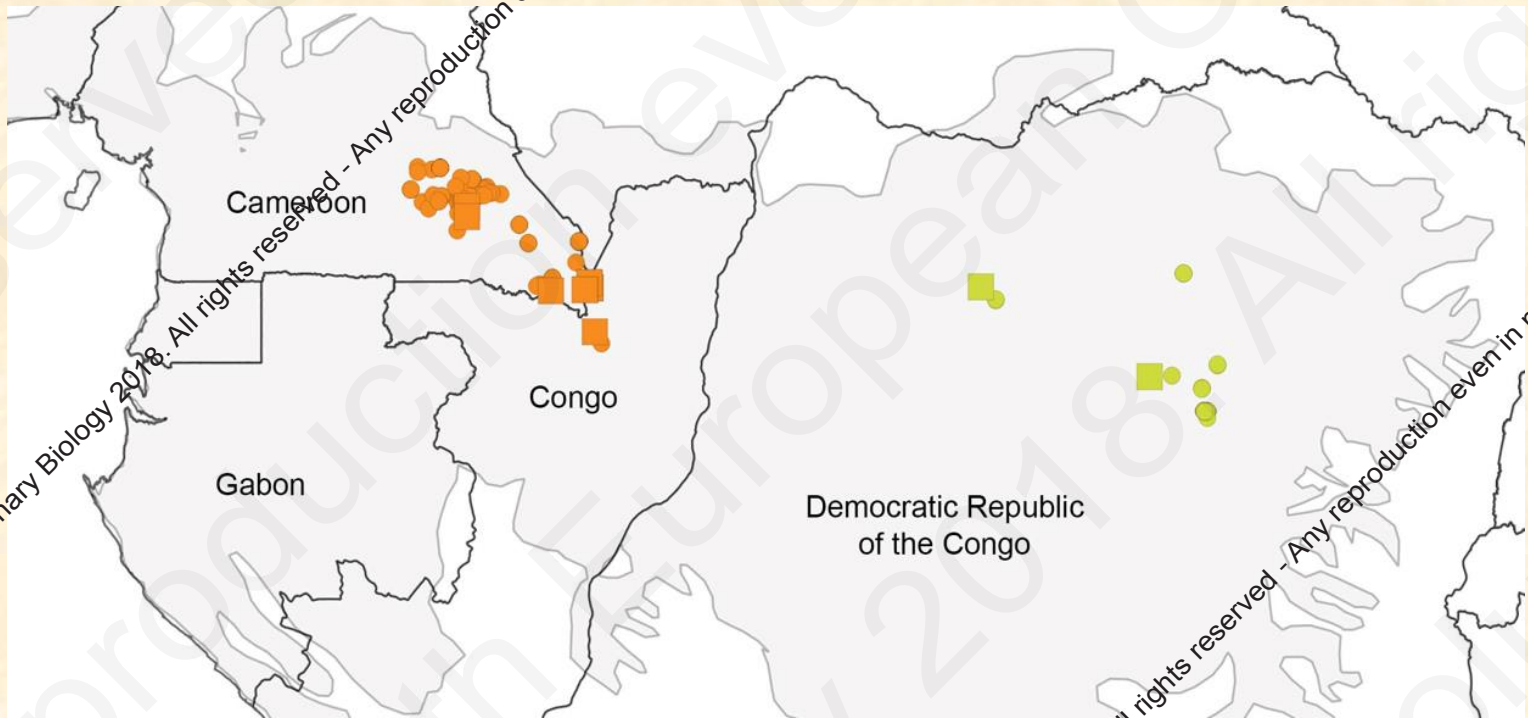
5 genetic clusters in parapatry

Admixture in contact zone

No known morphological differentiation

1 species with differentiated populations

Pericopsis elata (Fabaceae)



$F_{st} = 0.53$

STRUCTURE clustering (K=2)

2 genetic clusters in allopatry

No admixture

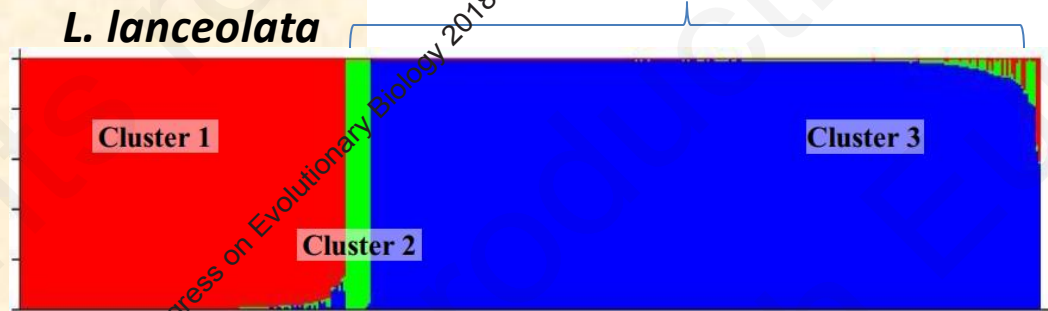
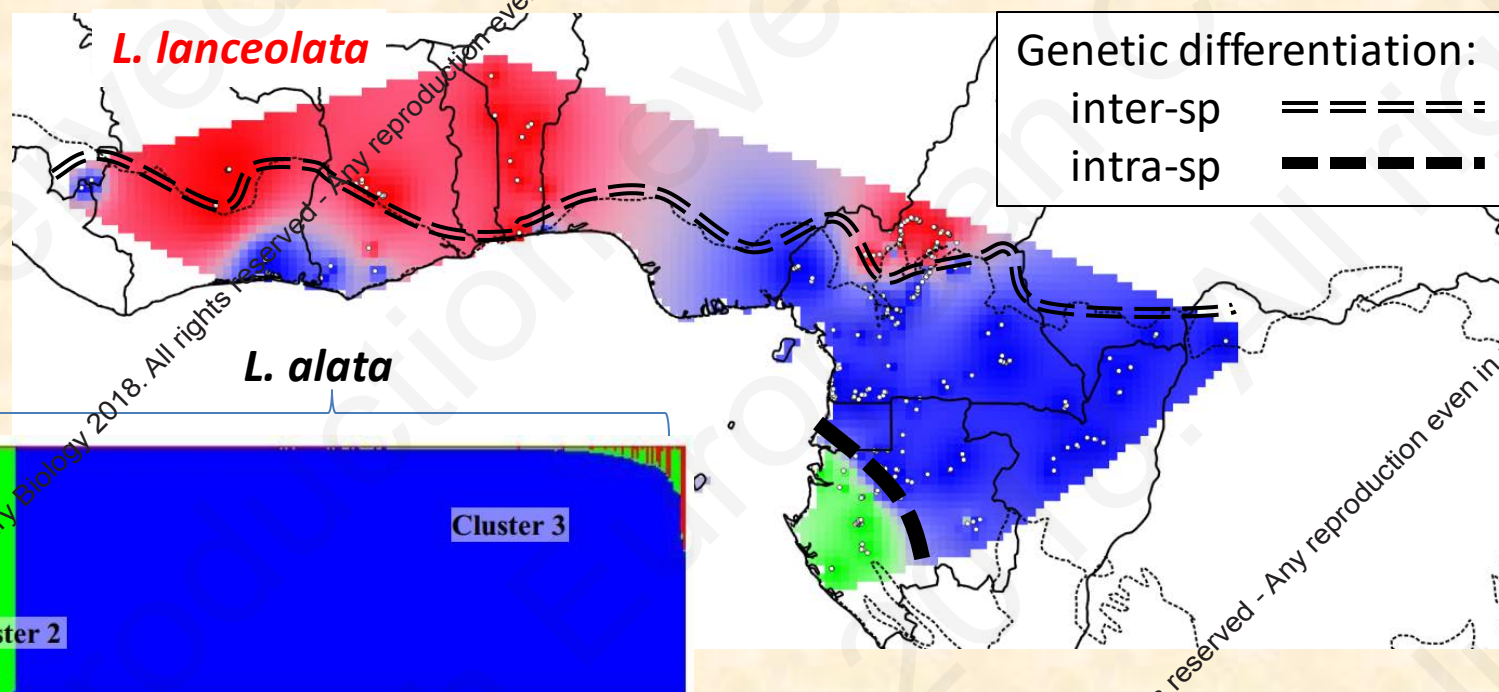
No known morphological differentiation

1 species

2018 © II Joint Congress on Evolutionary Biology 2018. All rights reserved - Any reproduction even in part is prohibited.

2018 © II Joint Congress on Evolutionary Biology 2018. All rights reserved - Any reproduction even in part is prohibited.

Lophira alata & *L. lanceolata* (Ochnaceae)



$F_{st} \text{ } L. \textit{alata} - L. \textit{lanceolata} = 0.38 \cong F_{st} \text{ between clusters } L. \textit{alata} = 0.33$

Genetic clusters in parapatry
 Rare admixture
 Morphological differentiation

2 species confirmed
 (*L. alata* / *L. lanceolata*)

2 genetic clusters in parapatry
 Rare admixture
 No known morphological differentiation

1 or 2 species within *L. alata*

Are we underestimating the species richness of trees in tropical Africa?

Taxonomic species	# BSC species	Diagnostic morphological trait
<i>Aucoumea klaineana</i>	1	
<i>Baillonella toxisperma</i>	1	
<i>Distemonanthus benthamianus</i>	1	
<i>Entandrophragma cylindricum</i>	1	
<i>Entandrophragma utile</i>	1	
<i>Erythrophleum ivorense</i>	1	
<i>Greenwayodendron oliveri</i>	1	
<i>Lophira lanceolata</i>	1	
<i>Milicia excelsa</i>	1	
<i>Pericopsis elata</i>	1	
<i>Scorodophloeus zenkeri</i>	1	
<i>Symphonia globulifera</i>	1	
<i>Terminalia superba</i>	1	
<i>Anthonotha macrophylla</i>	1-2	?
<i>Erythrophleum suaveolens</i>	1-2	?
<i>Lophira alata</i>	1-2	?
<i>Milicia regia</i>	1-2	?
<i>Pterocarpus santalinoides</i>	1-3	?
<i>Afzelia bipindensis</i>	2	Yes
<i>Entandrophragma angolense</i>	2	Yes
<i>Staudtia kamerunensis</i>	2-4	?
<i>Santiria trimera</i>	3-5	Yes
<i>Greenwayodendron suaveolens</i>	5	Yes
<i>Khaya anthotheca</i>	4-5	Yes

A sample of **24 tree species** as currently recognized by taxonomists



36 to 47 species following the Biological Species Concept according to new genetic data

- > Donkpegan 2017, PhD thesis
- > Monthe et al., in press, Tree Genetics & Genomes
- > Ikabanga et al. 2017, Phytotaxa
- > Lissambou et al. in press, PhytoKeys, Taxon
- > Bouka 2017, PhD thesis

Tree species richness in tropical African forests:

RAINBIO database: **3013 spp** recorded, c. 50% endemic (Sosef et al. 2017 BMC Biology)

Underestimated because (1) unsampled rare species

(2) “cryptic” species

Are our 24 species with population genetic data representative?

- different families, different ecology, not chosen for being species complexes
- **widespread**

⇒ Do species with narrow distribution contain fewer cryptic species ???

If not, the ‘actual’ number of tree species could be closer to **6000-8000 spp**

Conclusions

- (i) In general, plant taxonomists => not excessive splitters
- (ii) Chloroplast markers often not reliable to distinguish congeneric species
=> limits of plastid based DNA barcodes!
- (iii) Reciprocal monophyly => very conservative criterion to delineate species
(genealogical species concept) => not recommended for tree species
- (iv) “Cryptic” species common in the African flora, at least in widespread trees
=> may double the number of species
- (v) Putative “cryptic” species often reveal diagnostic characters => are not
cryptic *sensu stricto* => population genetics data can guide taxonomists

Thank you for your attention