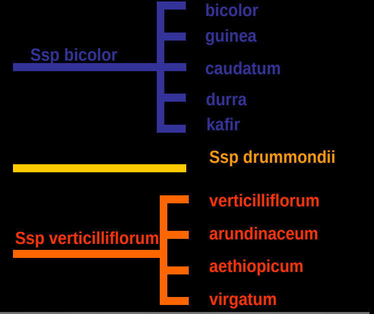
2012 International Sorghum Genomics Workshop
14-16th November


Ssp bicolor
- bicolor
- guinea
- caudatum
- durra
- kafir

Ssp drummondii

Ssp verticilliflorum
- verticilliflorum
- arundinaceum
- aethiopicum
- virgatum

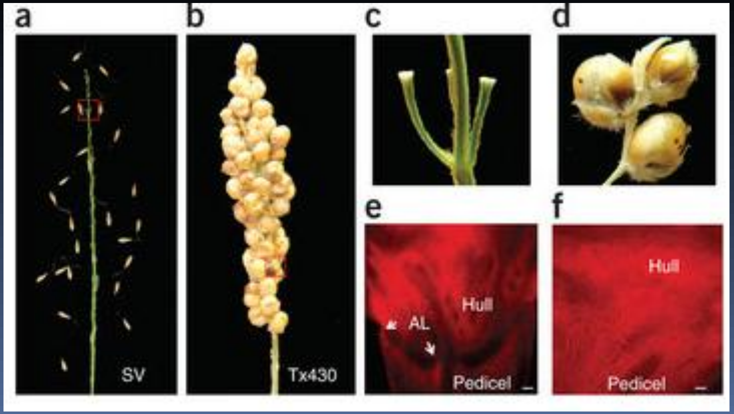
Sorghum bicolor, Comparison of wild and cultivated pools

David Pot
UMR AGAP
Genetic Resources team
david.pot@cirad.fr

Defining a neutral scenario of evolution for sorghum

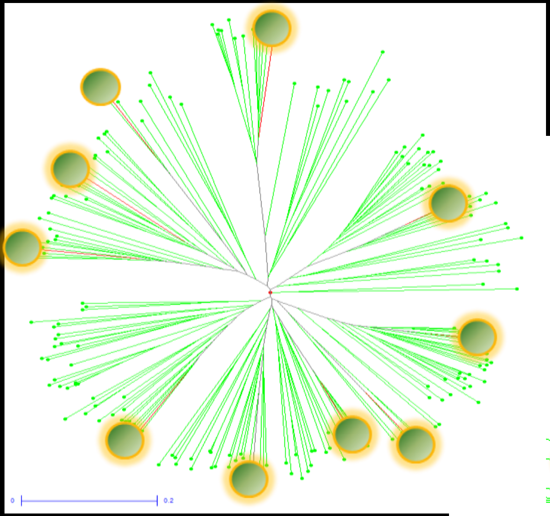
- Bottleneck intensity between wild and cultivated pools
- Cultivated pool expansion rate
- Gene flow between cultivated and wild pools

=> Provide a neutral hypothesis to identify Domestication / Adaptation genes



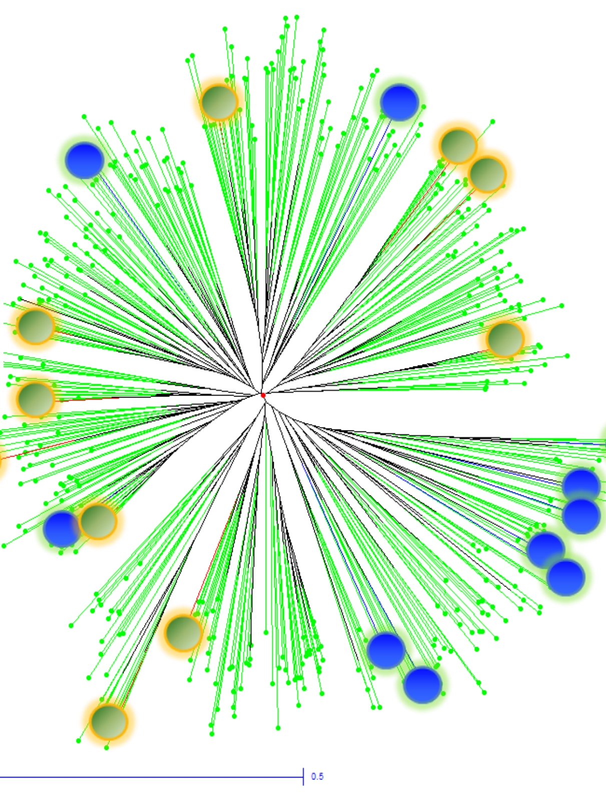
Lin et al, 2012

Selection of 10 cultivated and 10 Wild accessions

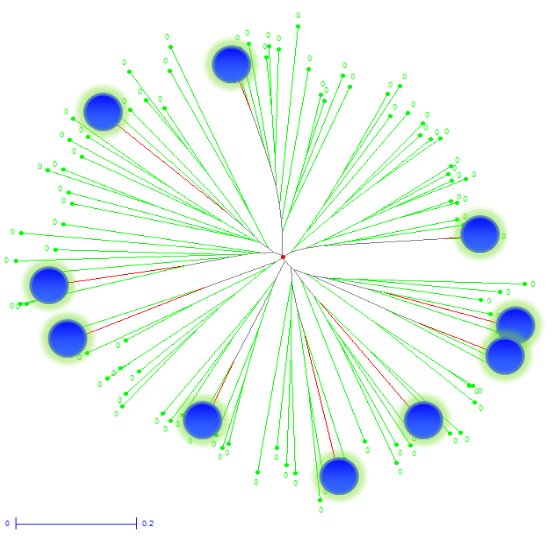


Cirad Cultivated Core collection (210 accessions, Deu et al 2006)

SSM1049	B	SEN
IS29876	C	ZWE
IS30436	C	CHN
SSM1123	C	NER
IS6193	D	IND
SSM973	D	SEN
IS14317	Gro	SWZ
IS29407	KC	LSO
SSM1057	Gma	SEN
IS26554	G	BEN



GCP reference Set + 100 wild accessions



100 wild accessions

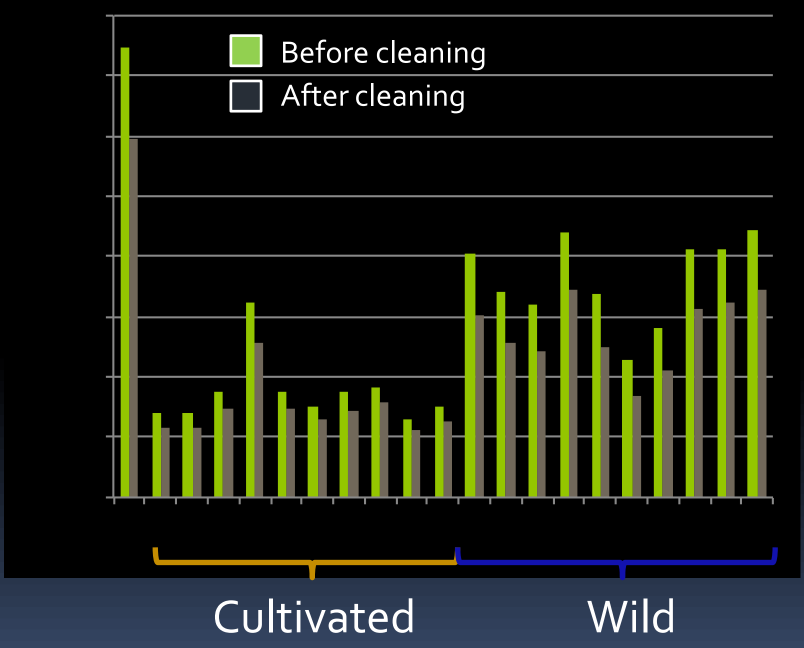
IS14564	aethiopicuum	SDN
IS18821	aethiopicuum	EGY
IS18909	aethiopicuum	TCD
IS18824	arundinaceum	CIV
IS18833	arundinaceum	MWI
IS14312	verticilliflorum	ZAF
IS14357	verticilliflorum	MWI
IS14719	verticilliflorum	ETH
IS18804	virgatum	USA
IS18812	virgatum	EGY

- + *Sorghum brachypodum* (2n = 10, annual, from Australian Tropical Crops and Forages Collection

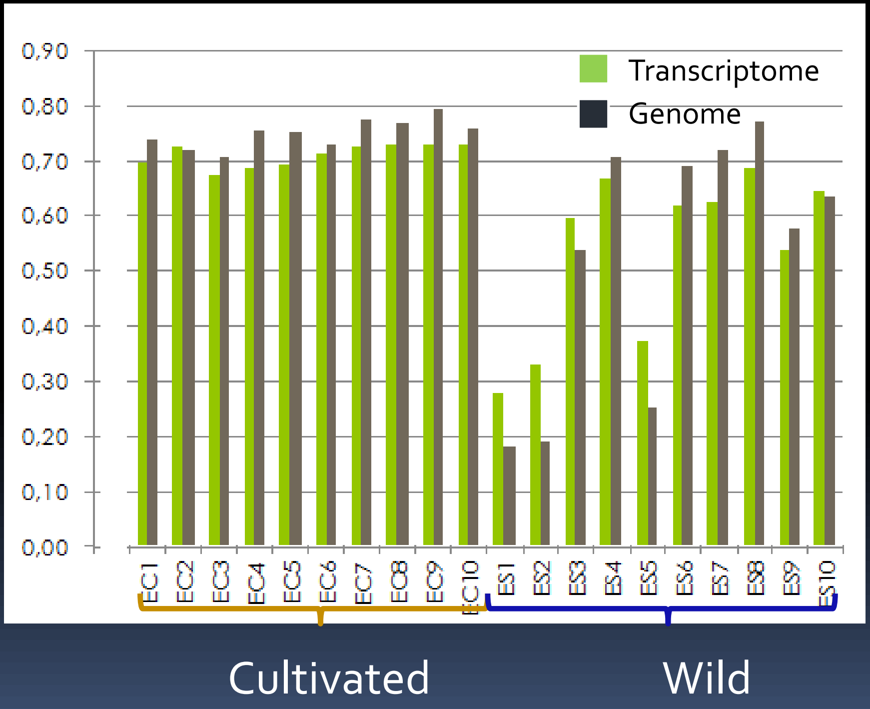
Sequencing and mapping

- RNA extraction : leaves, flower (anthesis), maturating seeds
- Hi Seq 2000 sequencing : one lane for cultivated, one lane for wild and ½ lane for *S. brachypodum*

Sequencing



Mapping on Btx623 (BWA, realigner GATK)



NB : The Low mapping proportion of wild accessions remains to be solved…

Polymorphism identification and filters

- GATK polymorphism detection (Only PASS considered)

- 8X coverage in at least 8 accessions

- Polymorphism with too high heterozygosities were discarded (>50 % in cultivated, >70 % in wild)

- Polymorphisms detected (20 Ksnp common with U(Zheng et al 2011 and Neslon et al 2011)=82 193 snp)

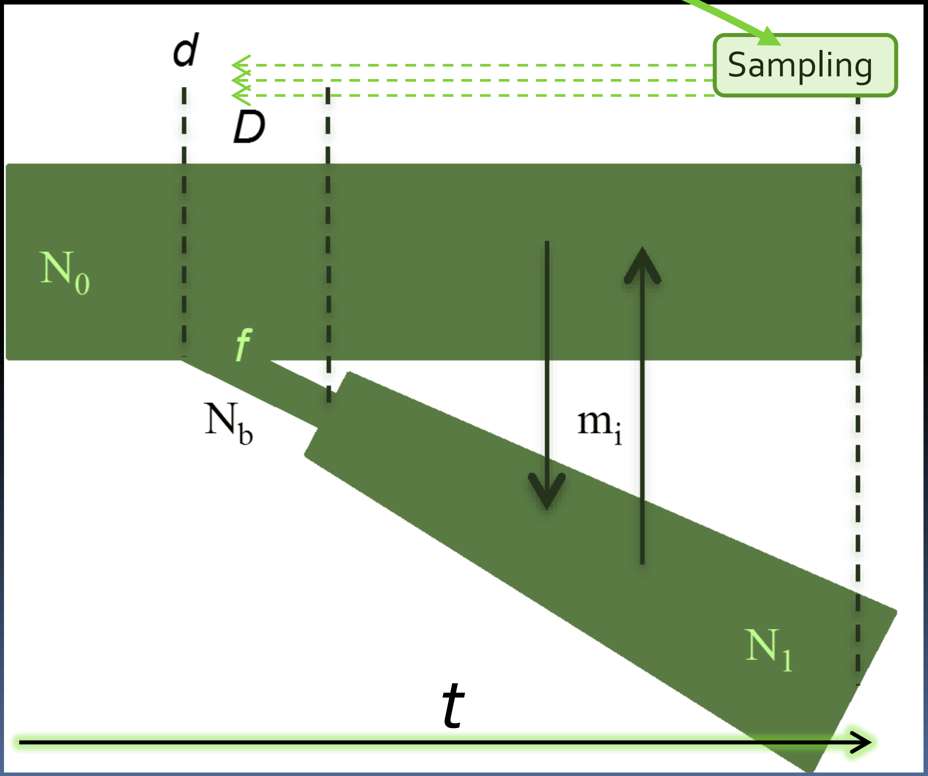
Type	Number
Wild	173 042
Specific to Wild	92 304
Cultivated	105 707
Specific to cultivated	25 919
In Wild and Cultivated	79 360

=> 171 450 polymorphisms from 14 150 genes selected for model analysis (information for at least 6 cultivated and 6 wild accessions). A total of 14 784 genes were considered

Nucleotide diversity in wild and cultivated pools

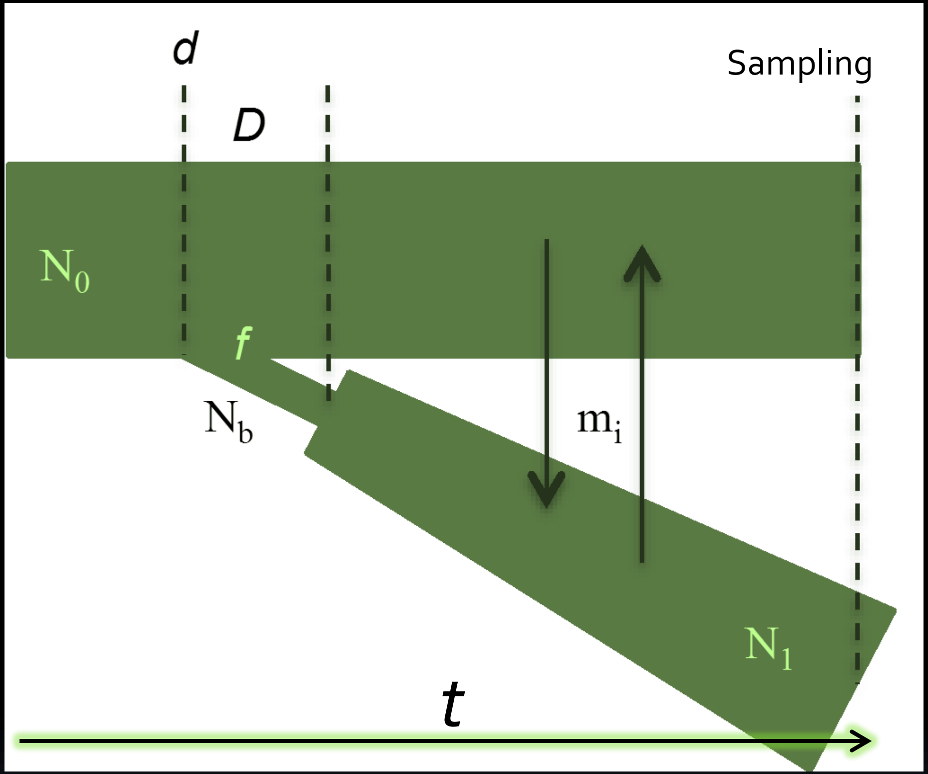
	Wild	Cultivated
θ	0.0032	0.0021
π	0.0032	0.0023
Tajima's D	-0.17	0.14
Kst	0.04	

- Reduction of diversity in cultivated : - 40 %
- Weak differentiation (low number of accessions)

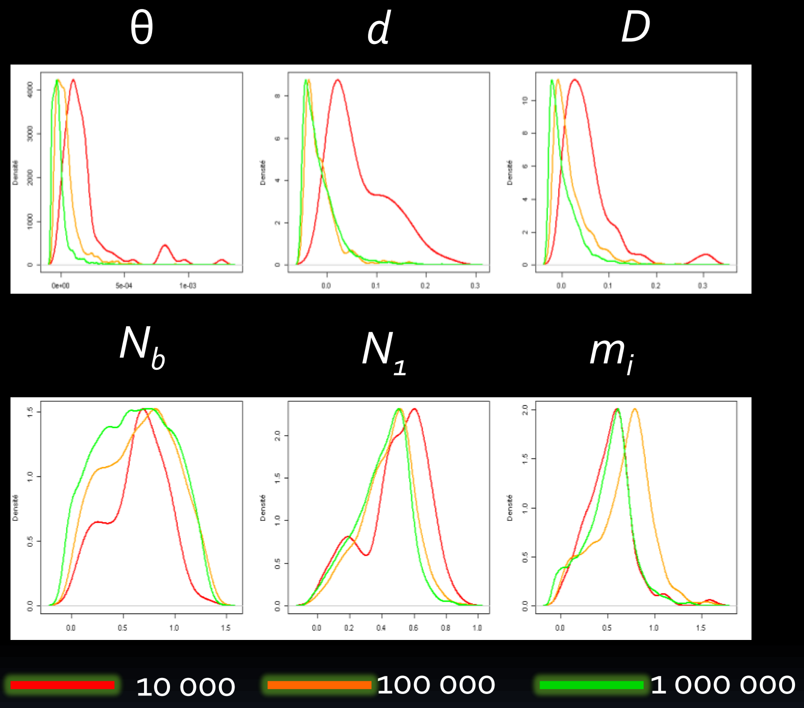


- Towards a neutral model of evolution
  - Approximate Bayesian Computation
  - Egglib (De Mita and Siol 2012)
  - Bottleneck intensity
  - Cultivatred pool expansion
  - migration

Neutral model of evolution (Simple !!!)



Optimal number of simulations

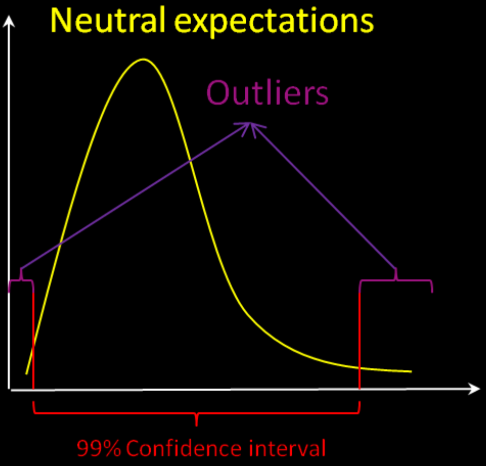


- Bottleneck intensity : Nb = 0.65 * No
- Expansion of cultivated pool : N1 = 1.001 * Nb
- Effective number of migrants per generation: 0.53

•With μ= 1.0e-08 , domestication time ≈ 6000 years (in accordance with archeological reports)

Domestication / Adaptation genes

- **Outlier identification strategy**
  - Neutral model
  - Distribution of the parameters under neutrality
  - Theshold definitions
  - for Kst, 22 (0.15 %) were higher than the 95% CI and 1276 (9%) were lower than the 95% CI



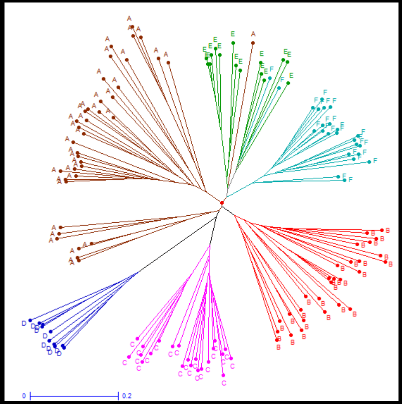
- **Some flowering genes exhibit a lack of diversity in the cultivated pool**
  - Auxin response factor 3 (Sb03g034850)
    - θ cult = 0 vs θ wild = 0.01436 (12 SNP)
    - affects developmental timing and patterning

  - Auxin response factor 4 (Sb06g011767)
    - θ cult = 0 vs θ wild = 0.0149 (7 SNP)
    - affects the development of floral organs and leaves

Domestication / Adaptation genes

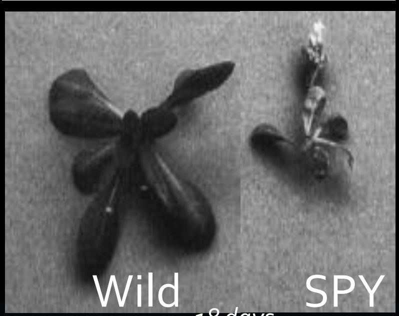
Convergence between evolutionnary and quantitative analyses

- Tetratricopeptide repeat domain (Sb01g019560): High differentiation between the 2 pools
  - $\theta_{Cult} = \theta_{wild} = 1.399x10^{-2}$
  - Kst =0.431



- Identified as an outlier in cultivated analysis (Bouchet et al 2012)
  - Fst = 0.55 (DArT marker at 7.7 kb from this gene)

- Homolog of SPINDLY (SPY) which is a negative regulator of the response to giberelins (Jacobsen et al 1996, Tseng et al 2001)
  - SPY mutant exhibits early flowering, pale green leaves, partial male sterility and pathenocarpic fruit development



18 days

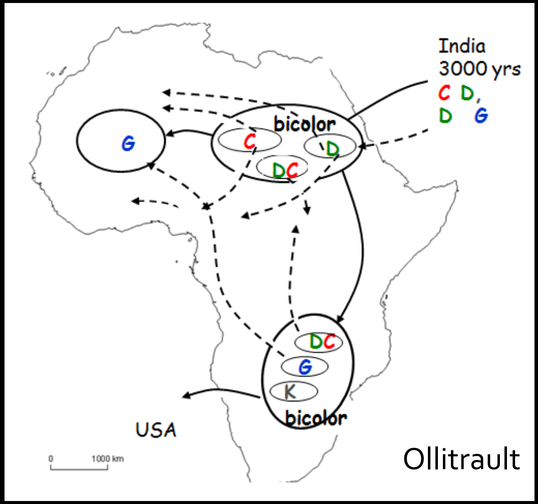
- Sb01g019560 is localized on Chr 1, 43 cM (21267707 – 21279160 bp)
  - A maturation metaQTL between 42.5-44 cM (Mace and Jordan, 2011)
  - « Significant » association with flowering time in two association studies trials (pvalue = 2.1 10^{E}-3 and 3.7 10^{E}-3.)

- These outliers need to be validated through larger diversity analysis, specific population development…

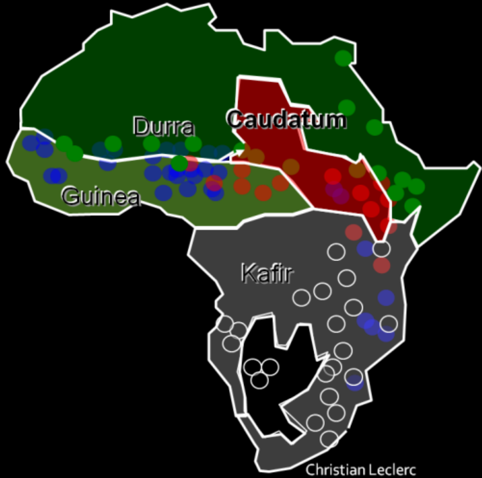
A first step, but several limitations

- An extremely simple and general neutral scenario that needs to be refined…

- Impossibility to test complex domestication scenarios due to the lack of diversity coverage
  - origin of Asian accessions…
  - anthropic impacts



Ollitrault



Christian Leclerc
From Harlan et Stemler

- Only 14 894 genes analyzed (not full lengths), requirement for « real » whole genome information **if outlier scan is a target**

- Would be interesting to sample elite commercialized accessions (grains, forage, biomass) to check what have been selected and what is missing !!!

- Availability of the polymorphism data (mid-2013)

Thanks for you attention !