

Chromosome evolution and the genetic basis of agronomically important traits in greater yam (*Dioscorea alata*)



Jessen Bredeson
Plant Comparative Genomics
DOE-JGI, LBNL
(fmr. University of California, Berkeley)

Read our publication to learn more



ARTICLE



<https://doi.org/10.1038/s41467-022-29114-w>

OPEN

Chromosome evolution and the genetic basis of agronomically important traits in greater yam

Jessen V. Bredeson ^{1,18}, Jessica B. Lyons ^{1,2,18}, Ibukun O. Oniyinde³, Nneka R. Okereke⁴, Olufisayo Kolade³, Ikenna Nnabue⁴, Christian O. Nwadili⁴, Eva Hřibová⁵, Matthew Parker⁶, Jeremiah Nwogha⁴, Shengqiang Shu ⁷, Joseph Carlson⁷, Robert Kariba^{8,9}, Samuel Muthemba ^{8,9}, Katarzyna Knop⁶, Geoffrey J. Barton ⁶, Anna V. Sherwood ^{6,16}, Antonio Lopez-Montes^{3,17}, Robert Asiedu ³, Ramni Jamnadass^{8,9}, Alice Muchugi^{8,9}, David Goodstein ⁷, Chiedozie N. Egesi^{3,4,10}, Jonathan Featherston¹¹, Asrat Asfaw ³, Gordon G. Simpson^{6,12}, Jaroslav Doležel ⁵, Prasad S. Hendre ^{8,9}, Allen Van Deynze ¹³, Pullikanti Lava Kumar³, Jude E. Obidiegwu ⁴✉, Ranjana Bhattacharjee ³✉ & Daniel S. Rokhsar ^{1,2,7,14,15}✉

<https://www.nature.com/articles/s41467-022-29114-w>

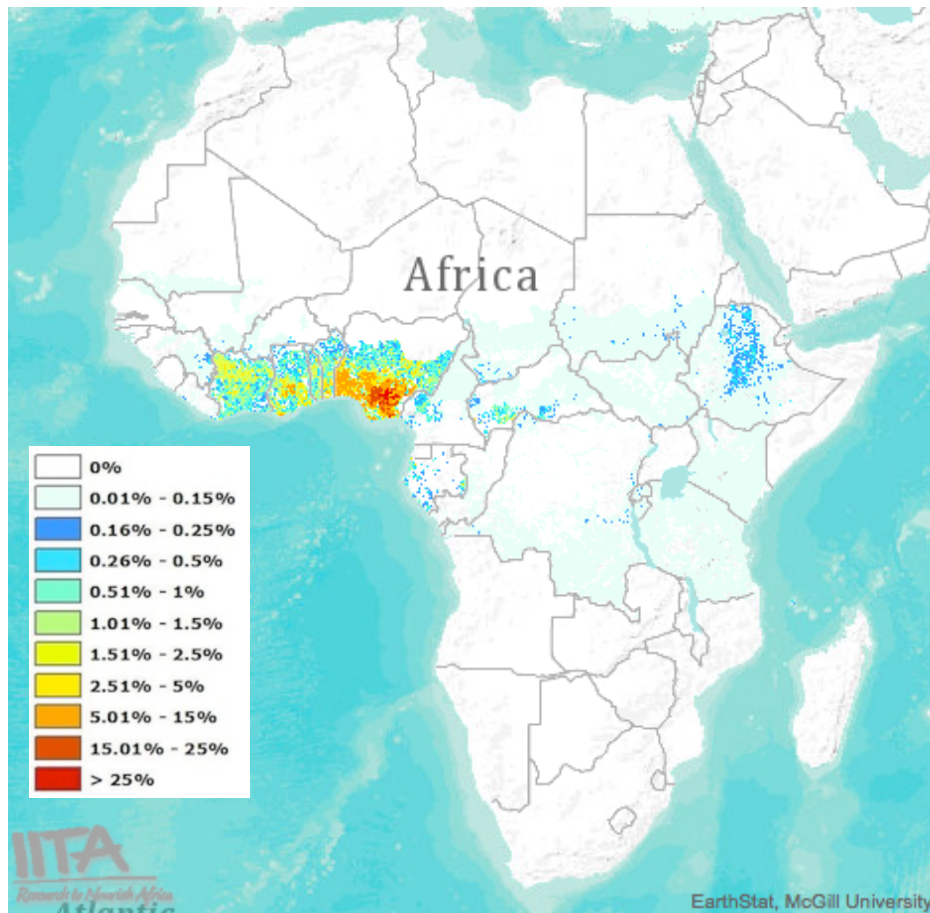
The greater yam, *D. alata*

- Water/winged/purple yam, ube (monocot, Dioscoreales)
- Most broadly cultivated *Dioscorea* species
- High nutritional content, low glycemic index
- Yields well in marginal soil
- Easily propagated, early vigor, low post-harvest losses
- Breeding difficult due to:
 - obligate outcrossing
 - long growth cycle
 - Poor/no flowering
 - polyploidy

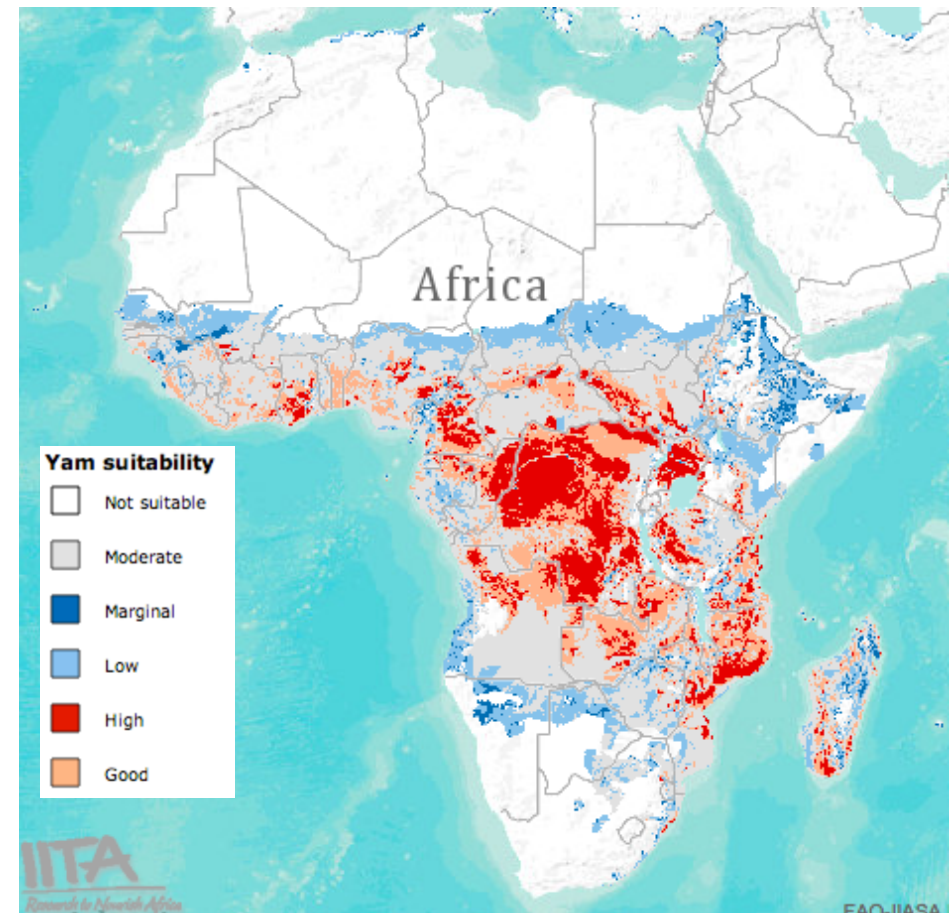


Huge potential for yam in Africa

Harvested area



Suitability index



Goal: Enhance food security for smallholder farmers

NSF BREAD Project aims:

- Accelerate the pace of improvement by:
 - Facilitating genome-enhanced breeding approaches
 - Connecting phenotype to genotype
 - Characterizing global genetic diversity
- Transfer advanced genomic technologies to African research institutions



TDa95/00328, a popular breeding line

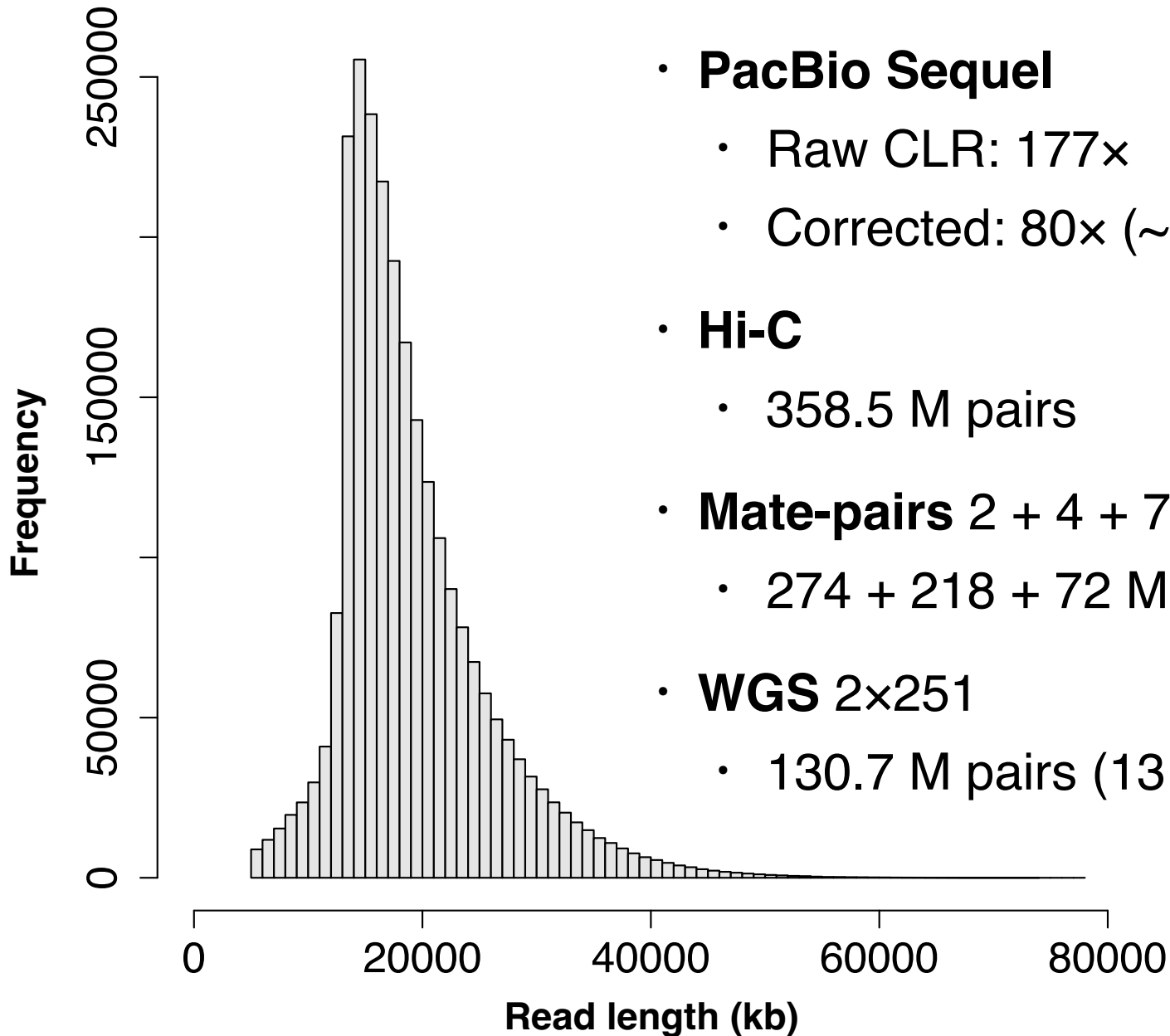


- Disomic inheritance
- Used in many crosses
- Anthracnose resistance

Anthracnose



Deep long reads and long-range linkages



- **PacBio Sequel**
 - Raw CLR: 177x
 - Corrected: 80x (~20kb)
- **Hi-C**
 - 358.5 M pairs
- **Mate-pairs 2 + 4 + 7kb**
 - 274 + 218 + 72 M pairs
- **WGS 2x251**
 - 130.7 M pairs (131x)



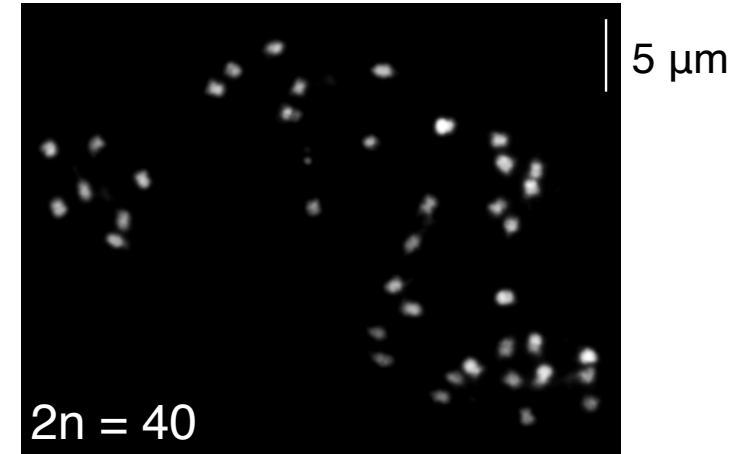
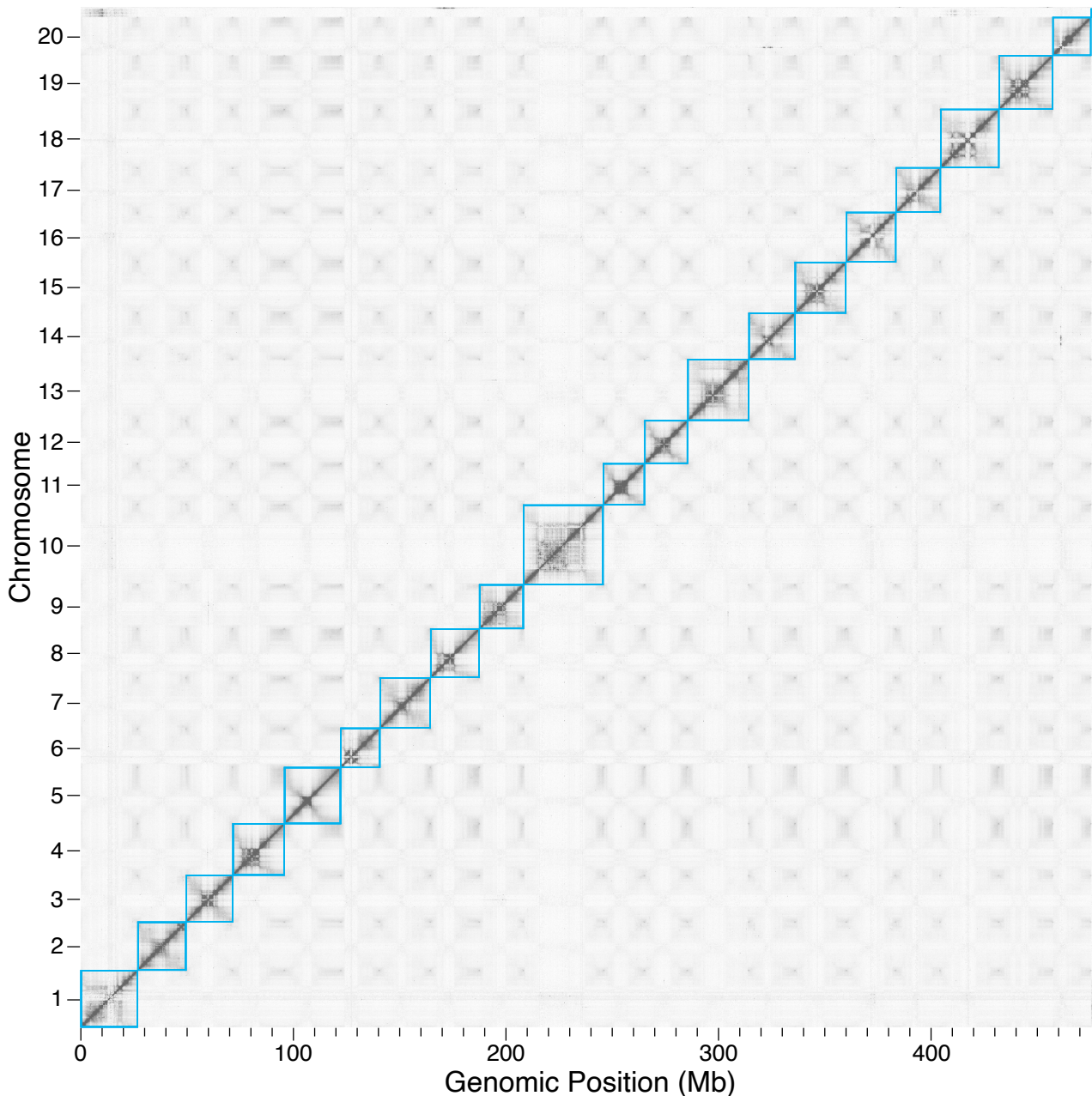
Prasad Hendre
(ICRAF, AOCC)



Jaroslav Doležel
(IEB Czech Rep.)

Eva Hřibová
(IEB Czech Rep.)

Assembly reconstructs 20 chromosomes



- Est. size: 455±39 Mb
- Contig: 479.5 Mb
- Chr: 475 Mb
- N50 length: 4.5 Mb

- 25,189 genes
- 38,603 total isoforms

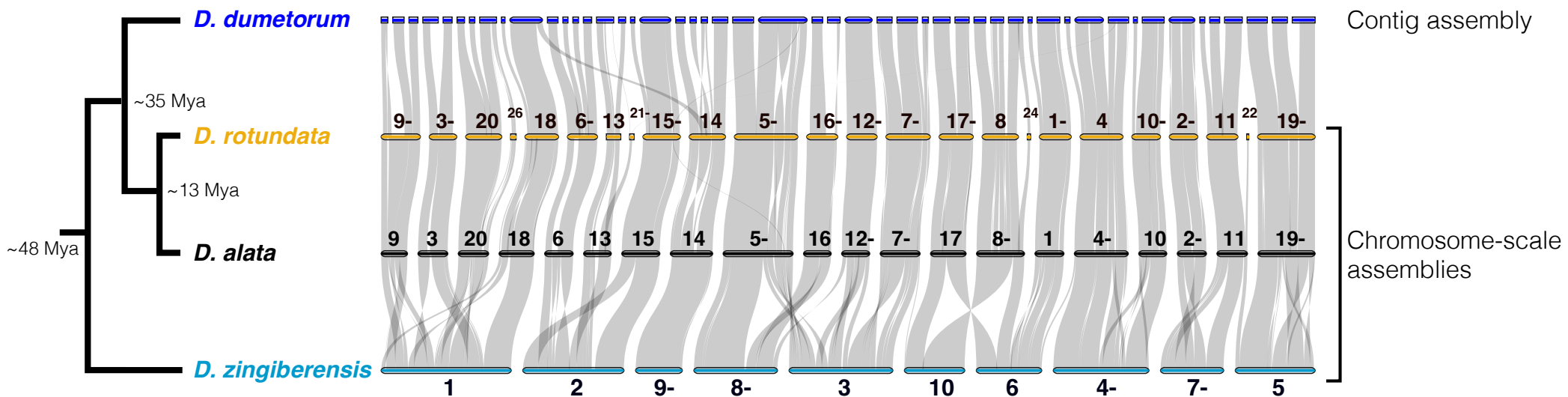
BUSCO (%)	Embryophyta	Liliopsida
Complete	97.8	93.4
Fragmented	1.5	3.8
Missing	0.7	2.8

Shengqiang Shu (DOE JGI)

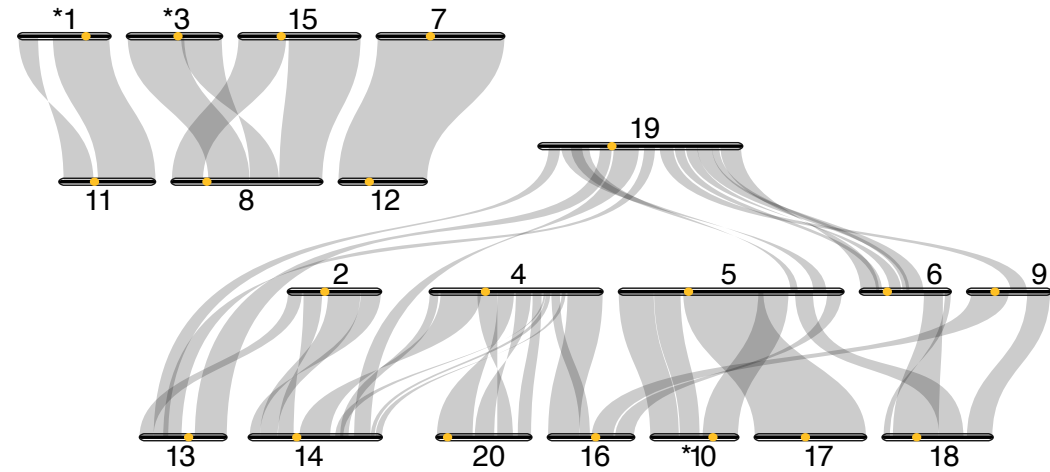
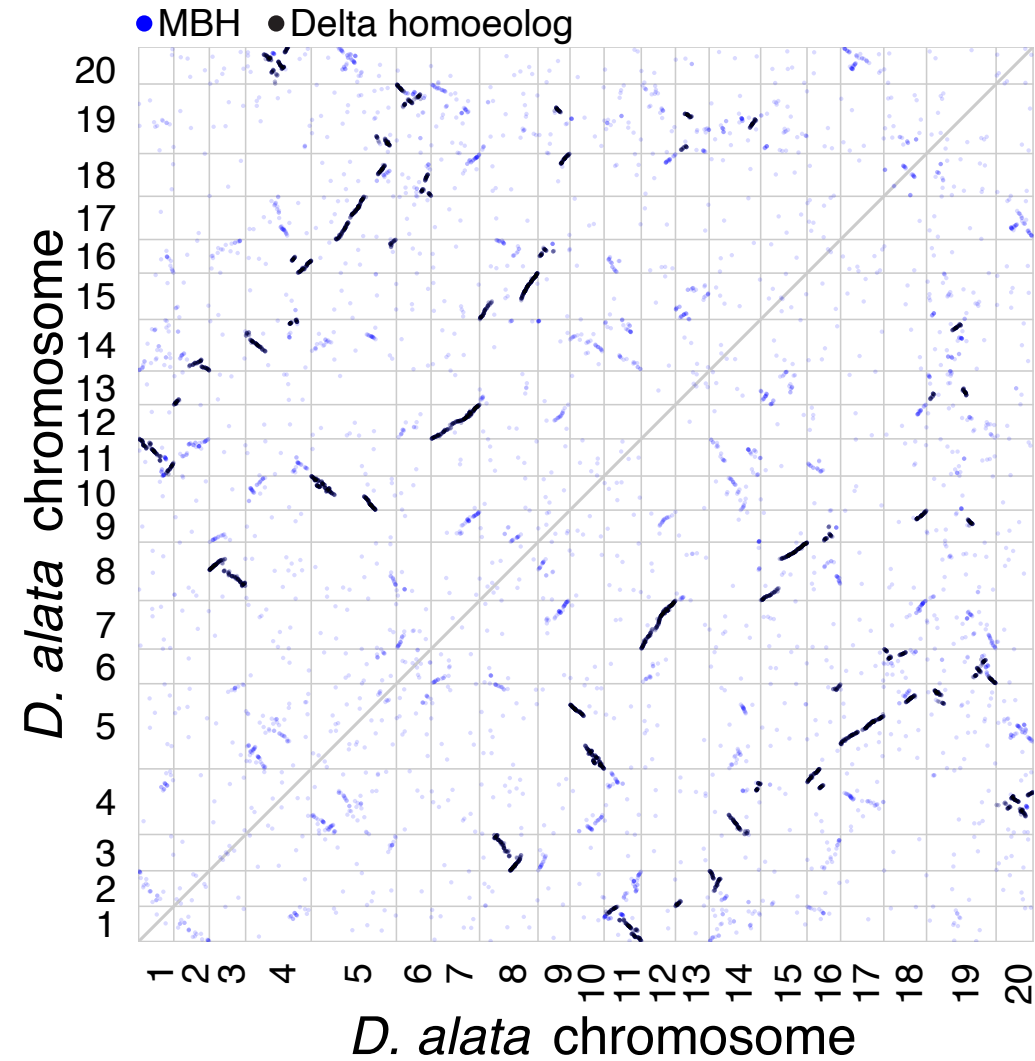
The background of the slide features a large, faint watermark of the University of California, Berkeley seal. The seal is circular and contains the text "UNIVERSITY OF CALIFORNIA BERKELEY" around the perimeter, "1868" at the bottom, and "THE GREAT SEAL OF THE UNIVERSITY OF CALIFORNIA" at the top. In the center of the seal is a shield with a book, a star, and the motto "LIGHT" and "THERE BE".

Inferring *Dioscorea*
genome evolutionary
history

Dioscorea 1:1 collinearity conserved

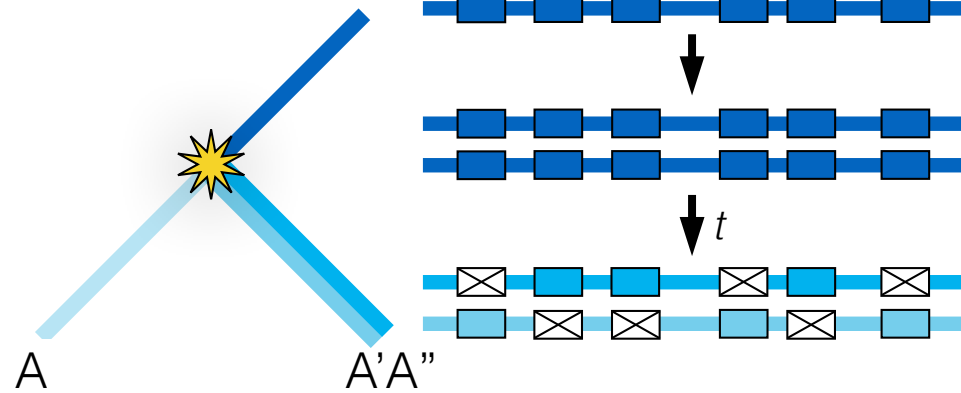


Dioscorea genomes share a WGD

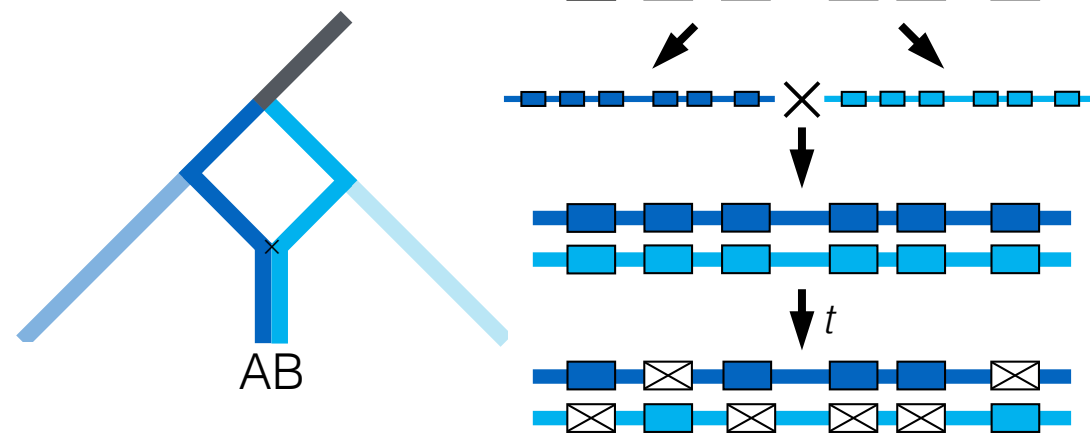


Delta: evidence of paleo-allotetraploidization

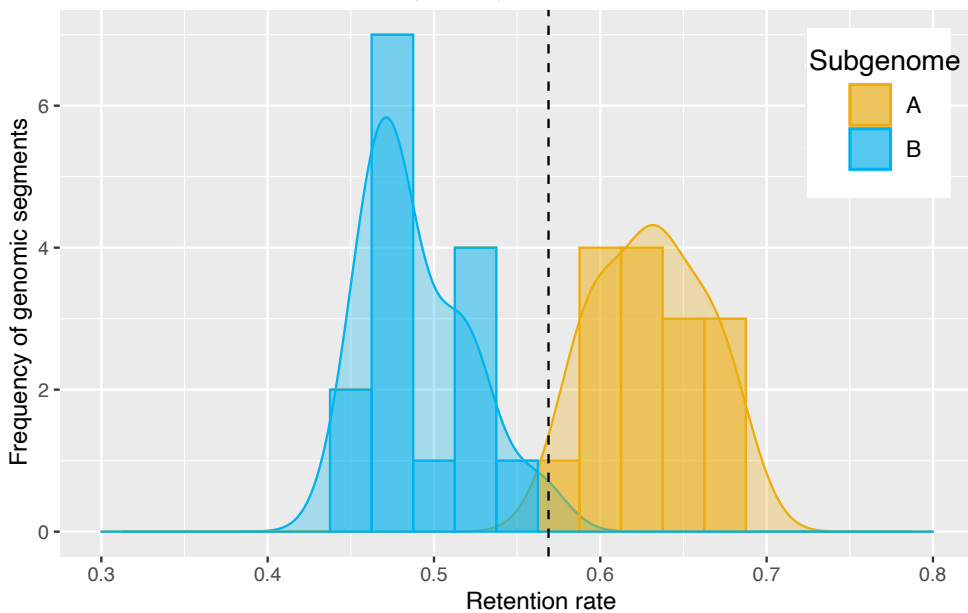
Autopolyploidy



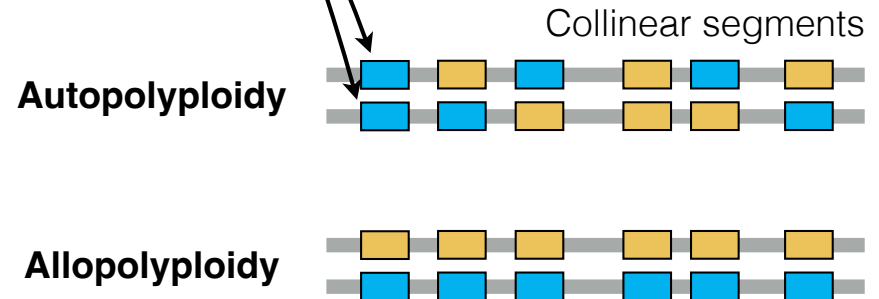
Allopolyploidy



$$r_{\text{retention}} = \frac{n_{\alpha} + n_{\alpha\beta}}{n_{\alpha} + n_{\alpha\beta} + n_{\beta}}$$

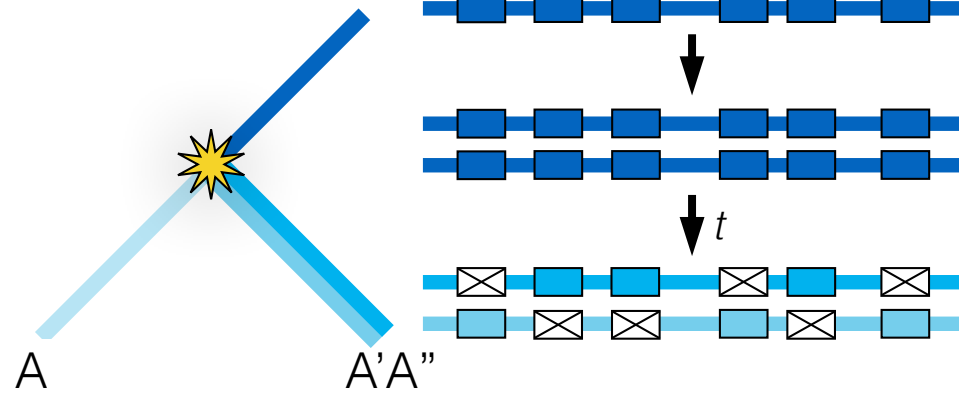


6 × Draw 2

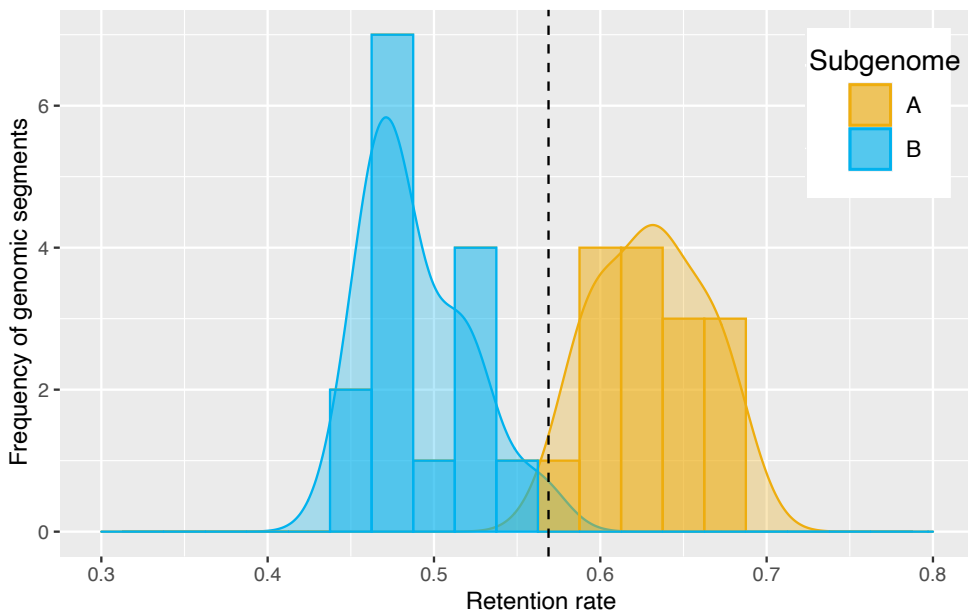
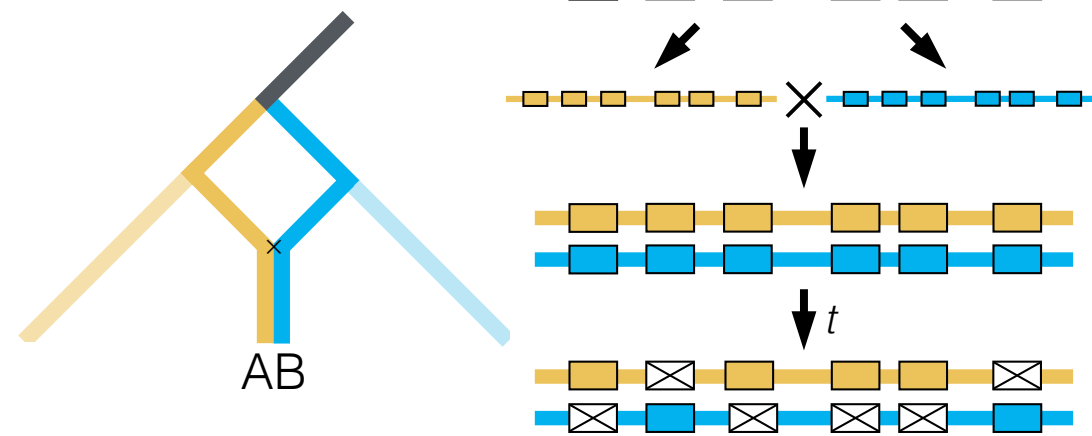


Delta: evidence of paleo-allotetraploidization

Autopolyploidy



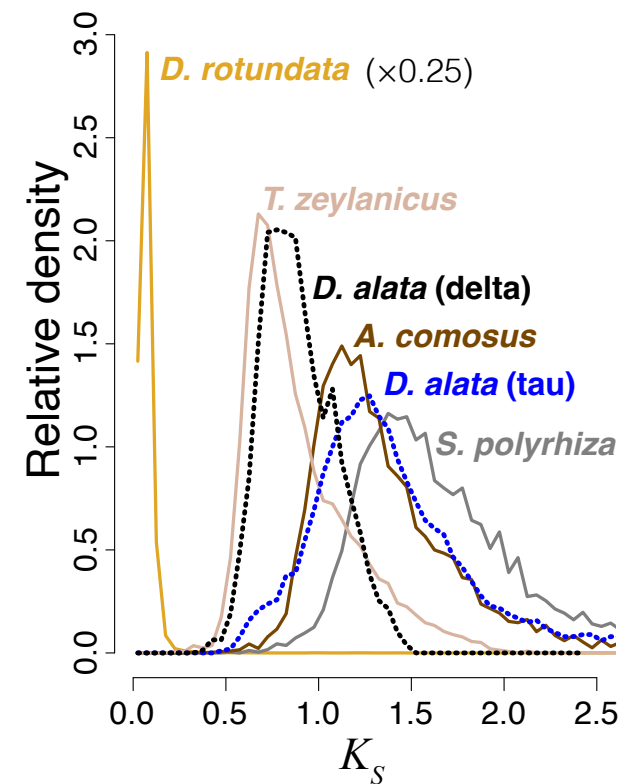
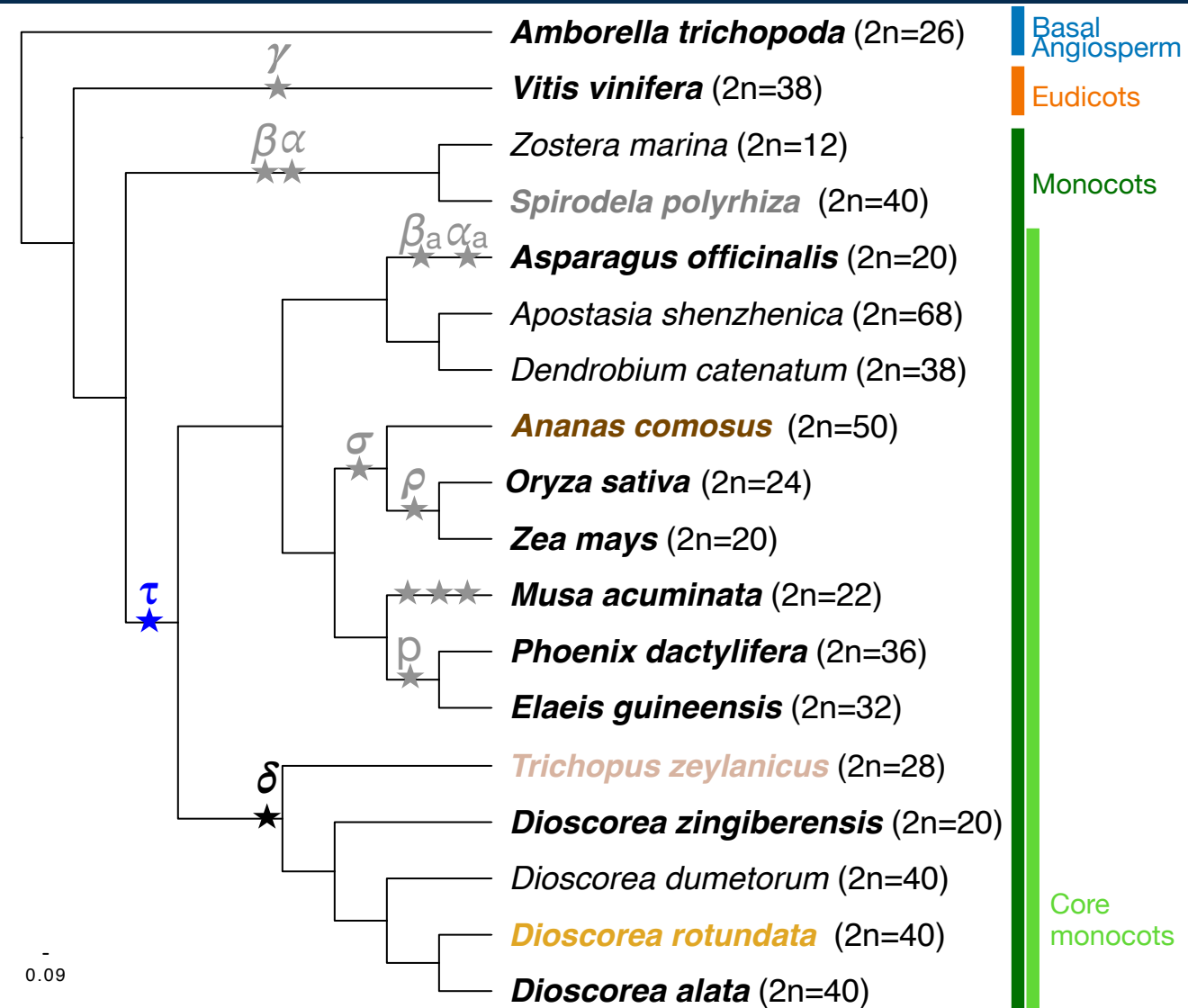
Allopolyploidy



$$\Pr(k; n) = 2^k \cdot n! / (k! \cdot \text{choose}(2n, n))$$

Species	k	n	A	B	Median	p
<i>D. alata</i>	11	11	0.632	0.484	0.552	2.9×10^{-3}
<i>D. rotundata</i>	11	11	0.623	0.493	0.545	2.9×10^{-3}
<i>D. dumetorum</i>	11	11	0.630	0.485	0.557	2.9×10^{-3}
<i>D. zingiberensis</i>	9	11	0.652	0.461	0.553	7.9×10^{-2}

WGDs in phylogenetic context



Species	n	Median K_s
<i>D. rotundata</i>	14,889	0.064
<i>T. zeylanicus</i>	9,013	0.804
<i>D. alata</i> (delta)	1,578	0.869
<i>A. comosus</i>	6,405	1.263
<i>D. alata</i> (tau)	404	1.316
<i>S. polyrhiza</i>	4,973	1.564

Dioscorea–*Trichopus*: ~68 Mya
Dioscorea–*Asparagus*: ~120 Mya
Dioscorea–*Spirodella*: ~128 Mya

The background of the slide features a large, faint watermark of the University of California, Berkeley seal. The seal is circular and contains the text "UNIVERSITY OF CALIFORNIA BERKELEY" around the perimeter. In the center, there is a shield with a book, a star, and the motto "LET THERE BE LIGHT". The year "1868" is also visible at the bottom of the seal.

Identifying agronomically important QTL

Linkage mapping and QTL for crosses segregating for key traits

Populations:

- 10 F₁, 6 outbred parents
- 83–320 progeny / cross

Target traits:

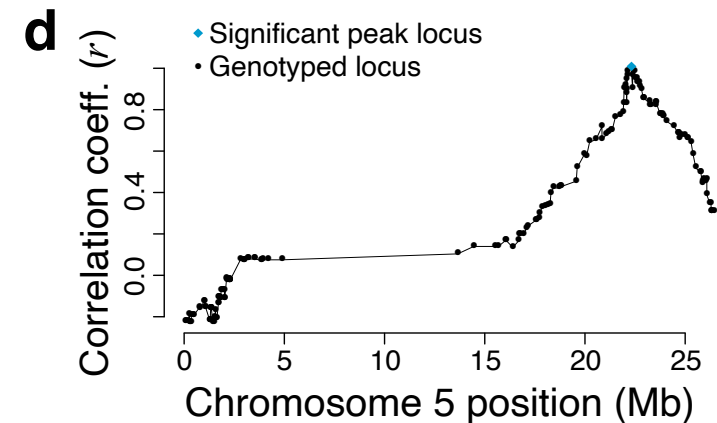
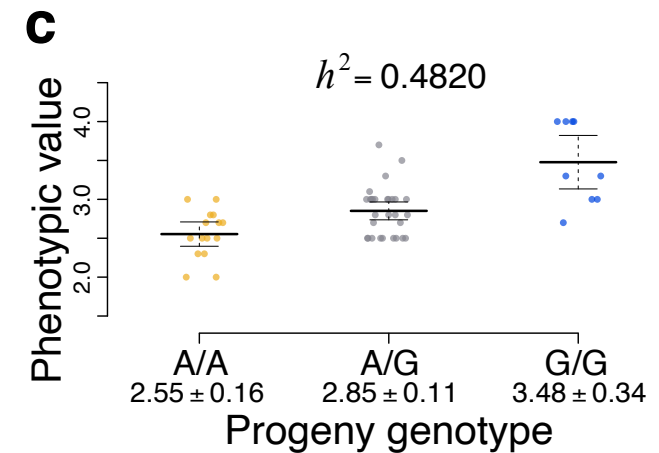
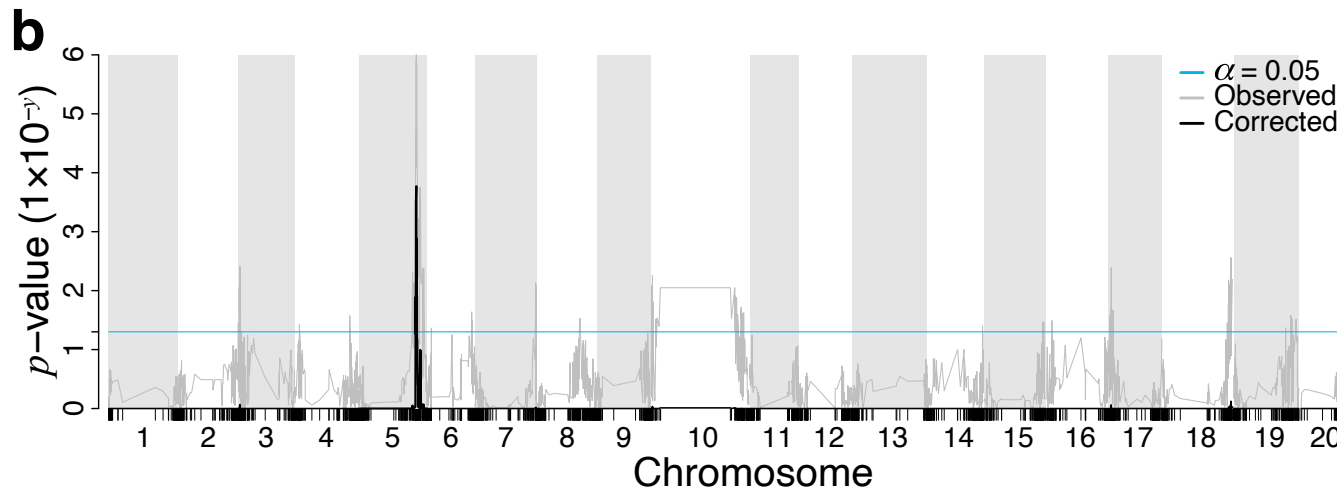
- Anthracnose resistance
- Tuber quality traits

QTL analysis:

- DArT genotyping at IGSS, BecA-ILRI, Kenya
- AlphaFamImpute imputation
- Logistic regression, Wald Test
- $\max(T)$ correction, 1×10^6 iterations



Anthracnose resistance (field trial)

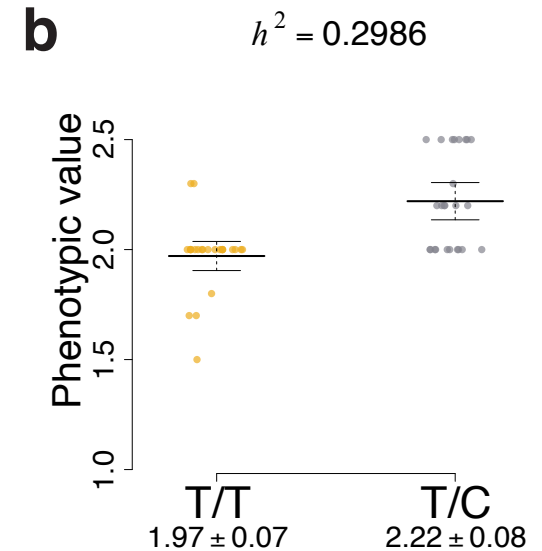
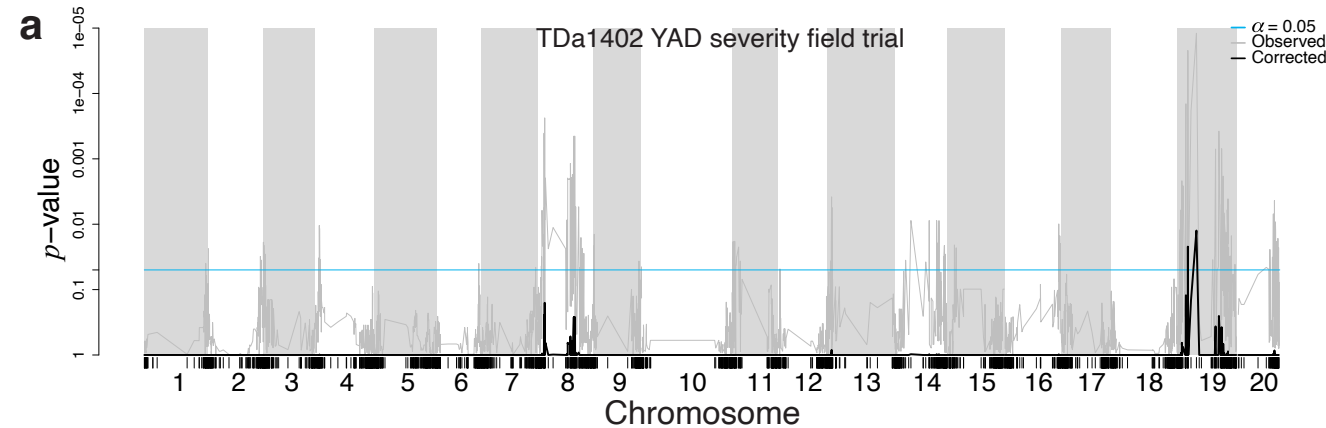


$p = 1.69 \times 10^{-4}$, $n = 53$ (additive effect)
Chr5:22,308,637 (21.93–22.83 Mb)

Dioal.05G18350:

EIX1/2 protein, LRR superfamily

Anthracnose resistance (field trial)



$$p = 1.25 \times 10^{-4}, n = 49$$

Chr19:8,369,514 (3.73–17.57 Mb)

Dioal.19G063700:

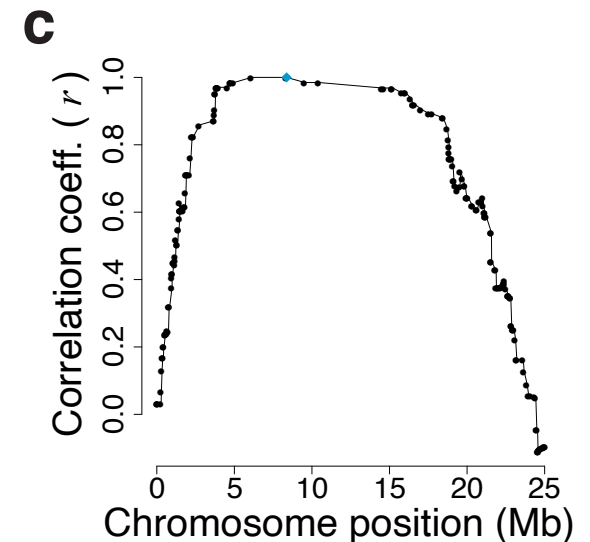
EMSY-LIKE (EML) family of immune regulators of fungal disease resistance

Dioal.19G073100, Dioal.19G074700, Dioal.19G08460:

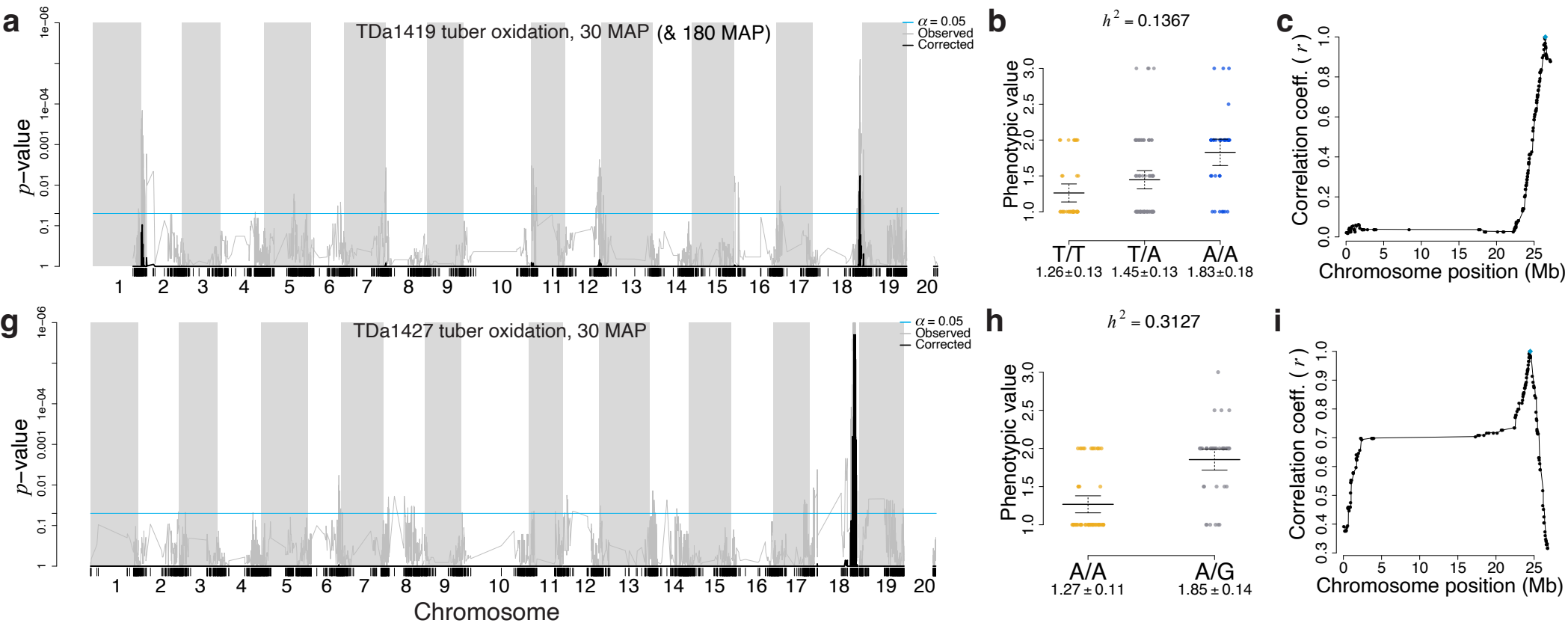
NB-ARC domain-containing R-gene analog (RGA) disease resistance protein-encoding genes

Dioal.19G066100, Dioal.19G066200:

proteins of unknown function that contain C-terminal domains of the ENHANCED DISEASE RESISTANCE 2 (EDR2) family that are negative regulators of plant pathogen response



Tuber oxidation



TDa1419 ($n = 151$):

Chr18: 26,496,992 (26.20–26.75 Mb)

30 mins: $p = 5.86 \times 10^{-3}$

180 mins: $p = 1.38 \times 10^{-2}$ ($h^2 = 0.1188$)

TDa1427 ($n = 97$):

Chr18: 24,495,033 (24.03–24.94 Mb)

30 mins: $p = 4.52 \times 10^{-6}$

Dioal.18G098800, Dioal.18G099400, Dioal.18G100900:
peroxidase-encoding genes

Potential marker-assisted breeding targets

Table 3 Significant QTL identified in this study.

Pop. ID	Trait	QTL peak position	<i>n</i>	<i>p</i> -value	Variant	<i>h</i> ²	Significance Window ^a
TDa1402	Anthracnose susceptibility (Field 2017)	Chr5: 22,308,637	53	1.69×10^{-4}	A/A,A/G,G/G	0.4820	21,931,073 22,825,712
TDa1402	Anthracnose susceptibility (Field 2018)	Chr19: 8,369,514	49	1.25×10^{-2}	T/T,T/C	0.2986	3,732,307 17,565,140
TDa1419	Anthracnose DLA 3-yr mean	Chr6: 61,001	243	1.28×10^{-2}	C/C,C/T	0.0734	38,157 1,418,849
TDa1419	Dry matter	Chr18: 25,069,928	150	2.27×10^{-2}	C/C,C/T	0.1020	24,779,355 25,415,124
TDa1419	Oxidation after 30 min ^b	Chr18: 26,496,992	151	5.86×10^{-3}	T/T,T/A,A/A	0.1367	26,199,630 26,749,589
TDa1419	Oxidation after 180 min ^b	Chr18: 26,496,992	151	1.38×10^{-2}	T/T,T/A,A/A	0.1188	26,199,630 26,749,589
TDa1427	Oxidation after 30 min	Chr18: 24,495,033	97	4.52×10^{-6}	A/A,A/G	0.3127	24,034,264 24,938,398
TDa1401B	Tuber size	Chr12: 310,852	53	4.19×10^{-2}	T/T,T/C,C/C	0.2894	76,400 489,583
TDa1512	Tuber shape	Chr7: 3,115,608	43	3.17×10^{-2}	A/A,A/G	0.3406	1,798,899 5,707,988

Pop. ID mapping population identifier, *n* the number of genotyped and phenotyped progeny used in QTL analysis, *p*-value empirical significance ($\alpha = 0.05$) of the genotype-phenotype association at the peak locus, calculated by Wald statistic-based logistic regression and corrected for family-wise multiple testing by the max(*T*) method, *Variant* alleles segregating at QTL peak position, *h*² narrow-sense heritability.

^aCalculated as haplotypic linkage disequilibrium ≥ 0.9 relative to the peak QTL marker.

^bSame QTL for both oxidation time points in TDa1419.

The background of the slide features a large, faint watermark of the University of California, Berkeley seal. The seal is circular and contains the text "UNIVERSITY OF CALIFORNIA BERKELEY" around the perimeter. In the center, there is a shield with a book, a star, and the motto "LET THERE BE LIGHT". The year "1868" is also visible at the bottom of the seal.

A sneak peek: African yam diversity

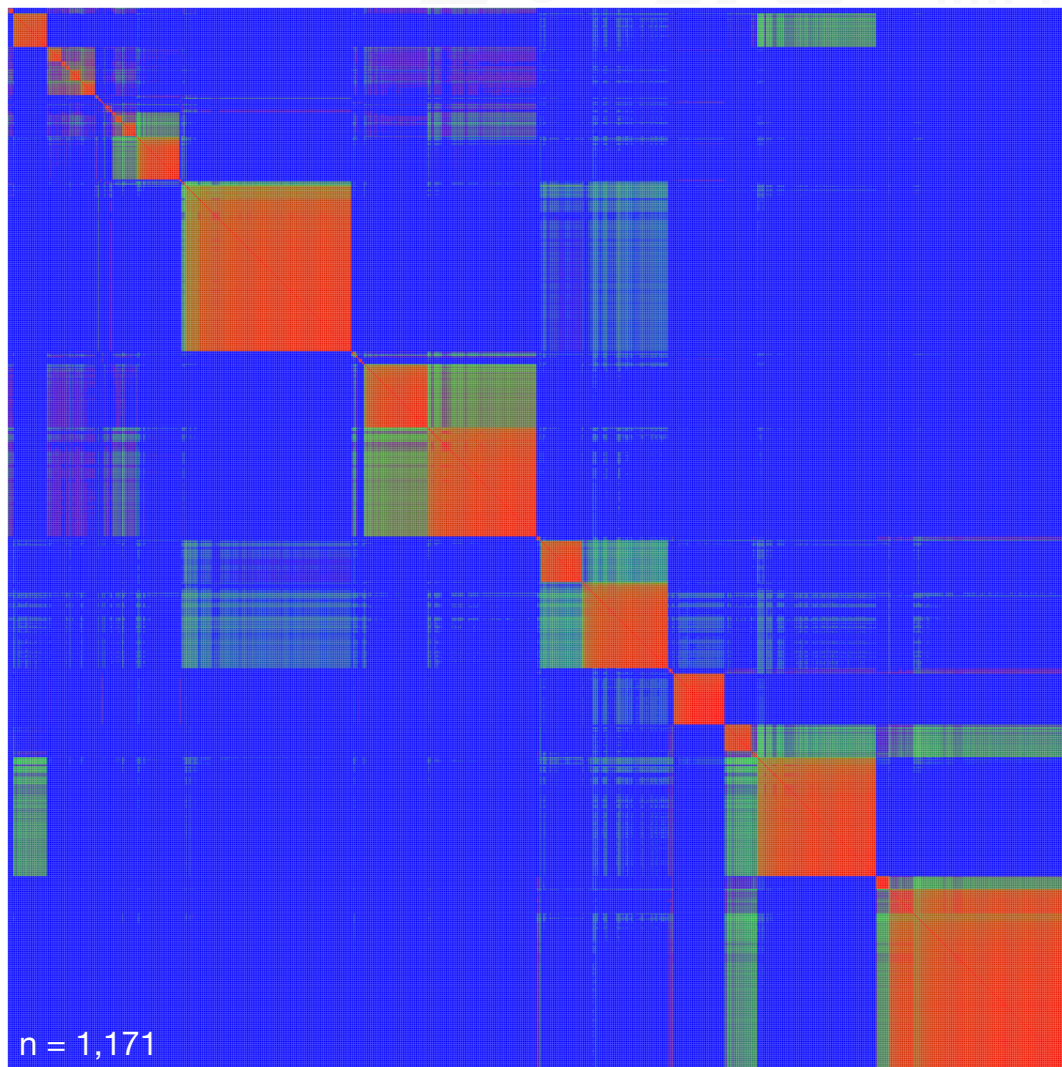
Dense sampling of African Yam Belt

- 1,272 *D. alata* samples
- IGSS, BecA-ILRI, Kenya
- 87 replicated samples (181 total replicates)
- 101 contaminated samples
- 7,441 QC-passing loci @ $MAF \geq 5\%$

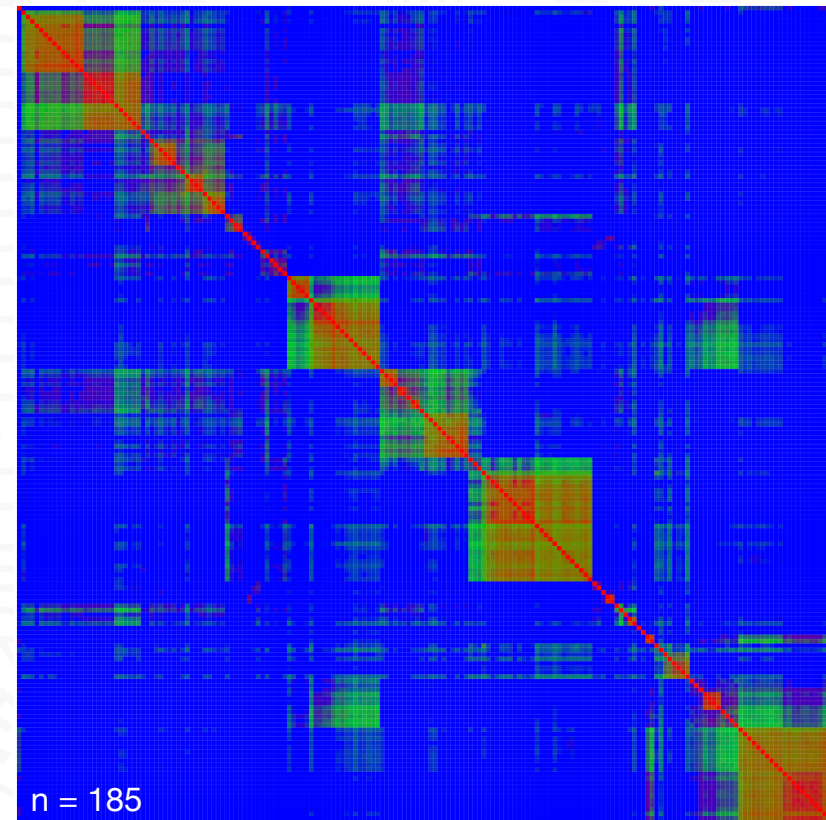
Nominal source	Count
Total	1,272
Nigeria (1,026 IITA; 28 NRCRI)	1,054
Cote d'Ivoire	90
Tanzania	82
Cameroon	18
Japan	16
Vietnam	12

Diversity sample is highly clonal

■ IBS2 ■ IBS1 ■ IBS0



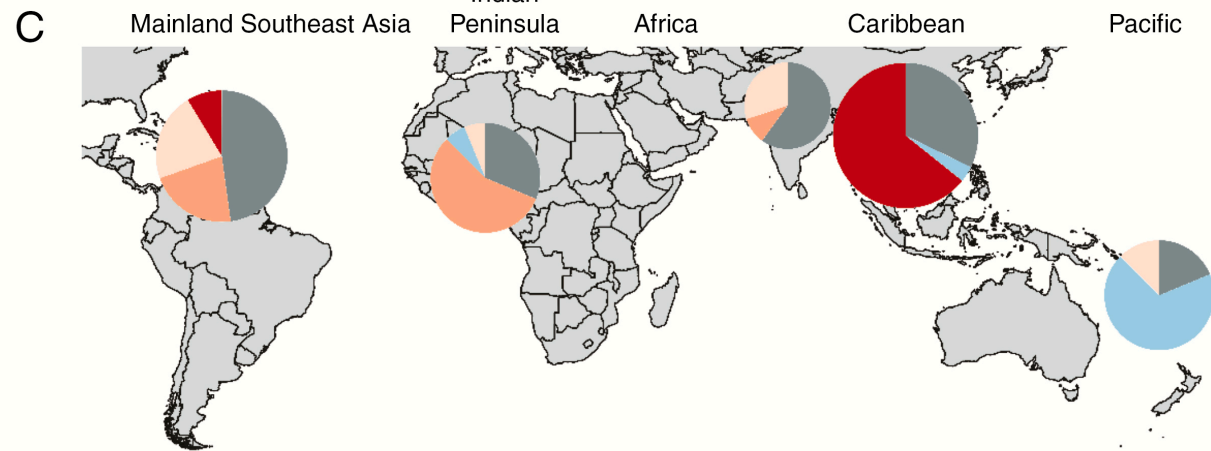
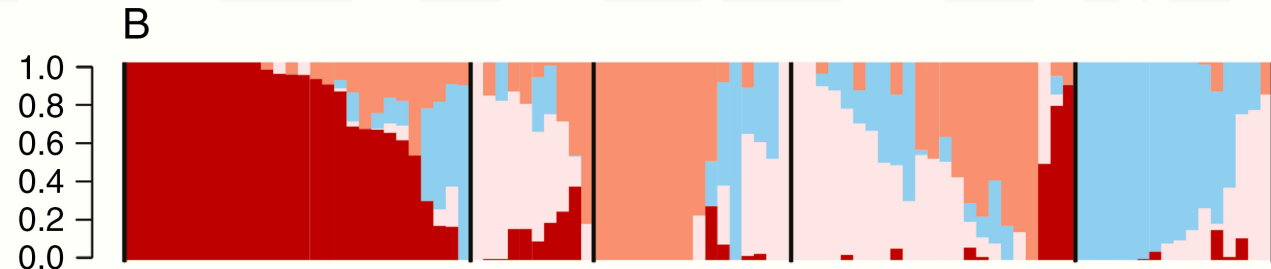
Remove
identicals
(n = 986)
→



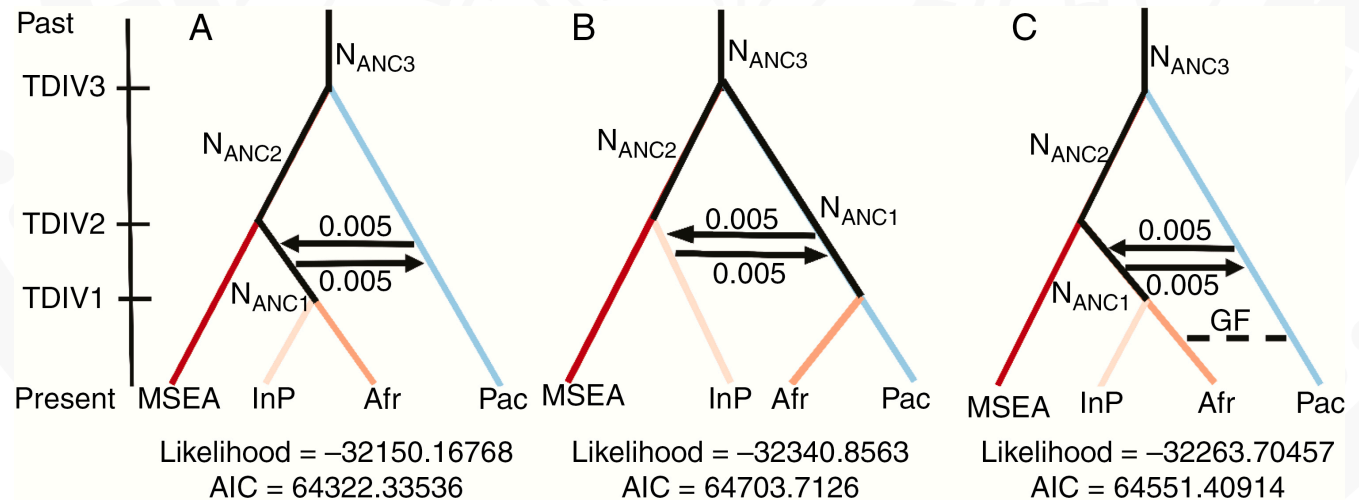
→ Remove
≤1° relatives
(n = 158) → Unrelated
samples
(n = 27)

Can we analyze in a broader global context?

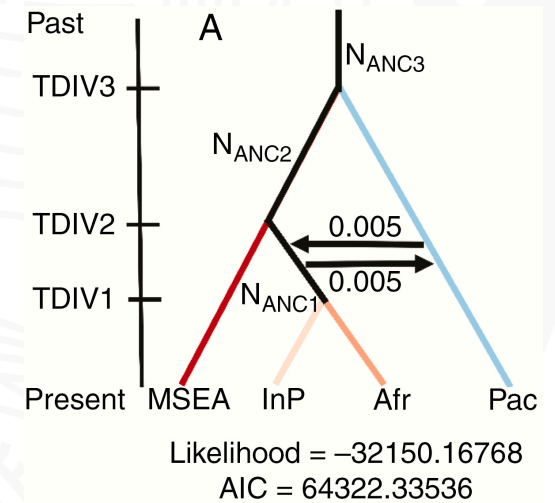
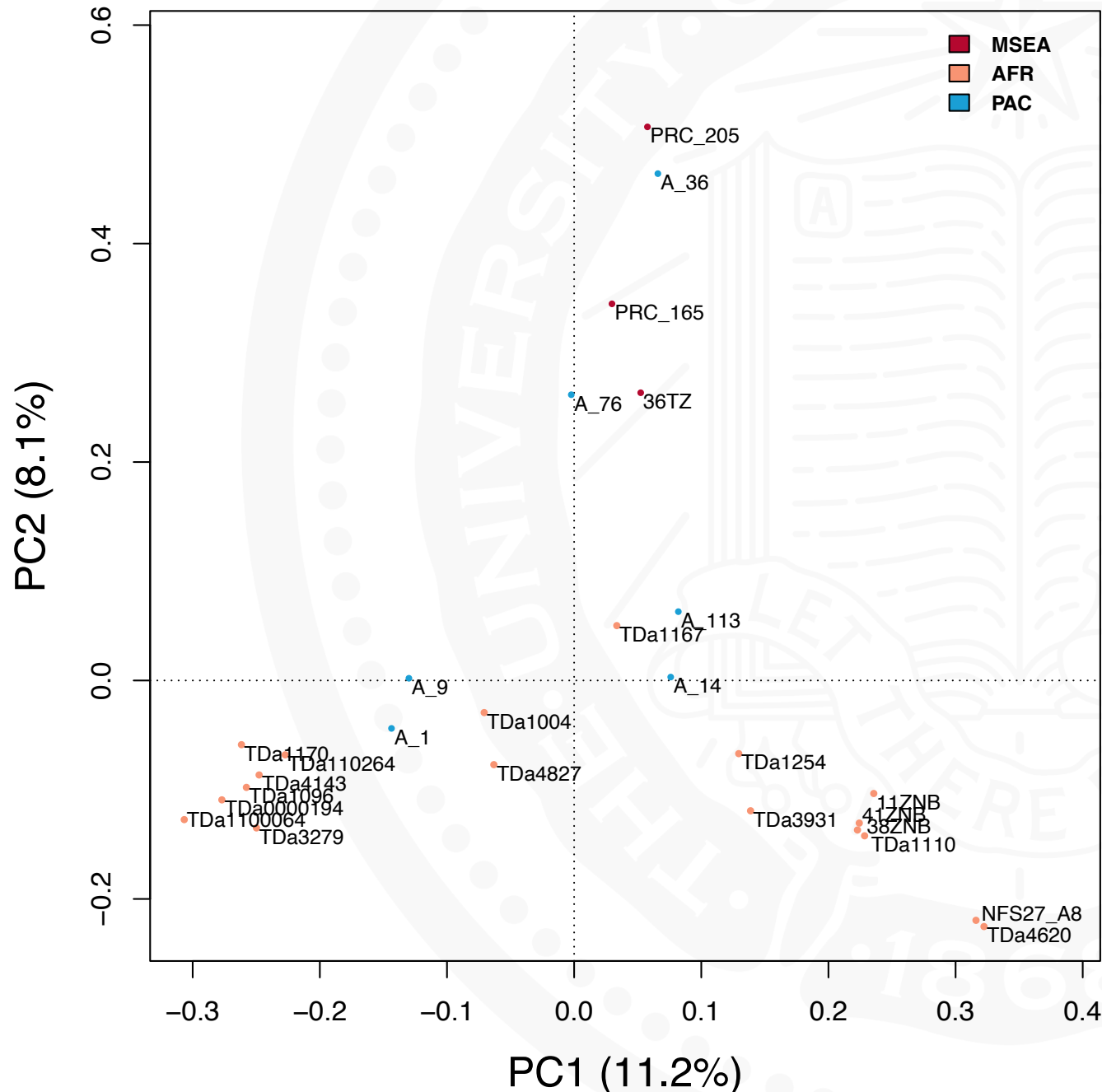
- 93 unrelated samples
- 6k GBS SNPs



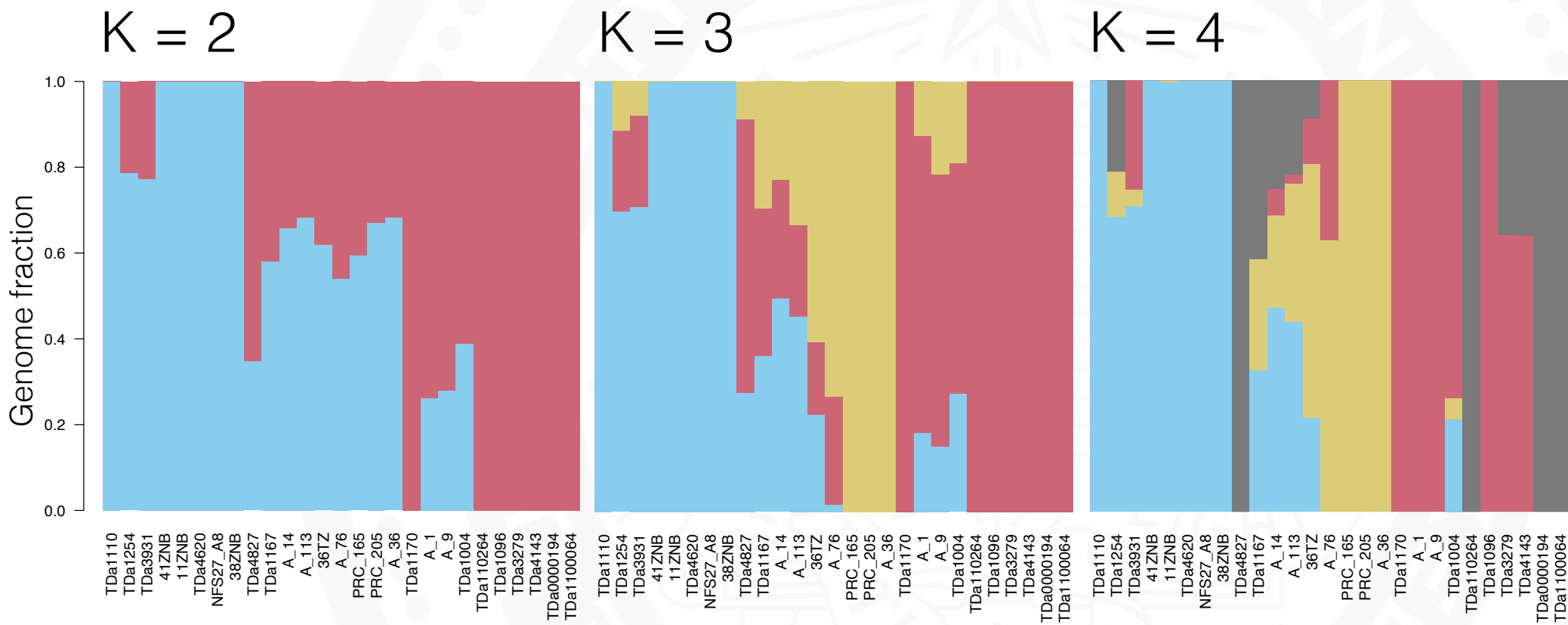
FASTSIMCOAL
continuous time
coalescent simulator



Accessions semi-differentiated by continent



To do: segmental population assignment



D. alata genome project



Dan Rokhsar (UC Berkeley, JGI)

Jess Lyons (fmr. UC Berkeley)

Ranjana Bhattacharjee (IITA Ibadan)

Ibukun Ogunleye (IITA Ibadan)

Jude Obidiegwu (NRCRI)

Chiedozie Egesi (NRCRI, IITA, Cornell)

Shengqiang Shu (JGI)

Joe Carlson (JGI)

David Goodstein (JGI)

Therese Mitros (UC Berkeley)

Allen Van Deynze (UC Davis, AOCC)

Lukas Mueller and colleagues (BTI)

Prasad Hendre (ICRAF, AOCC)

Ramni Jamnadass (ICRAF, AOCC)

Alice Muchugi (ICRAF, AOCC)

Jaroslav Doležel (IEB Czech Rep.)

Eva Hřibová (IEB Czech Rep.)

Gordon Simpson (JHI, Dundee Univ.)

Mathew Parker (Dundee Univ.)

Kasia Knop (Dundee Univ.)

Nick Schurch (Dundee Univ.)

Geoff Barton (Dundee Univ.)

Jonathan Featherston (ARC S. Africa)



Acknowledgements

Lava Kumar (IITA)
Kolade Olufisayo (IITA)
Asrat Amele (IITA)
Agre Paterne (IITA)
David De Koeyer (IITA)
Robert Asiedu (IITA)
Antonio Lopez-Montes (former IITA)
Michael Abberton (IITA)

Clay Sneller (IGSS)
Jackline Chepkoech (IGSS)
Mercy Elohor (IGSS)
Andrzej Kilian (DArT)
Lutz Froenike (UC Davis)
Oanh Nguyen (UC Davis)
Dovetail Genomics LLC
Shana McDevitt (UC Berkeley VCGSL)



**World
Agroforestry**



YamBase



IGSS Africa

UC DAVIS
UNIVERSITY OF CALIFORNIA



BILL & MELINDA
GATES *foundation*

Univ. Dundee GCRF Challenge Fund
Illumina Greater Good Initiative

Thank you for listening!



jbredeson@lbi.gov