Genomics assisted breeding in alfalfa

Bernadette Julier, Marie Pégard, Julien Leuenberger, Philippe Barre

▶ To cite this version:

Bernadette Julier, Marie Pégard, Julien Leuenberger, Philippe Barre. Genomics assisted breeding in alfalfa. EUCLEG Workshop on Genomics assisted breeding in forage and grain legumes, Sep 2021, online, United Kingdom. hal-03365357

HAL Id: hal-03365357

https://hal.inrae.fr/hal-03365357

Submitted on 5 Oct 2021

Genomics assisted breeding in alfalfa

Bernadette Julier

Marie Pégard, Julien Leuenberger, Philippe Barre

www.eucleg.eu

Alfalfa - Lucerne

A major legume species
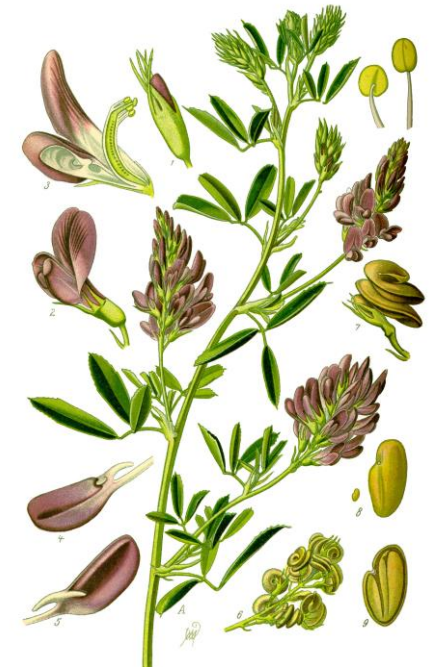
- Highest protein production/ha in temperature climates

- Drought tolerant

- Protein/energy

- Ruminant health

- Positive effects in the rotation

Allogamous reproduction, synthetic varieties

$2n = 4x = 32$

Julier et al. 2017, CABI Publishing
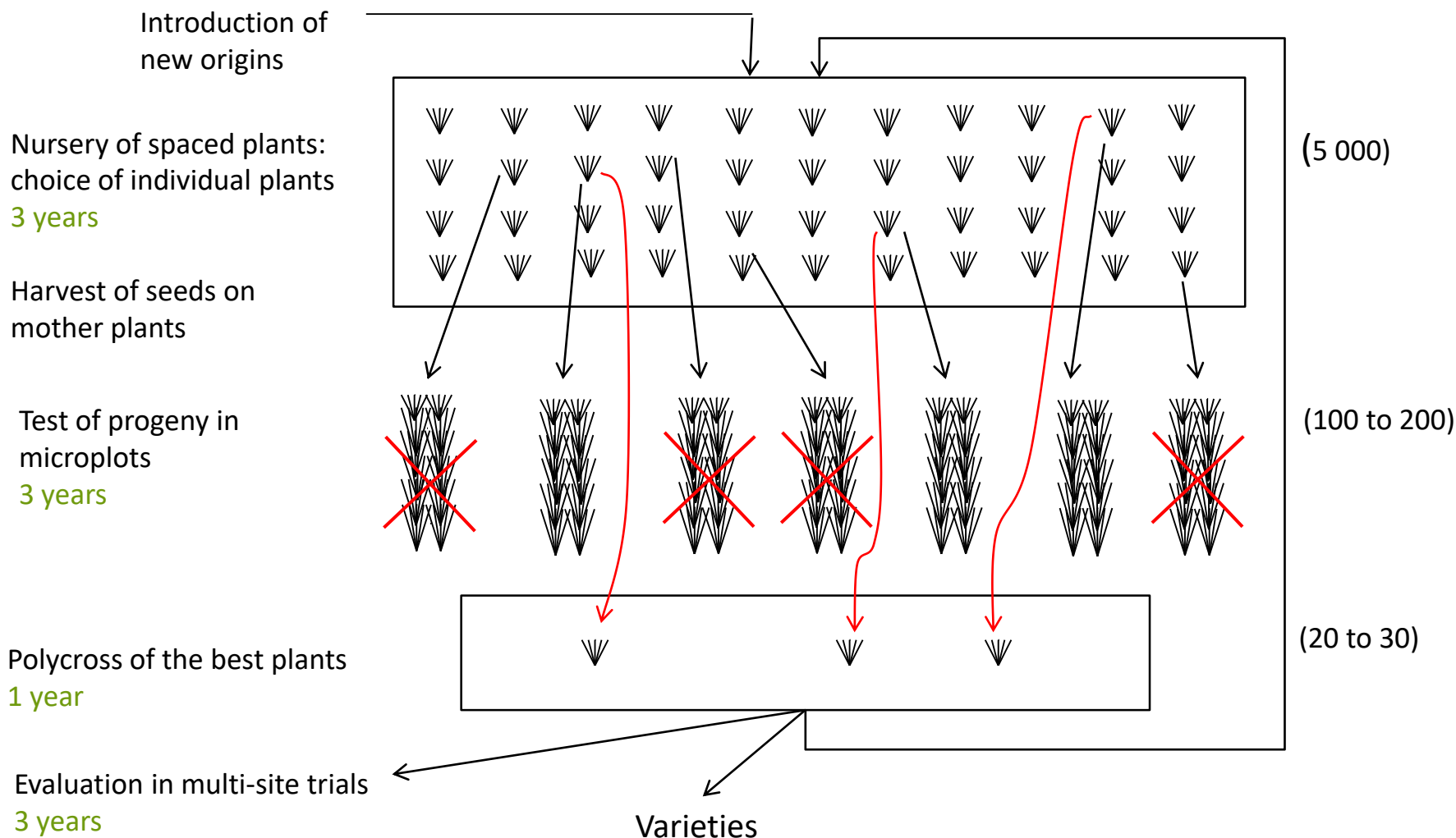
Traditional breeding methodology

- Based on phenotypic traits

- First step in nursery of spaced plants

- Second step in progeny testing

Traditional breeding methodology

Introduction of new origins

Nursery of spaced plants: choice of individual plants
3 years

Harvest of seeds on mother plants

Test of progeny in microplots
3 years

Polycross of the best plants
1 year

Evaluation in multi-site trials
3 years

Varieties

(5 000)

(100 to 200)

(20 to 30)

THE FRAMEWORK PROGRAMME FOR RESEARCH AND INNOVATION

HORIZON 2020

4

Traditional breeding methodology

Introduction of new origins

Nursery of spaced plants: choice of individual plants
3 years

(5 000)

Harvest of seeds on mother plants

Test of progeny in microplots
3 years

(100 to 200)

Polycross of the best plants
1 year

(20 to 30)

Evaluation in multi-site trials
3 years

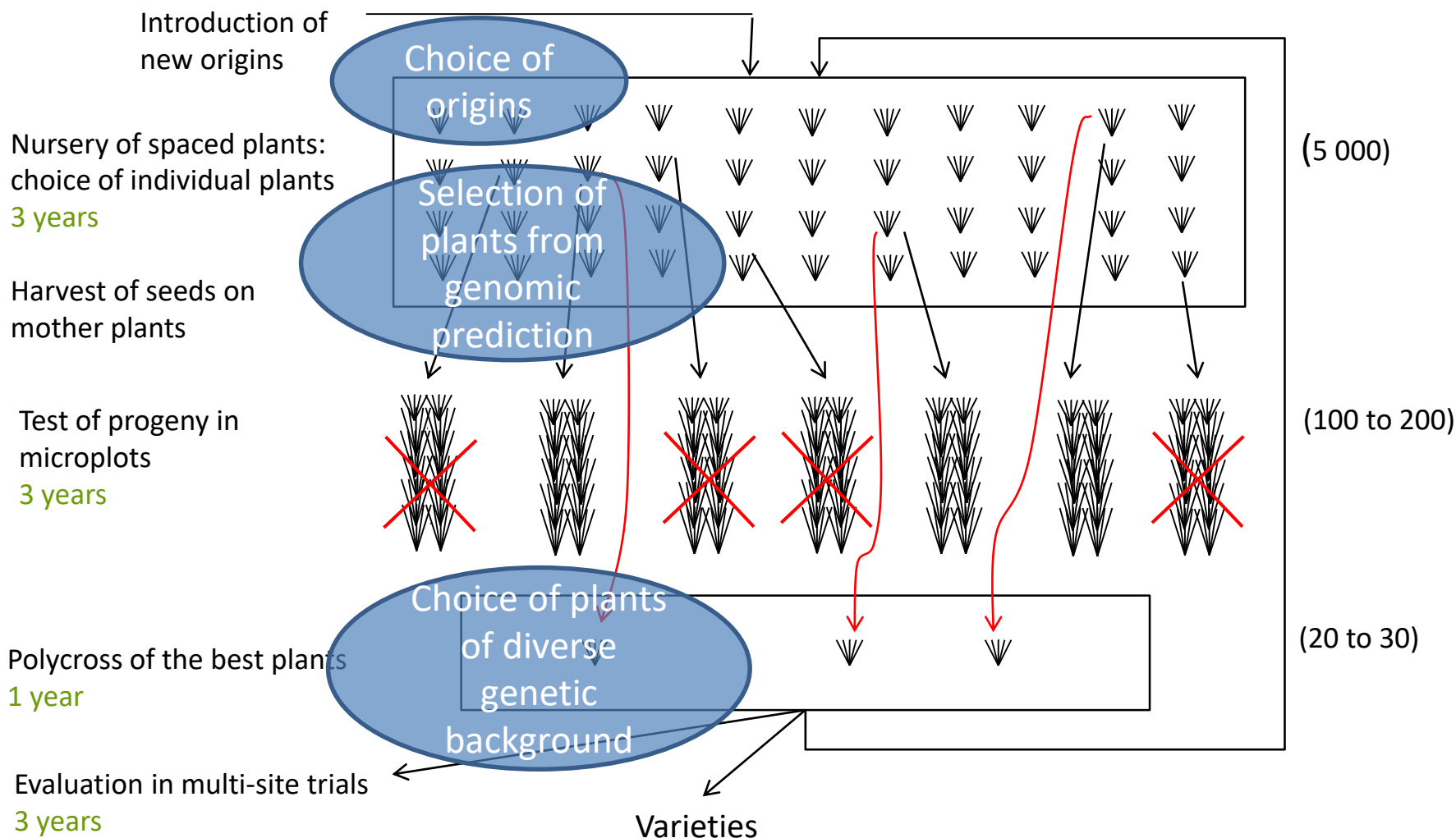Varieties

Traditional breeding methodology

Strength

- Scoring of many traits
- Early selection for heritable traits
- Skilled staffs

Weakness

- Some traits are scored in case of stress occurrence only
- Number of years
- Cost
- Fixation of positive alleles is slow

Genomics assisted breeding

Introduction of new origins

Choice of origins

Nursery of spaced plants: choice of individual plants
3 years

Selection of plants from genomic prediction

Harvest of seeds on mother plants

(5 000)

Test of progeny in microplots
3 years

(100 to 200)

Choice of plants of diverse genetic background

Polycross of the best plants
1 year

(20 to 30)

Evaluation in multi-site trials
3 years

Varieties

Genomics assisted breeding

- Marker development

- Management of genetic diversity

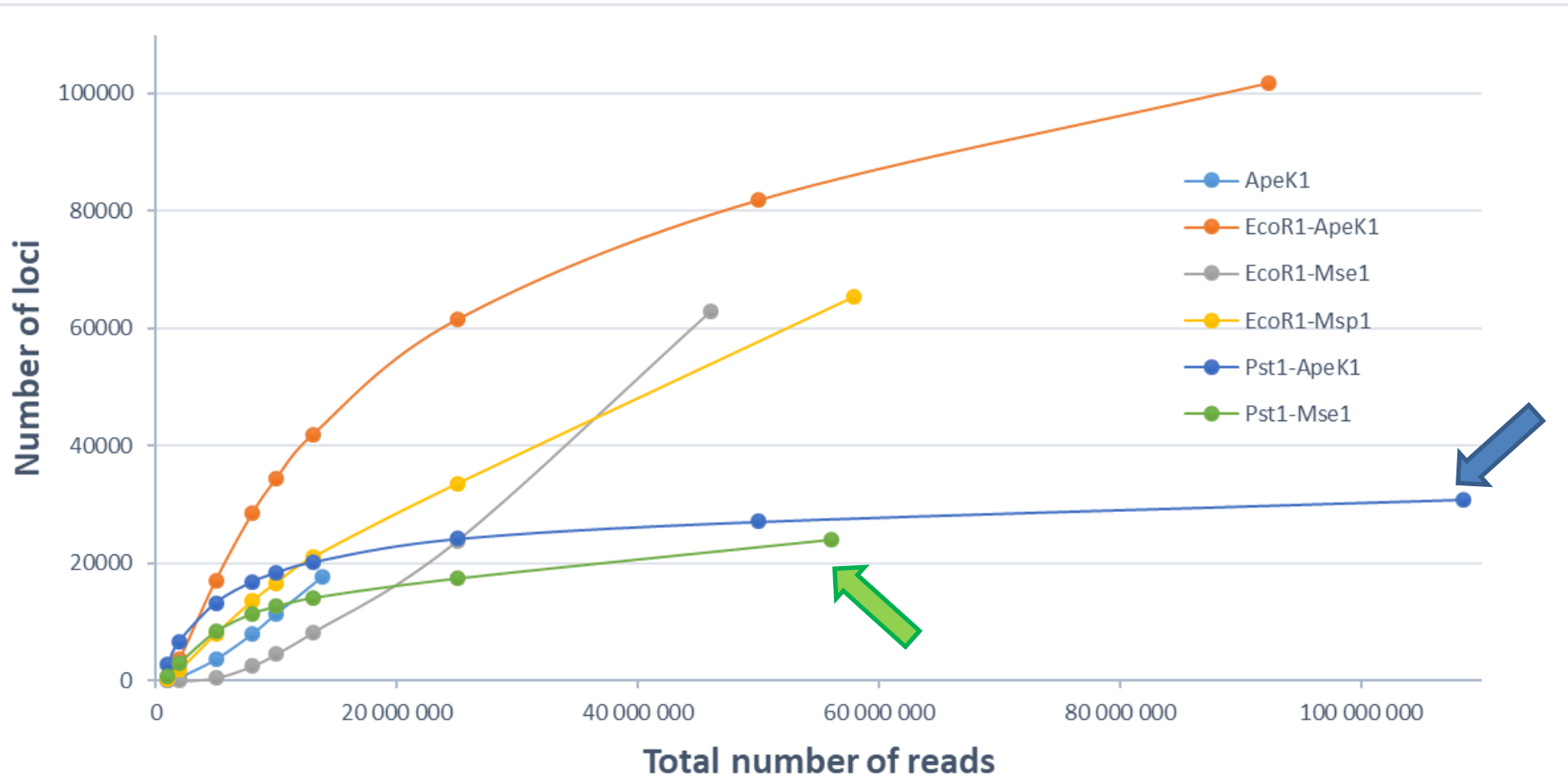- Genome wide association study

- Genomic selection

Marker development

Before EUCLEG:

- Low throughput markers: SSR, AFLP…

- 10k SNP array: too expensive

- GBS: < 40K markers, risk of missing data

Marker development

EUCLEG: an improved GBS protocol

- Choice of restriction enzymes to reduce missing data
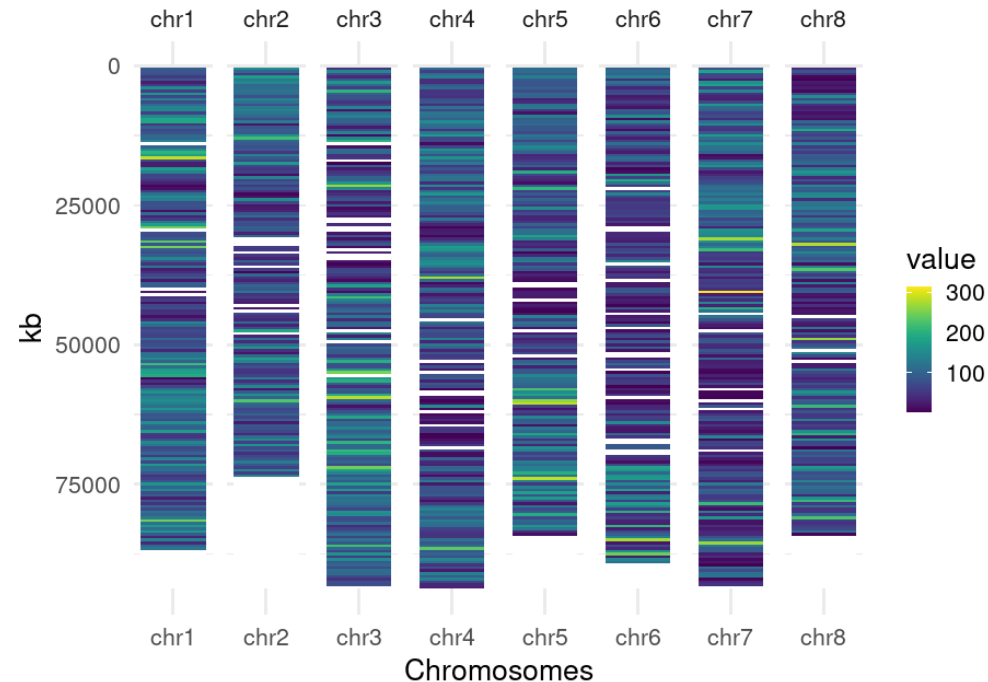
Marker development

Marker development

EUCLEG: an improved GBS protocol

- Choice of restriction enzymes to reduce missing data

- Use of a reference genome sequence: Chen et al. 2020

- Allele frequency of each accession
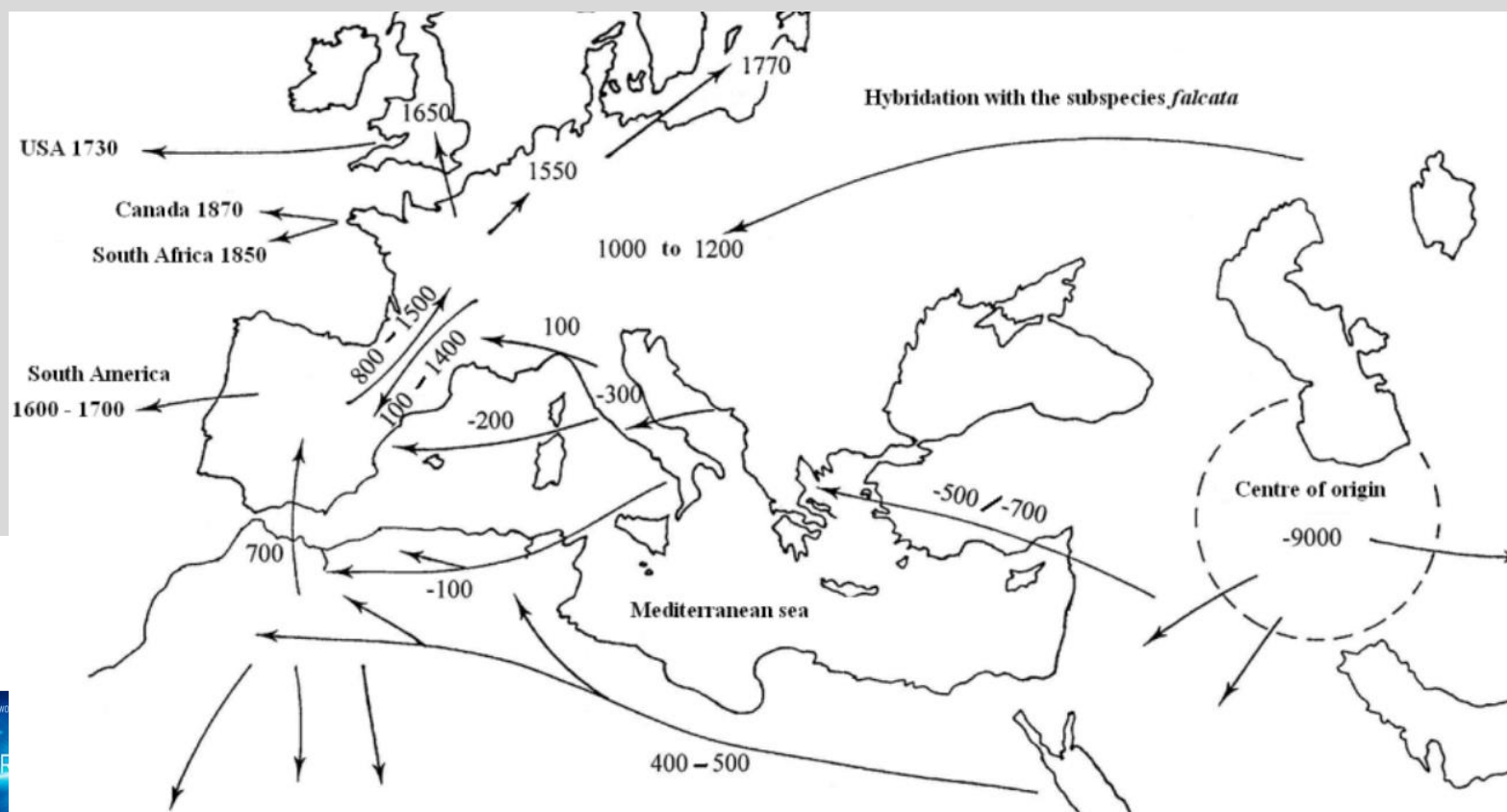
Marker development

On 1 061 accessions:

- 31 743 loci

- 228 568 SNP with less than 5% missing data per SNP

- 118 421 SNP without missing data

Management of genetic diversity

Before EUCLEG:

- Overview of world diversity

Management of genetic diversity

Before EUCLEG:

- Overview of world diversity

- Large among-accession diversity

- Huge within-accession diversity

	10 populations, 40 indiv/pop		11 populations, 7-20 indiv/pop
	5 SSR	Plant height	Yield
Variance among-varieties	0.02	0.10	1.7
Variance within-varieties	7.56	0.30	27.7
		No structure	

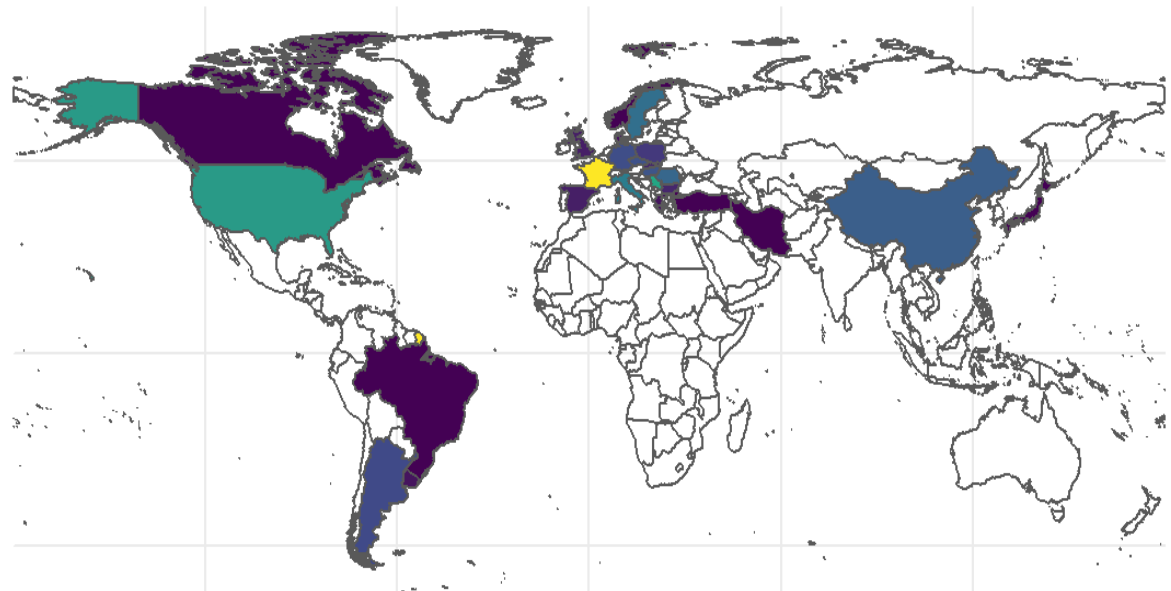Herrmann et al., 2010   Julier et al. 2000

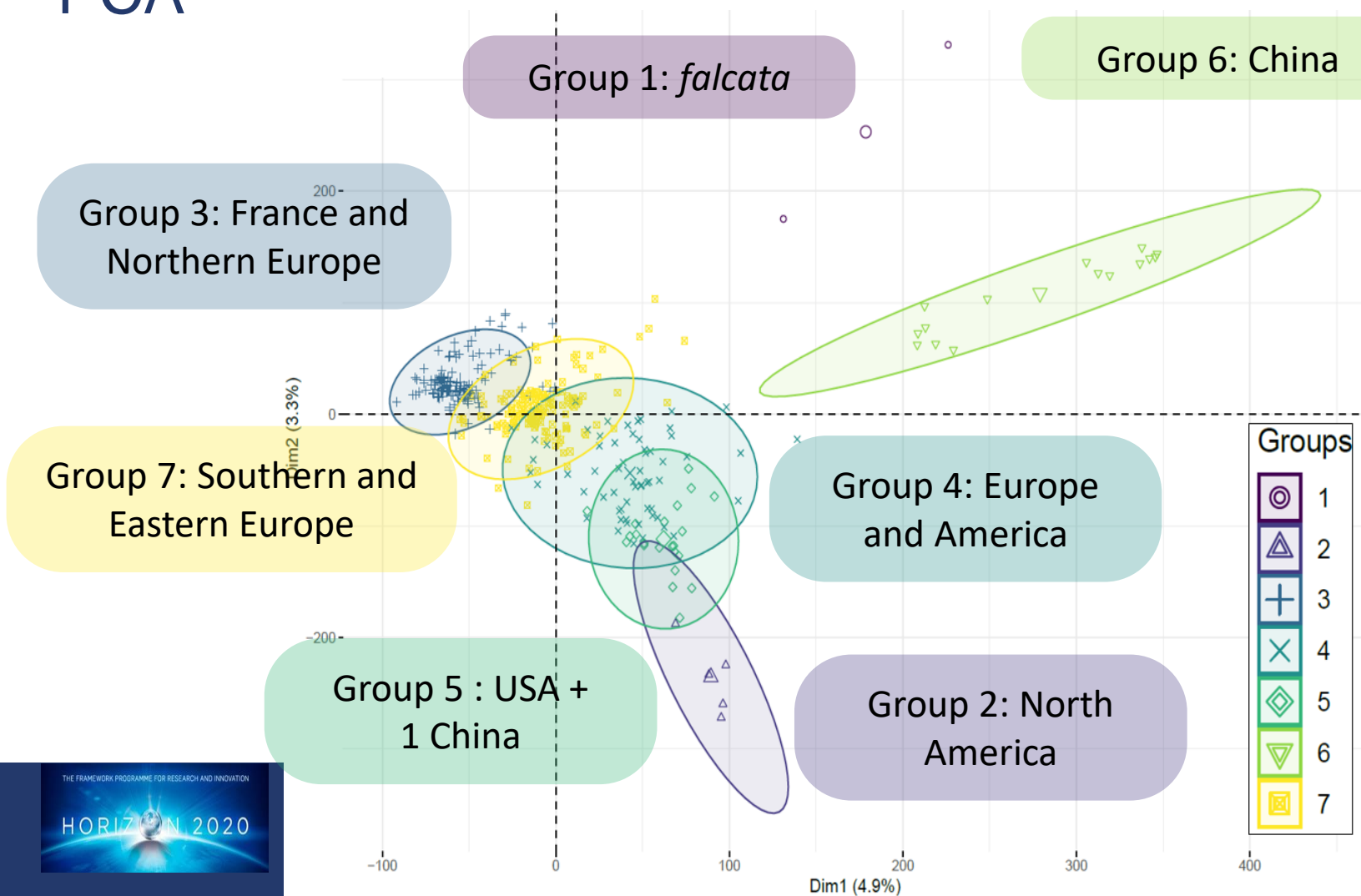## EUCLEG: a revision of diversity overview

400 accessions: landraces and cultivars, dormancy 3 – 7:

- Europe : 313
- North America : 45
- South America : 16
- China : 17
- Middle East : 3
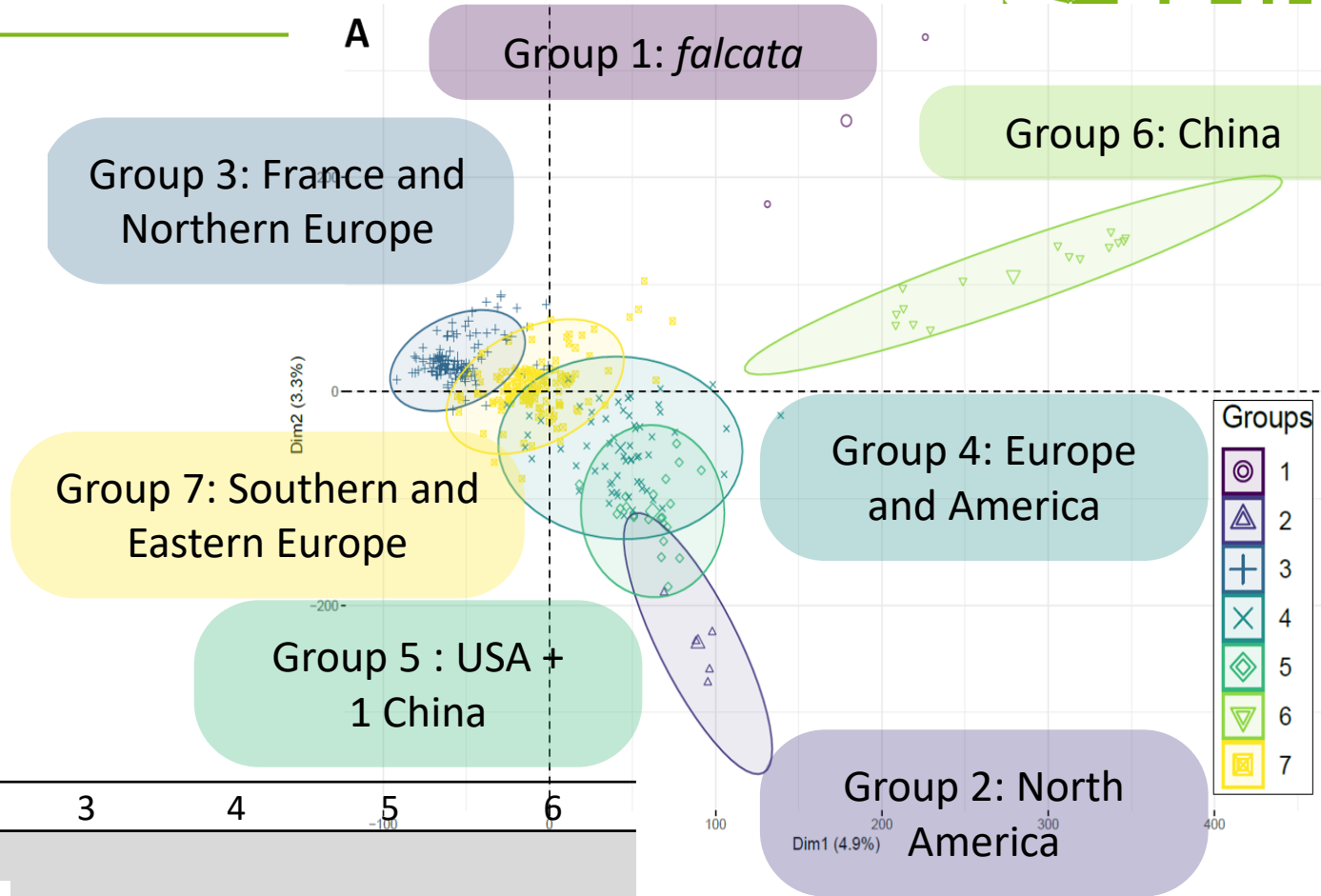- Japan : 1

Management of genetic diversity

PCA

Group 1: *falcata*

Group 6: China

Group 3: France and Northern Europe

Group 7: Southern and Eastern Europe

Group 4: Europe and America

Group 5 : USA + 1 China

Group 2: North America

Management of genetic diversity

PCA

A

Group 1: *falcata*

Group 6: China

Group 3: France and Northern Europe

Group 7: Southern and Eastern Europe

Group 4: Europe and America

Group 5 : USA + 1 China

Group 2: North America

Dim2 (3.3%)

Dim1 (4.9%)

Groups: 1, 2, 3, 4, 5, 6, 7

Clusters	1	2	3	4	5	6
1						
2						
3		0.017				
4		0.013	0.003			
5		0.004	0.006	0.002		
6		0.026	0.019	0.012	0.015	
7		0.016	0.001	0.001	0.005	0.015

18

Management of genetic diversity
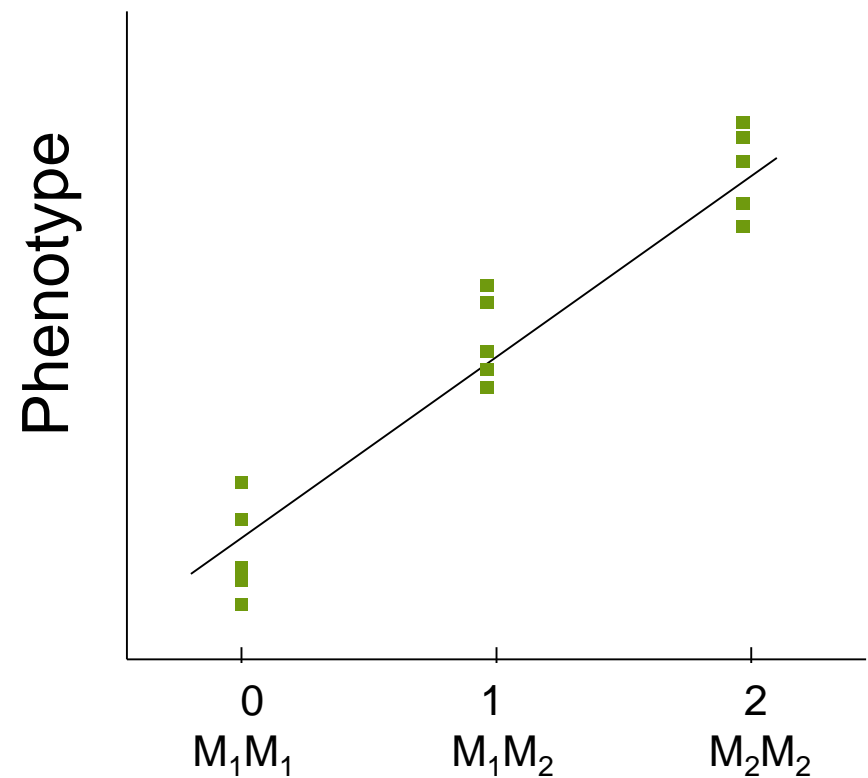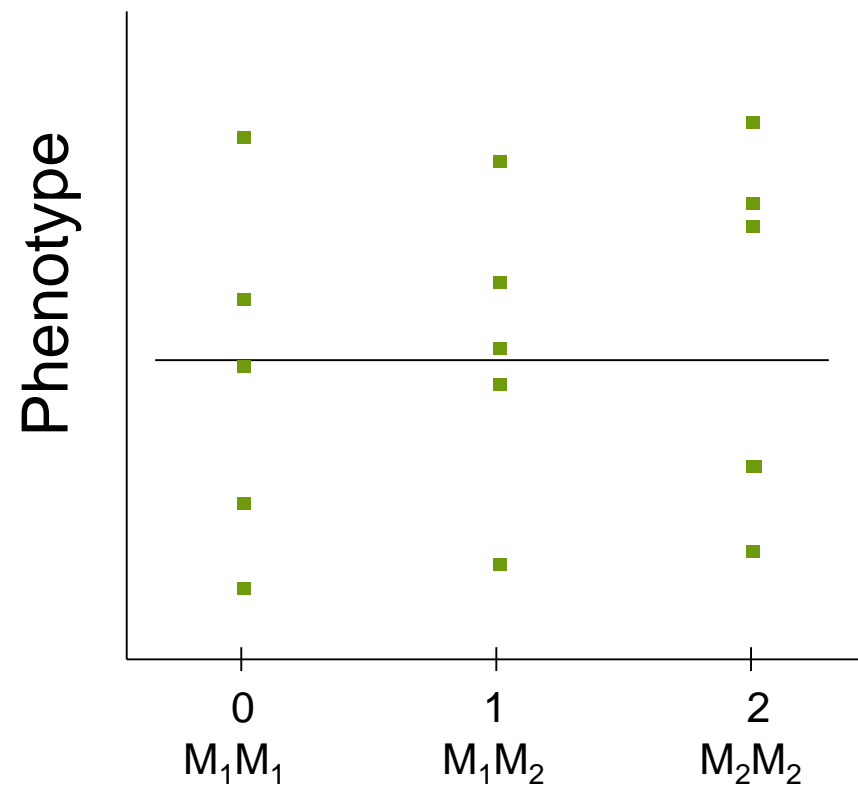
PCA : accessions colored with autumn dormancy score

Management of genetic diversity
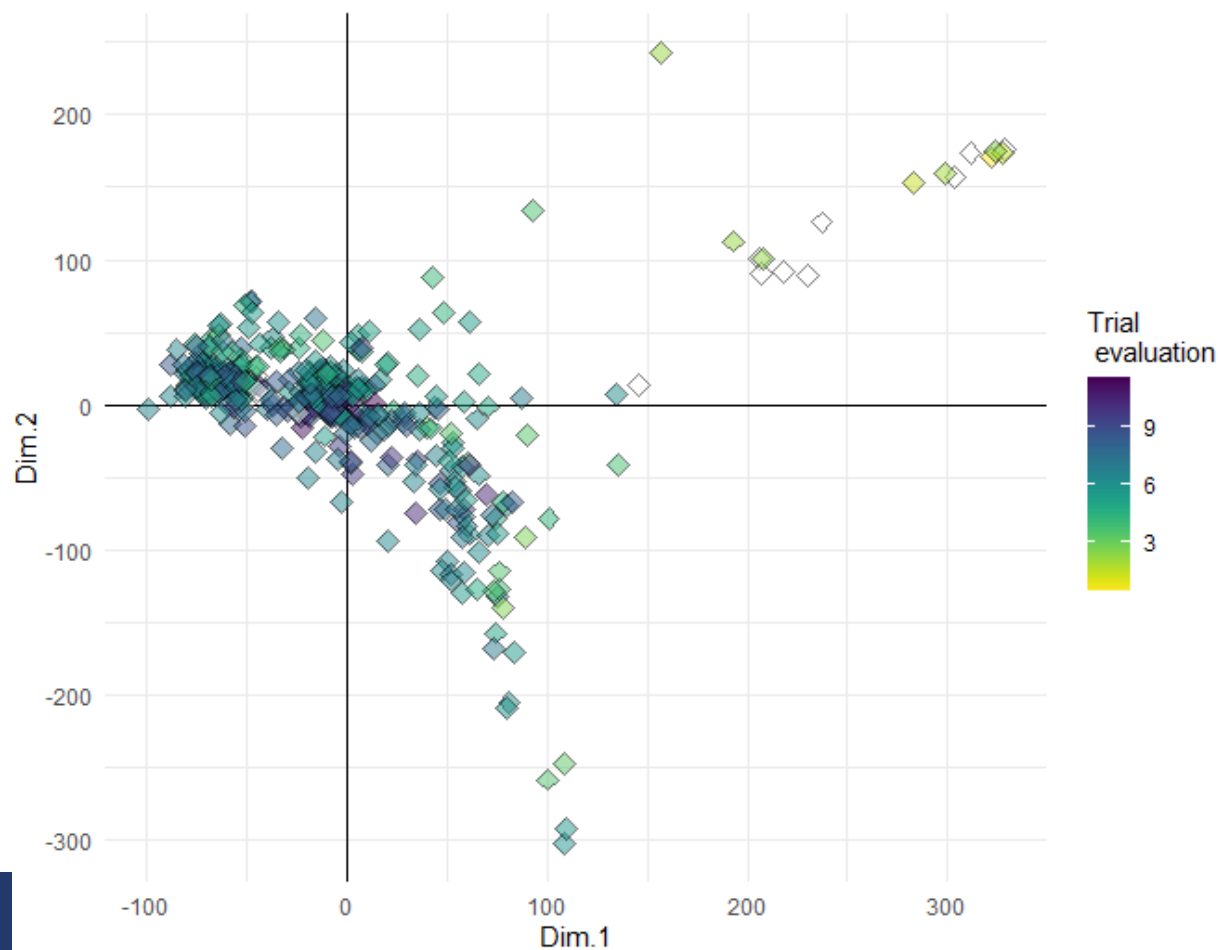
EUCLEG: a revision of diversity overview

- Diversity: China < > Europe + America

- Diversity: Europe <> America

- Structure is not associated to autumn dormancy

Genome wide association study
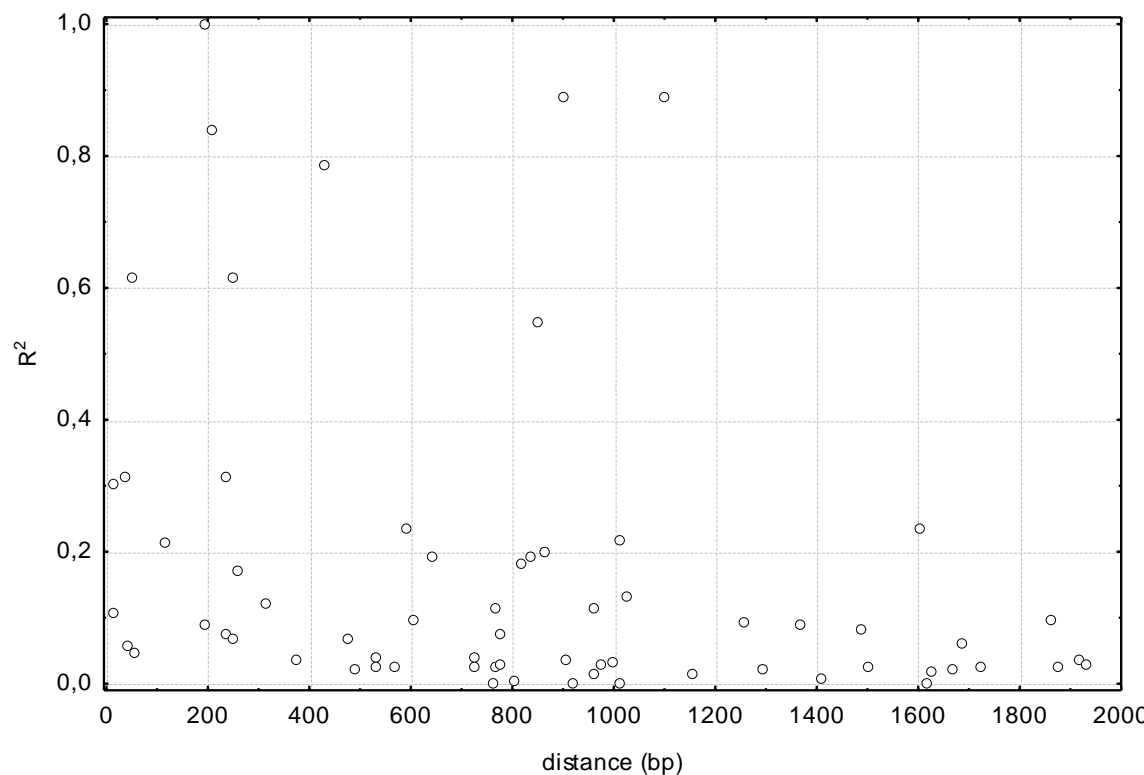
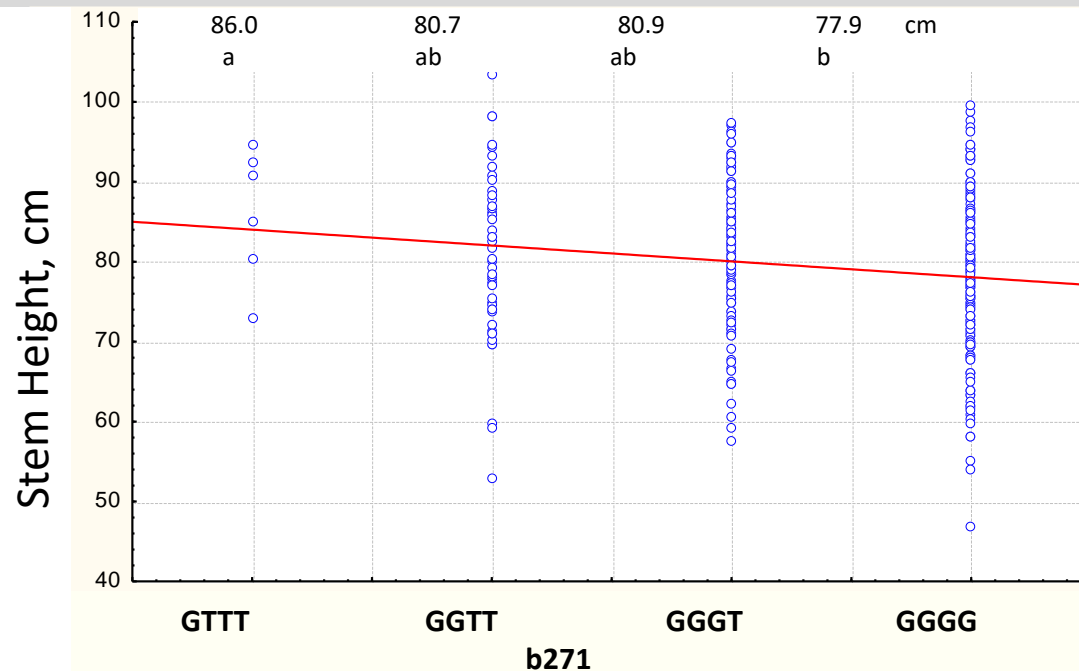For each marker: is it associated to trait variation?

Genome wide association study

Before EUCLEG:

- Low marker density

- Short linkage disequilibrium

12 SNP (66 pairs) in Constans-like gene (Herrmann et al., 2010)

to decrease the EU's
e European Union's
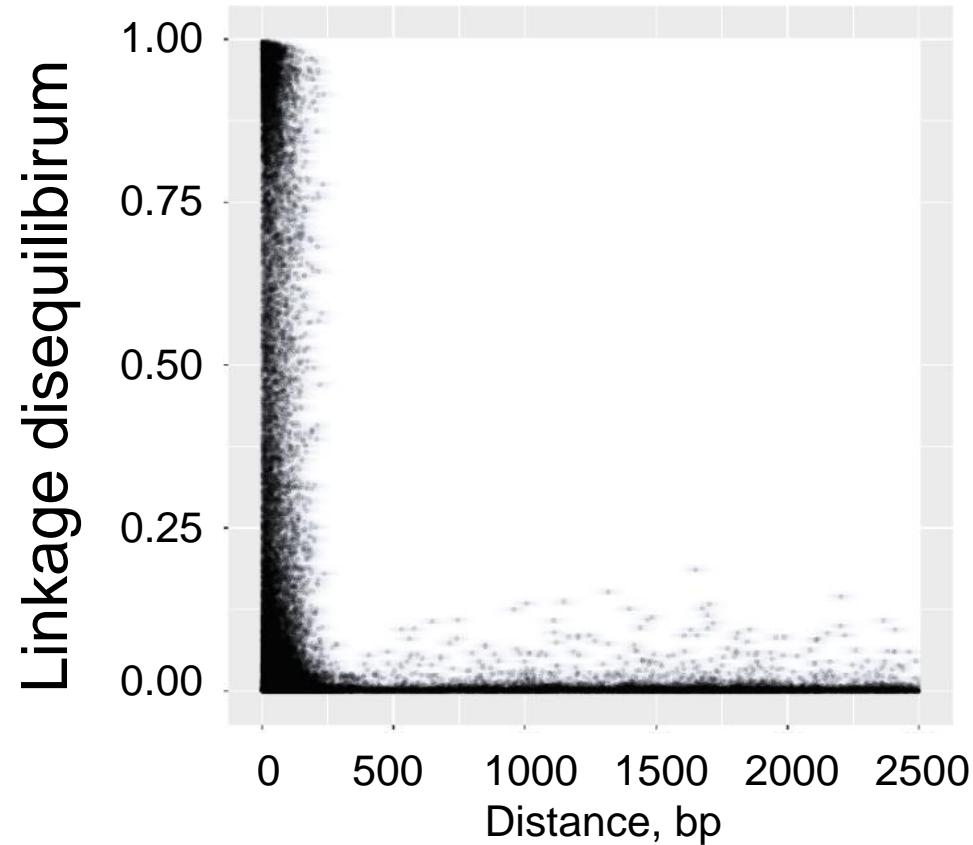
Before EUCLEG:

- Low marker density
- Short linkage disequilibrium
- → Candidate gene approach only

Constans-like gene (Herrmann et al., 2010)

EUCLEG: Short linkage disequilibrium over the genome

Horizon 2020 of European Union: Call 2016, SFS 44 : "A joint plant breeding programme to decrease the EU's and China's dependency on protein imports" I *This project has received funding from the European Union's Horizon 2020 Programme for Research & Innovation under grant agreement n°727312.*

24

Yield and quality (proteins, fibres, saponins)

400 accessions x 2 locations x 2 years

+   100 accessions x 3 locations x 2 years

Germination

Disease resistance: fusarium, anthracnose

Drought and P tolerance

Drought x fusarium

Genome wide association study

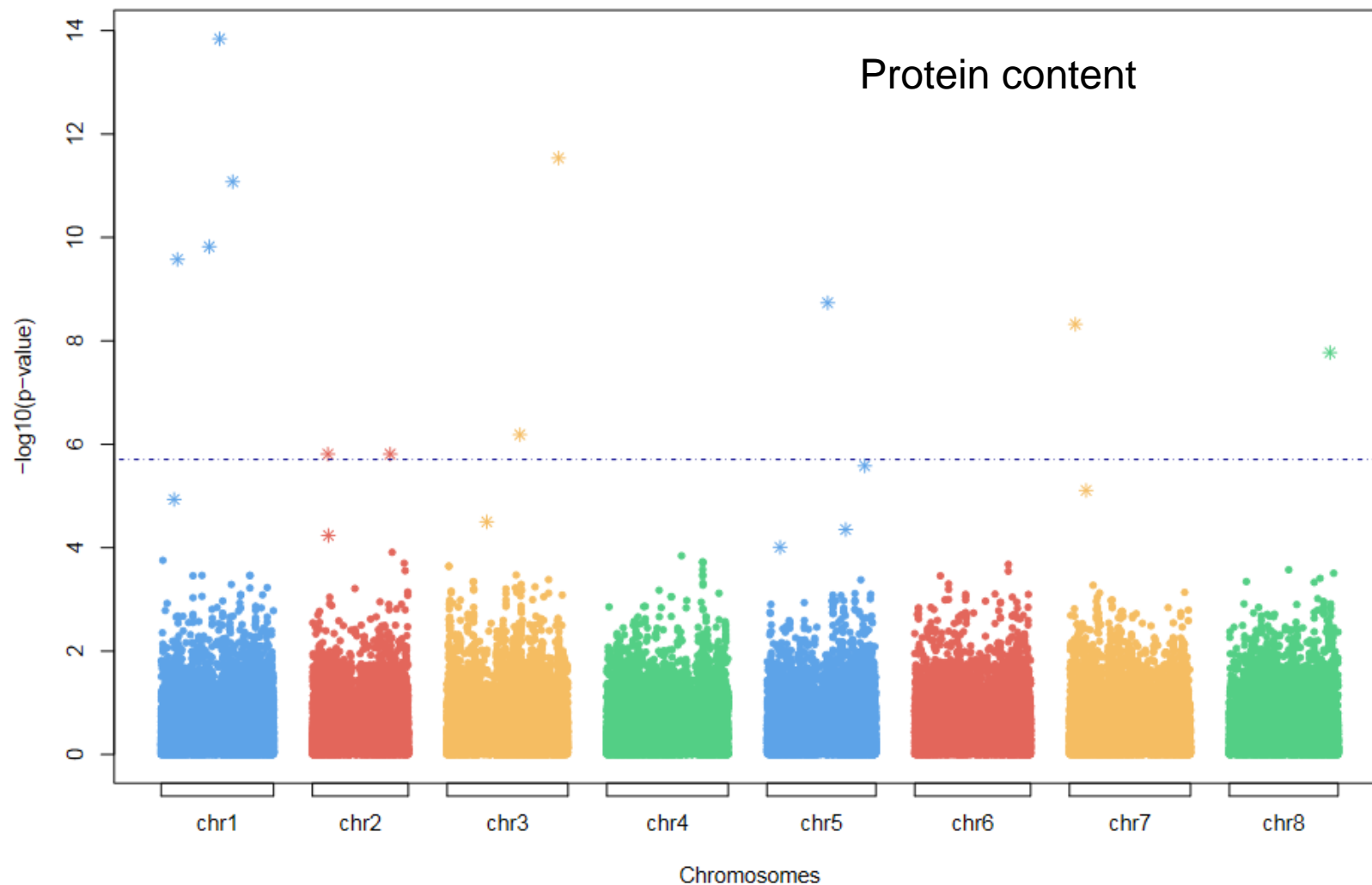EUCLEG: Detection of major QTL

400 accessions x 2 locations x 2 years

Method: Multi Locus Mixed Model (MLMM)
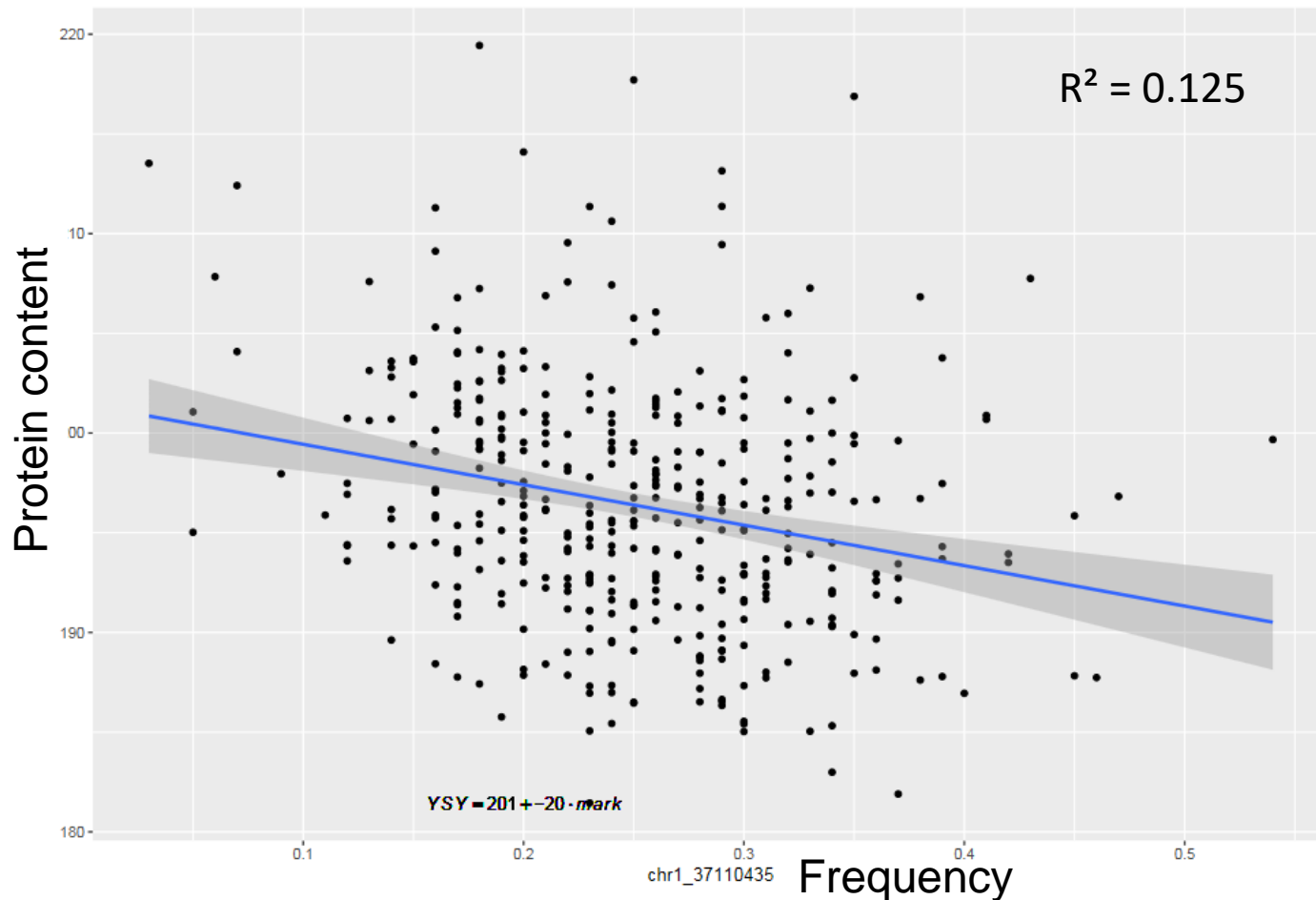
Genome wide association study

EUCLEG: Detection of major QTL

Protein content

Genome wide association study

EUCLEG: Detection of major QTL

- Up to 10 – 20% of variation

Genomic selection

Breeding population

Crossings

Candidates for selection

Genotyping

Phenotyping

Cross-validation

Prediction model

Genomic selection

Genotyping
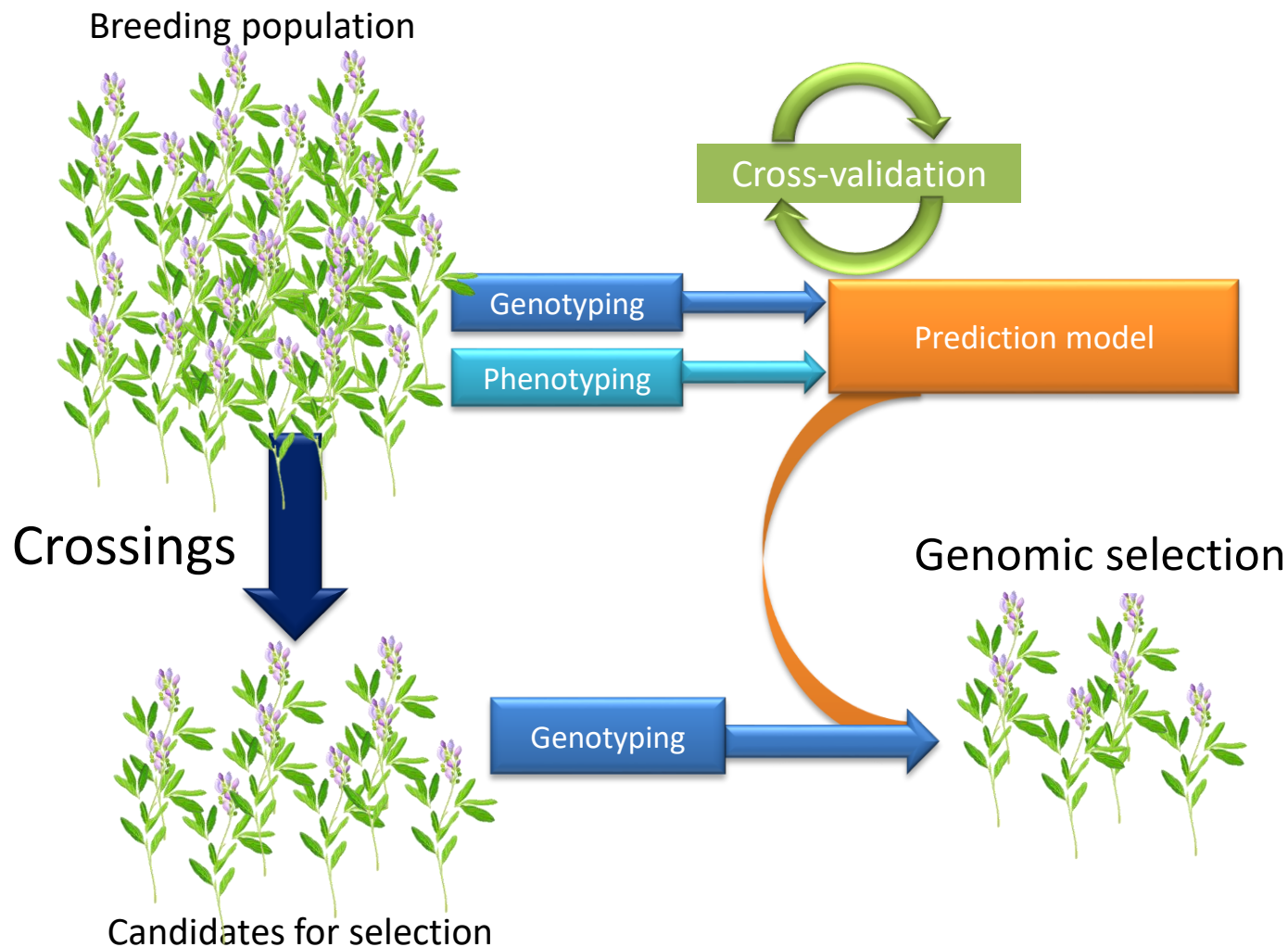
Genomic selection

Before EUCLEG

- 8 – 44 K SNP, 75 – 244 individuals

- Promising results, predictive ability ~ 30%

Genomic selection

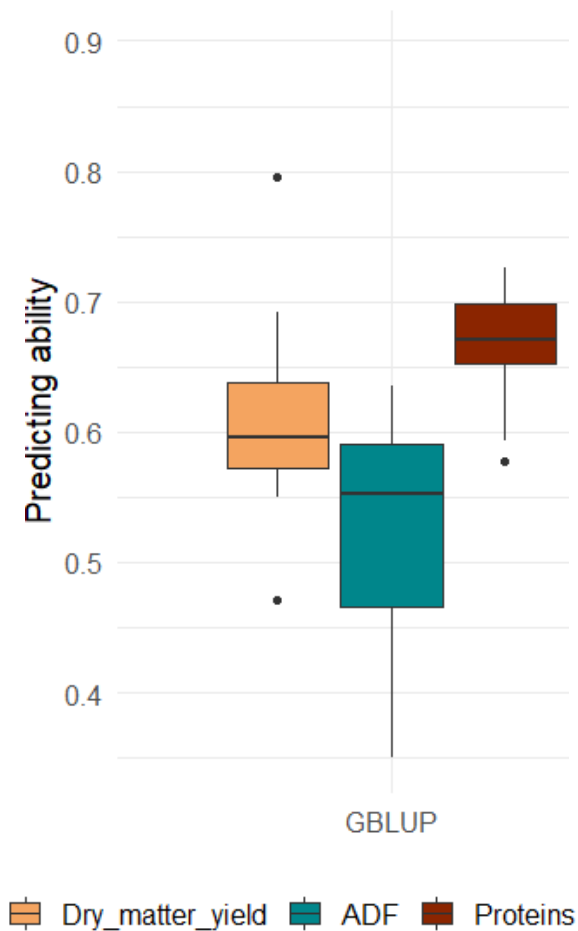EUCLEG
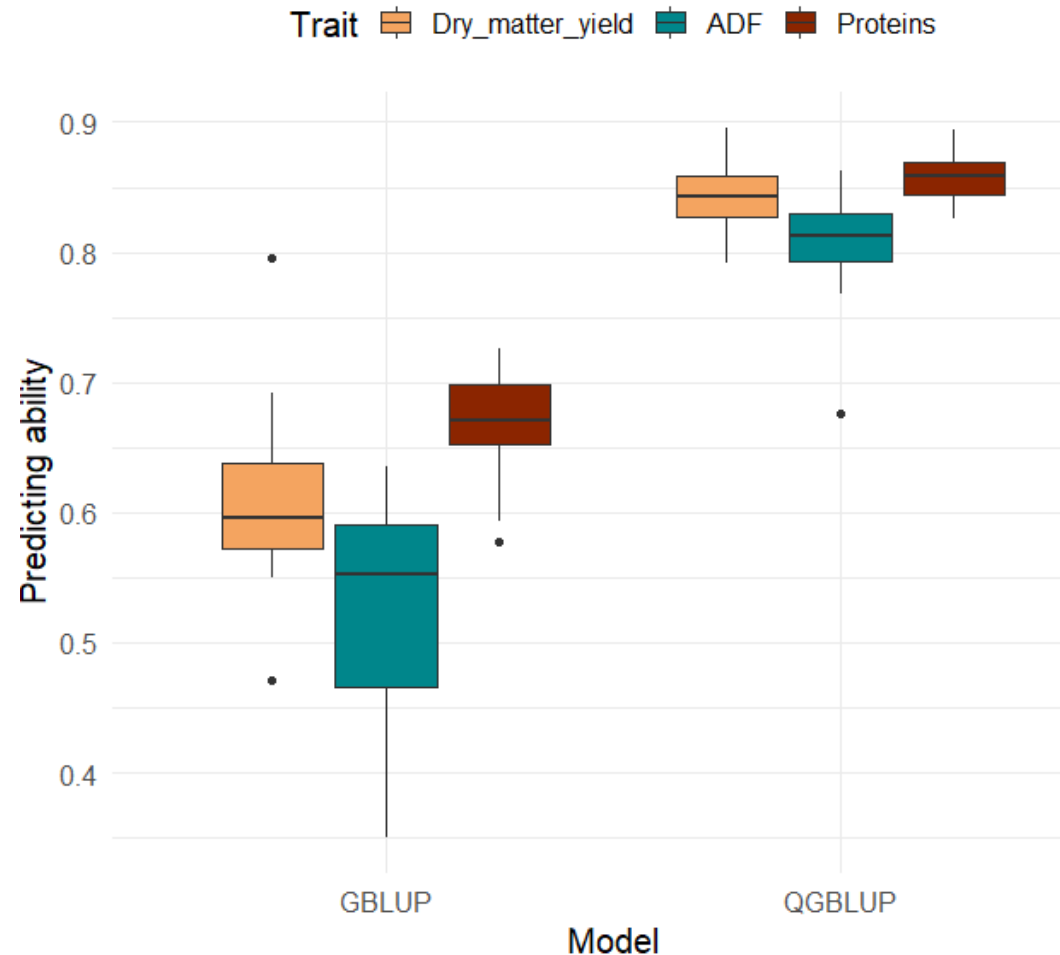
- GBLUP

- A good predicting ability:
0.52 < P < 0.66

Genomic selection

Integration of the QTL information in the prediction model (QGBLUP)

→ P > 0.80

Consequence in breeding

GBS are efficient markers

- Allele dosage in individuals

- Allele frequency in populations

- Genome coverage

- Low missing data

→ useful to analyse genetic diversity and genetic determinism of breeding traits

Consequence in breeding

Management of genetic diversity

- Some specialisation of the breeding pools in EU, America, China

- GBS markers to decide on the introduction of new genetic diversity in a breeding pool

GS models provide high predictive ability

- Even higher with the inclusion of QTL effect

→ To be used to select promising individuals in breeding pools

Still to be done

- Extend the analysis of alfalfa diversity from dormancy 3-7 to the whole species complex

- Improve cost-efficiency of genotyping

- Calculate genetic gain with GS

- Estimate cost-efficiency of GS

- Implement genomic selection in breeding programmes

GS in breeding programmes

Introduction of new origins

Genotyping
Analysis of diversity
Genomic prediction
→ Choice of plants
0.5 year
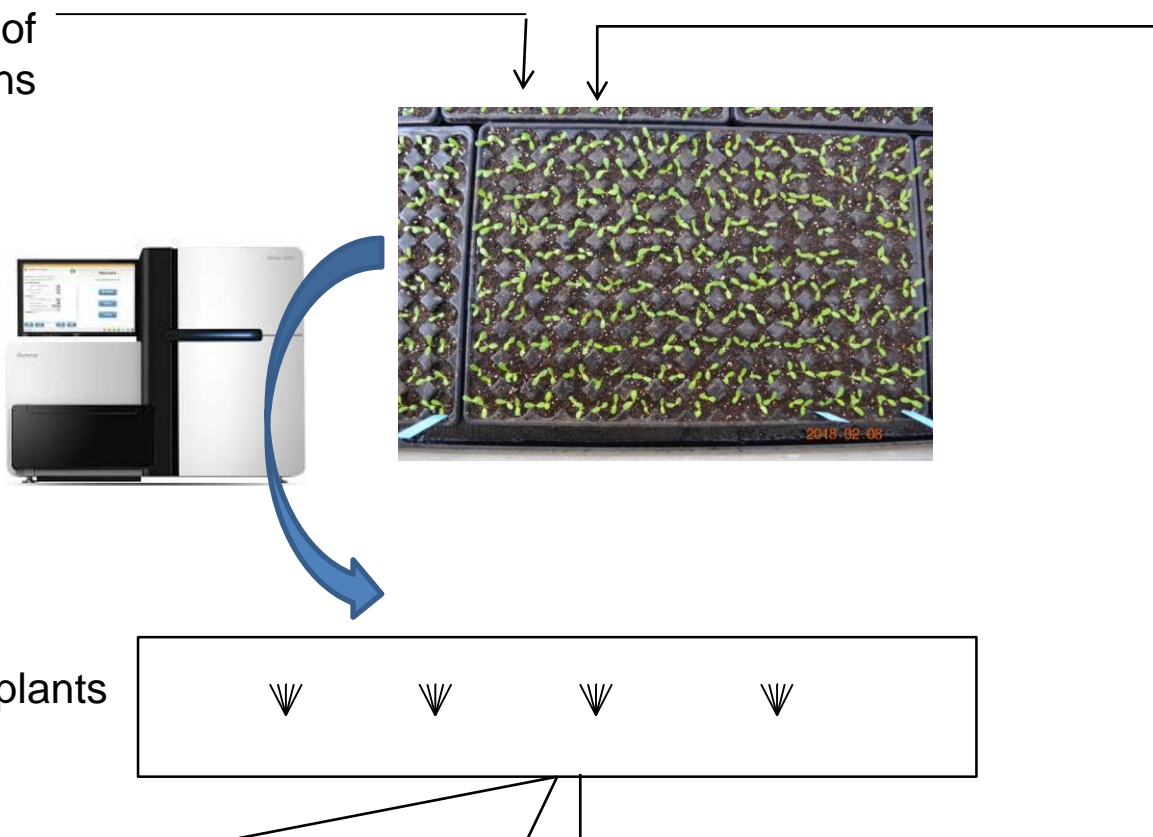
(5 000 to 15 000)

Polycross of the best plants
1 year

(20 to 30)

Evaluation in multi-site trials
3 years

Varieties

GS in breeding programmes

EUC LEG

Strength

- Reduced field work
- Early selection for all predicted traits
- Reduced number of years
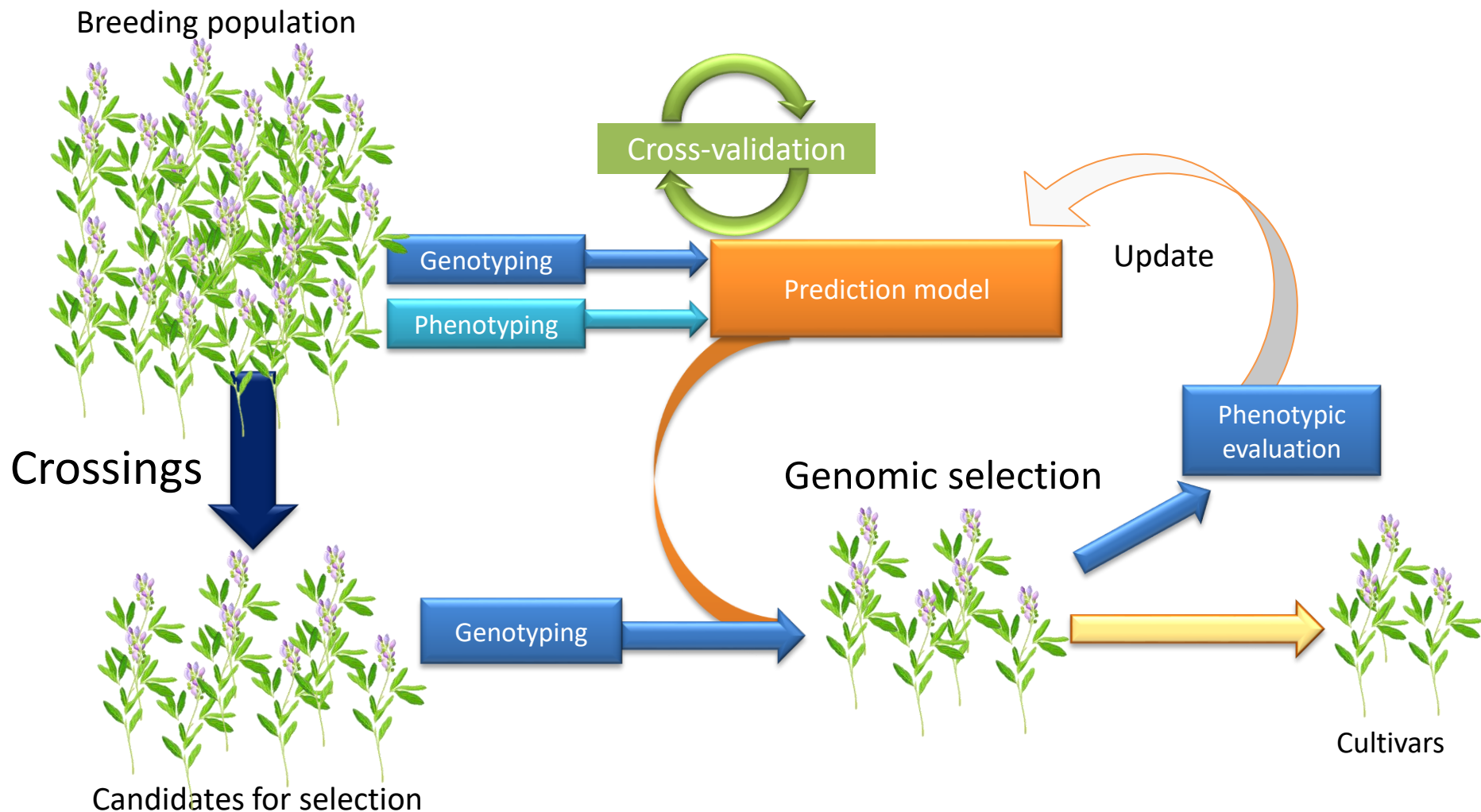- Fixation of positive alleles is quick

Weakness

- No prediction for some traits
- Staffs have to get new skills

Genetic gain?

Cost efficiency ?

Still to be done

- Extend the analysis of alfalfa diversity from dormancy 3-7 to the whole species complex

- Improve cost-efficiency of genotyping

- Calculate genetic gain with GS

- Estimate cost-efficiency of GS

- Implement genomic selection in breeding programmes

- Imagine the updating of GS equations

Genomic selection

Horizon 2020 of European Union:
Call 2016, SFS 44: "A joint plant breeding programme to decrease the EU's and China's dependency on protein imports"

EUCLEG.eu