Conservation of genetic diversity in natural populations and ecosystems

René Smulders


1

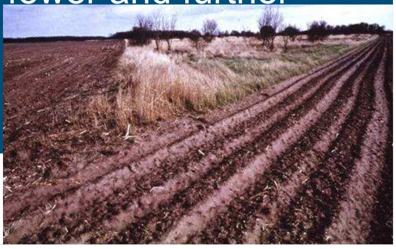

2

Nature in the Netherlands during the 20th century

- Strong decrease in area
- Loss of quality
- Strong increase in fragmentation

Hence

- Populations became smaller, fewer and further apart



Isolation-by-distance analysis of Moor frog (*Rana arvalis*) populations is sensitive to landscape heterogeneity

- Paul Arens, Wendy van ’t Westende, Ben Vosman, René Smulders (PRI)
- Theo van der Sluis, Claire Vos (Alterra)

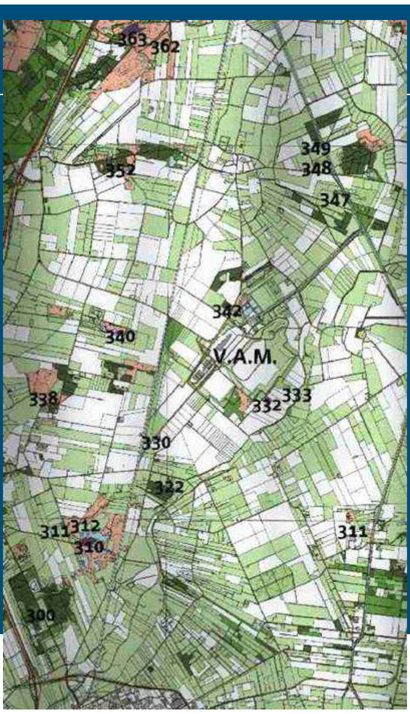
Moor frog (*Rana arvalis*) studies

- Endangered
- Breeds in ponds > populations defined
- Breeds in 2nd and 3rd year
- Habitat preferences known

Research questions:

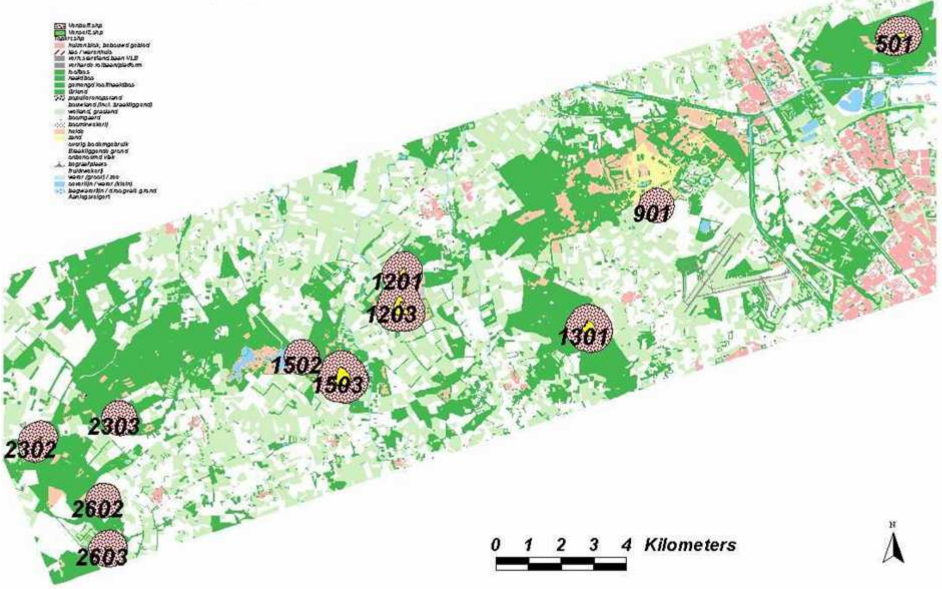
- How much population differentiation?
- Is there a relationship
  - Between geographic and genetic distance?
  - For certain landscape types?

Study area in Drenthe



PLANT RESEARCH INTERNATIONAL
WAGENINGEN UR

Onderzoekspopulaties Heikikker Noord-Brabant



Moor frog (*Rana arvalis*) studies

- Drenthe: fragmentation since 1930's ~ 20-30 generations
  - 11 populations sampled (232 egg clumbs)
- Kempen, Noord-Brabant: more severe fragmentation, since 1900
  - 12 Populations sampled (202 egg clumbs)
- Studied with 8 microsatellite markers

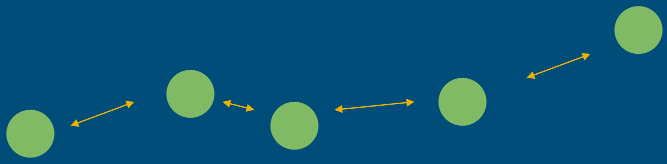
PLANT RESEARCH INTERNATIONAL
WAGENINGEN UR

Genetic differentiation

	Drenthe		
	F_{ST}	F_{is}	F_{it}
Across all loci	**0.022**	**0.096**	**0.116**
95% CI	**0.013 – 0.034**	**0.027 – 0.180**	**0.051 – 0.198**
Pairwise distances	**-0.011 to 0.061**		
	Noord-Brabant		
	F_{ST}	F_{is}	F_{it}
Across all loci	**0.060**	**0.086**	**0.141**
95% CI	**0.041 – 0.092**	**-0.008 – 0.188**	**0.059 – 0.230**
Pairwise distances	**0 to 0.175**		

PLANT RESEARCH INTERNATIONAL
WAGENINGEN UR

Isolation-by-distance (IBD)



Genetic similarity between neighbouring populations

decreases with distance

IBD

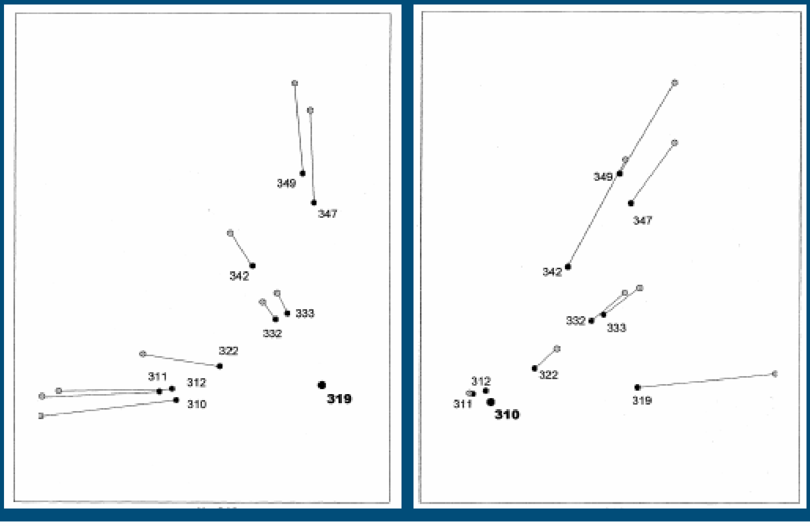
- IBD analysis assumes
  - uniformity of drift between populations
  - a homogeneous landscape structure, and
  - a common history

Landscape structure

- Geographical distance between pairs of ponds
- Relative amounts of:
  - negative linear elements (roads, railways) RL
  - negative area (agricultural fields) RA
  - positive linear elements (hedgerows) CL

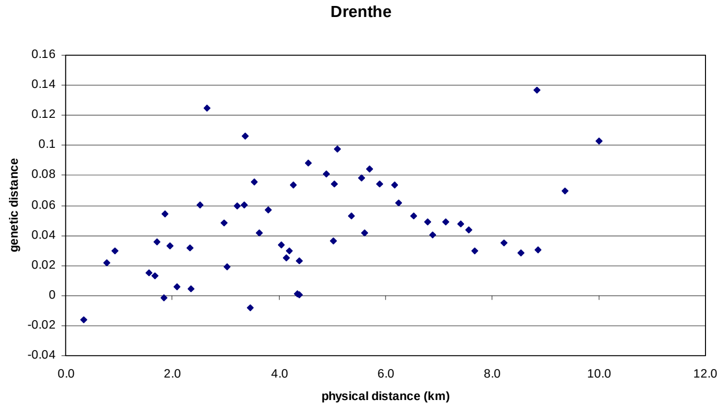
PLANT RESEARCH INTERNATIONAL
WAGENINGEN UR

Visualisation of effect of landscape (D*RL)

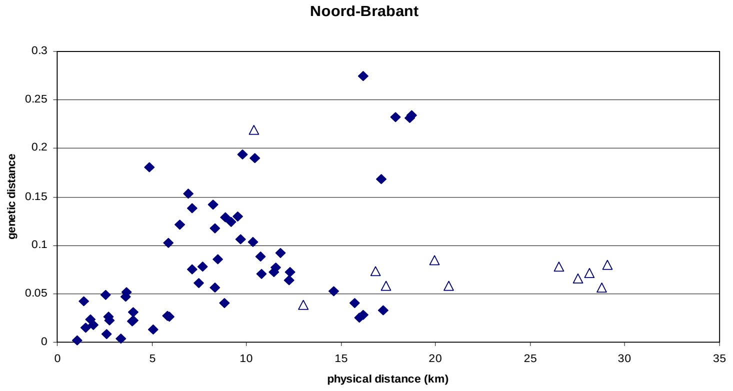


PLANT RESEARCH INTERNATIONAL
WAGENINGEN UR

Isolation by distance



Isolation by distance



Single Mantel tests

Variable	Description of variable	Correlation with genetic distance	Percentage of variation explained	Significance of correlation (single Mantel test)
Drenthe				
D	Distance	0.66	42.2	P<0.0001
D*RL	Distance corrected for negative linear elements	0.62	37.7	P<0.0001
D*RNA	Distance corrected for negative area	0.64	39.7	P<0.0001
D*CL	Distance corrected for positive linear elements	0.44	17.7	P<0.0001
D^*R_{INT}	Distance corrected for positive and negative elements	0.47	20.5	P<0.0001
Noord-Brabant[1]				
D	Distance	0.69	47.3	0.0003
D*RL	Distance corrected for negative linear elements	0.69	46.1	0.0003
D*RNA	Distance corrected for negative area	0.62	37.6	0.0004
D*CL	Distance corrected for positive linear elements	0.70	48.1	0.0003
D^*R_{INT}	Distance corrected for positive and negative elements	0.65	41.2	0.0006

[1] Population 0501 excluded

PLANT RESEARCH INTERNATIONAL
WAGENINGEN UR

Multiple Mantel tests

- in Drenthe P > 0.24, in Noord-Brabant P > 0.22

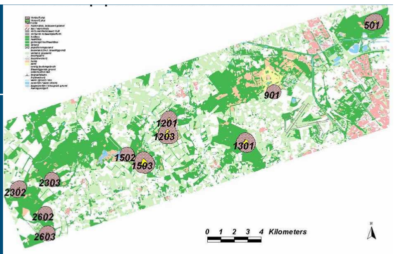
- An earlier, smaller, study in Drenthe did show that D*RL added significantly to a model with geographical distance D
- When jack-knifing populations in Noord-Brabant, D*RL did add significantly to a model with D when when population 0901 was excluded

PLANT RESEARCH INTERNATIONAL
WAGENINGEN UR

1999 vs 1995 Drenthe study

- 1999 sampling was larger: more samples, more populations, more markers
- Study produced more accurate results (smaller variation in F-statistic estimates)
- yet - effect of roads (D*RL) not significant

Noord-Brabant



- Population 0501 was isolated much earlier, and up to 17% of the habitat has remained (elsewhere 1-2%)
  - Outside the set of populations for IBD
- Population 0901 is in a military training area, has the largest distance to other populations, and is surrounded by a high proportion of unsuitable land
  - Distorts D*RL significance

Conclusion

  - *IBD analysis assumes*
    - *uniformity of drift between populations*
    - *a homogeneous landscape structure, and*
    - *a common history*
- Differences in census population sizes (and in drift) can not be weighed in current IBD analysis
- Heterogeneity in the landscape can be compensated
  - by using permeability measures instead of geographical distances
  - By jackknifing all possible subsets of populations local outlier populations can be identified and significant IBD unveiled
- Landscape history data, such as old maps, can help understanding history

Multilocus microsatellite haplotype sharing as a tool to estimate gene flow and introgression.
The case of *M. domestica* and *M. sylvestris*

- René Smulders, Li Yinghui, Eric van de Weg, Wim Koopman, Ben Vosman (PRI)
- Els Coart, Isabel Roldán-Ruiz (INBO-Plant, Melle, Belgium)

The case of *M. domestica* and *M. sylvestris*

- Only a few 100’s wild apple trees remain in the Netherlands
- They occur in forests no longer used for coppice, hence they grow reproduce poorly
- Conserve them as one population?

- Are they threatened by introgression from cultivated apple?

- Collection was done by Hennie Ketelaar, Chris Rövekamp, René van Loon, and Bert Maes
- Based on previous inventories

PLANT RESEARCH INTERNATIONAL
WAGENINGEN UR

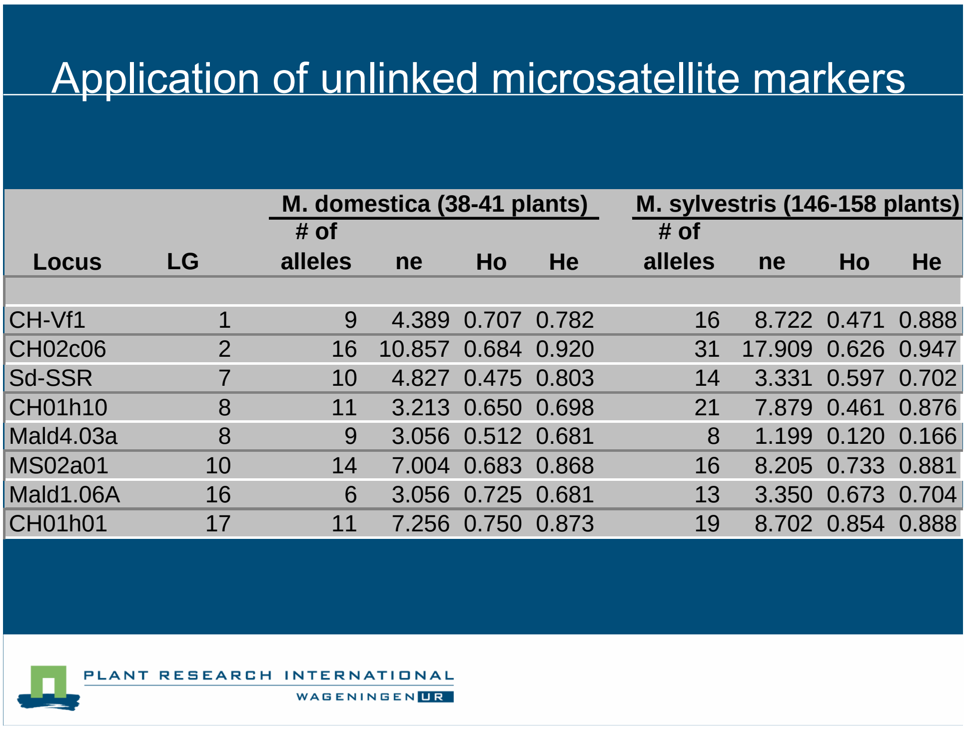
Minister opent 'bronnen voor nieuwe natuur'

Genenbank in Dronten voor behoud van inheemse bossen en struiken

Minister Veerman van Landbouw, Natuur en Voedselkwaliteit heeft maandag de genenbank 'bronnen voor nieuwe natuur' voor inheemse bomen en struiken geopend. De genenbank bestaat uit 3500 bomen en struiken van bijna 60 verschillende soorten en is aangeplant in een bestaand bos, het Roggebotzand in Dronten.

LNV-minister Veerman (midden) bij de opening van de genenebank van Staatsbosbeheer nabij Dronten (foto's Staatsbosbeheer)

12



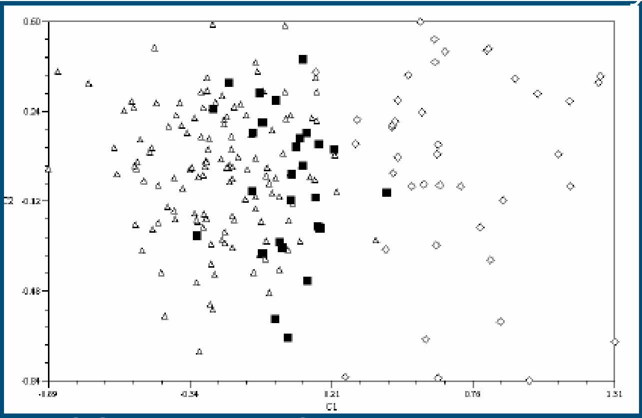
Application of unlinked microsatellite markers

		M. domestica (38-41 plants)				M. sylvestris (146-158 plants)			
Locus	LG	# of alleles	ne	Ho	He	# of alleles	ne	Ho	He
CH-Vf1	1	9	4.389	0.707	0.782	16	8.722	0.471	0.888
CH02c06	2	16	10.857	0.684	0.920	31	17.909	0.626	0.947
Sd-SSR	7	10	4.827	0.475	0.803	14	3.331	0.597	0.702
CH01h10	8	11	3.213	0.650	0.698	21	7.879	0.461	0.876
Mald4.03a	8	9	3.056	0.512	0.681	8	1.199	0.120	0.166
MS02a01	10	14	7.004	0.683	0.868	16	8.205	0.733	0.881
Mald1.06A	16	6	3.056	0.725	0.681	13	3.350	0.673	0.704
CH01h01	17	11	7.256	0.750	0.873	19	8.702	0.854	0.888

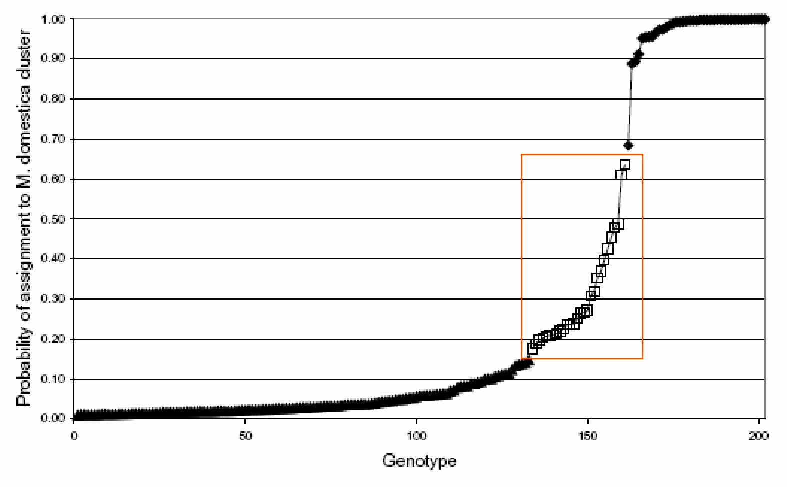
PLANT RESEARCH INTERNATIONAL
WAGENINGEN UR

13

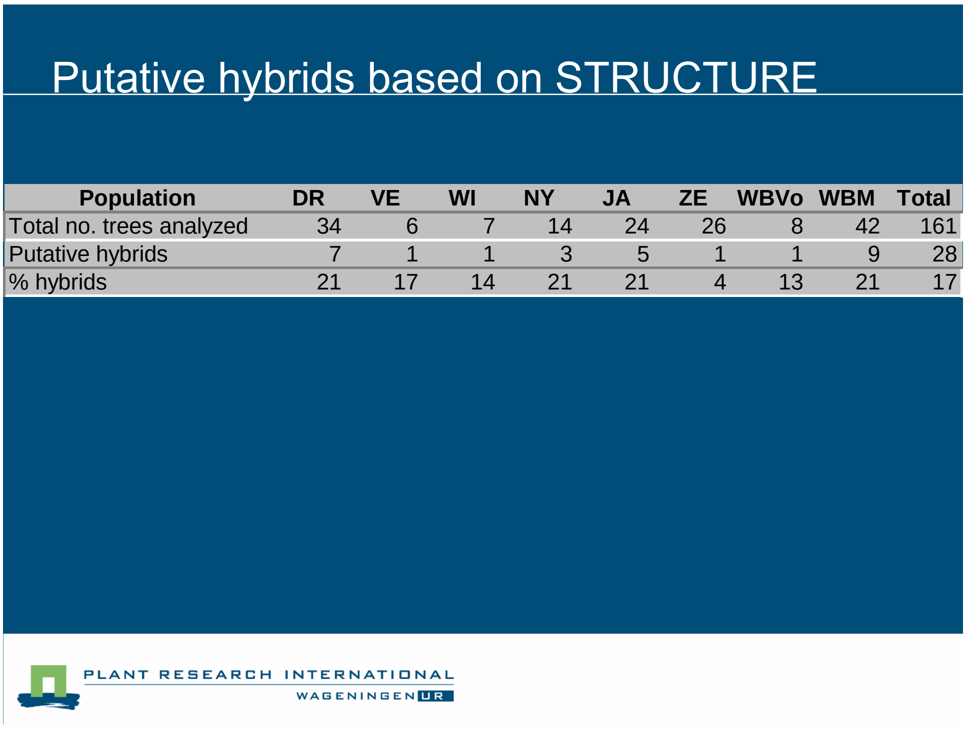
The case of *M. domestica* and *M. sylvestris*



- A PCO analysis of 8 unlinked microsatellite markers indicated two clear groups, but also a number of intermediate plants. Introgression?



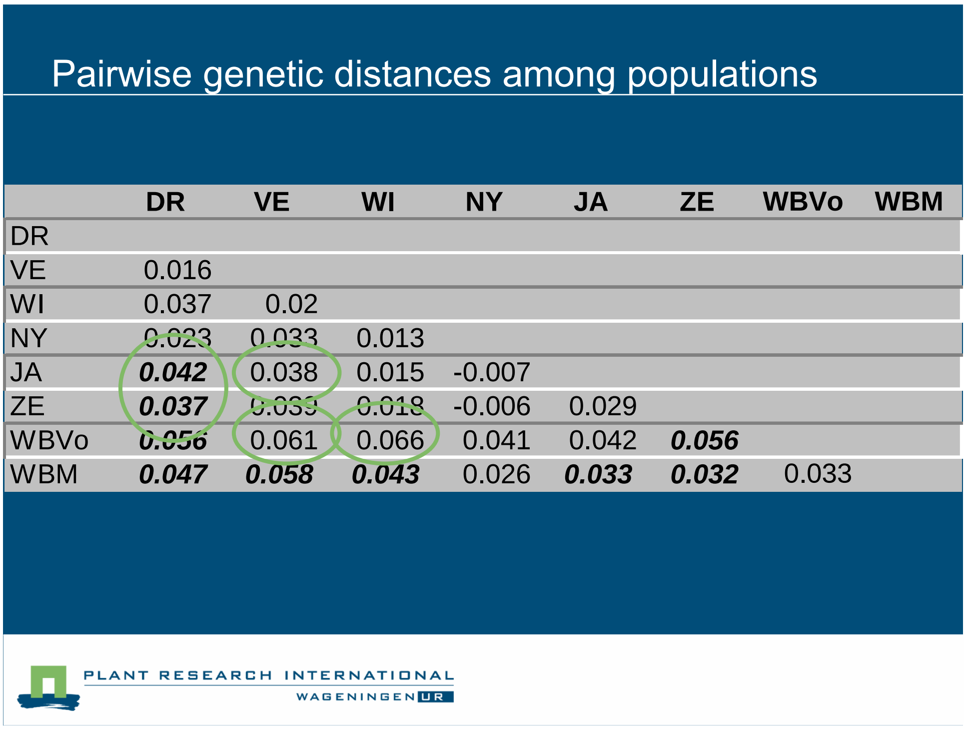
- *structure* analysis of 8 unlinked microsatellite markers indicated two clear groups, but also a number of intermediate plants. Introgression?



Putative hybrids based on STRUCTURE

Population	DR	VE	WI	NY	JA	ZE	WBVo	WBM	Total
Total no. trees analyzed	34	6	7	14	24	26	8	42	161
Putative hybrids	7	1	1	3	5	1	1	9	28
% hybrids	21	17	14	21	21	4	13	21	17

PLANT RESEARCH INTERNATIONAL
WAGENINGEN UR



Pairwise genetic distances among populations

	DR	VE	WI	NY	JA	ZE	WBVo	WBM
DR								
VE	0.016							
WI	0.037	0.02						
NY	0.023	0.033	0.013					
JA	***0.042***	0.038	0.015	-0.007				
ZE	***0.037***	0.039	0.018	-0.006	0.029			
WBVo	***0.056***	0.061	0.066	0.041	0.042	***0.056***		
WBM	***0.047***	***0.058***	***0.043***	0.026	***0.033***	***0.032***	0.033	

PLANT RESEARCH INTERNATIONAL
WAGENINGEN UR

Problems with population genetics of natural populations

- Populations are small --> difficult to obtain significant differentiation
- Genetic changes in populations are lagging behind (alleles are not lost instantaneously)
- Exact history of populations is not known
- Therefore: estimates of genetic differentiation (Fst) cannot directly be translated into extent of dispersal

PLANT RESEARCH INTERNATIONAL
WAGENINGEN UR

Problems with introgression

- Alleles that are common in one species but rare in another, can be indicative of introgression
- But they can also be there from the origin of the soecies onward, or came into existence through mutations afterwards

PLANT RESEARCH INTERNATIONAL
WAGENINGEN UR

Use of multilocus haplotypes

- A haplotype has a limited life span, as it is destroyed by recombination
- Hence, haplotype sharing can be used to determine
  - Recent introgression between species
  - Level of recent gene flow

PLANT RESEARCH INTERNATIONAL
WAGENINGEN UR

Linkage Group 10 of M. domestica

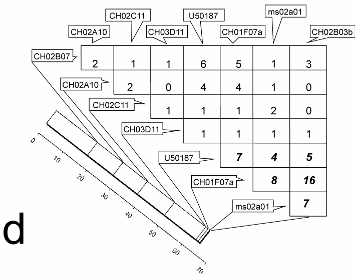
		M. domestica (91-97 plants)				**M. sylvestris (146-161 plants)**			
Locus	**LG (cM)**	**# of alleles**	**ne**	**Ho**	**He**	**# of alleles**	**ne**	**Ho**	**He**
CH02b07	10 (17.1)	16	5.348	0.790	0.817	16	6.417	0.800	0.847
CH02a10	10 (28.0)	17	5.363	0.802	0.818	14	6.497	0.727	0.849
CH02c11	10 (38.1)	15	9.422	0.785	0.899	13	6.219	0.747	0.842
CH03d11	10 (49.5)	13	3.948	0.742	0.751	10	4.078	0.635	0.757
U50187	10 (62.7)	11	4.149	0.537	0.763	26	8.156	0.634	0.880
CH01f07a	10 (63.6)	12	6.104	0.629	0.841	19	8.575	0.707	0.886
MS02a01	10 (64.2)	15	5.174	0.701	0.811	16	8.205	0.733	0.881
CH02b03b	10 (64.2)	13	6.189	0.760	0.843	17	8.485	0.647	0.885

PLANT RESEARCH INTERNATIONAL
WAGENINGEN UR

Use of multilocus haplotypes

- 8 microsatellite markers on LG10
- Variable distance (cM) -> variable expected life span (generations)

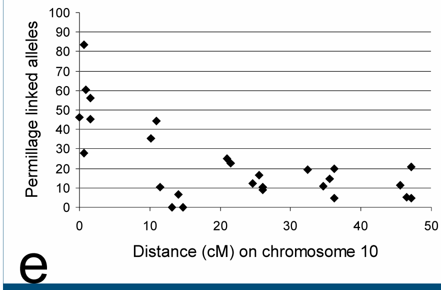
LG 10 / M. sylvestris



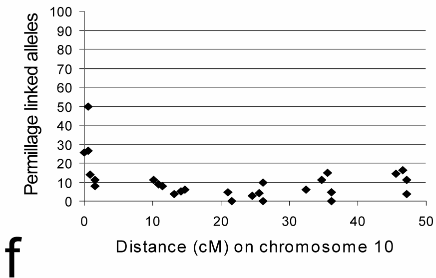
	CH02A10	CH02C11	CH03D11	U50187	CH01F07a	ms02a01	CH02B03b
CH02B07	2	1	1	6	5	1	3
CH02A10		2	0	4	4	1	0
CH02C11			1	1	1	2	0
CH03D11				1	1	1	1
U50187					***7***	***4***	***5***
CH01F07a						***8***	***16***
ms02a01							***7***

LD

LG 10 / M. domestica



LG 10 / M. sylvestris



- Steps:
  - Infer haplotypes from genotypic data using PHASE
  - Identify haplotypes present in more than one copy
  - Calculate sharing of haplotypes between M. domestica and M sylvestris, and among M. sylvestris populations
  - Repeat this procedure for each haplotype length
  - Calculate total length of haplotypes shared

PHASE (Stephens et al. 2001, 2003)

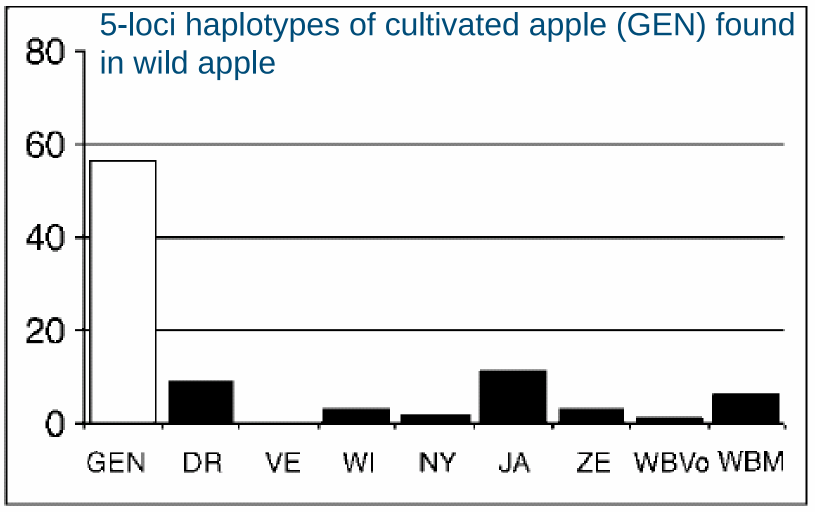
```
BEGIN BESTPAIRS1
0 #Cox
(6) 5 8 (8) 22 14 5 12
(13) 5 8 (11) 22 3 6 1
0 #Discovery
10 (9) 13 8 27 13 8 12
16 (5) 11 8 22 13 9 11
0 #Elstar
6 8 11 8 22 14 5 13
6 8 16 8 17 14 5 12
0 #Fiesta
7 22 14 10 22 13 20 11
13 5 8 8 22 3 6 1
0 #Golden
6 8 16 8 17 5 16 3
10 8 9 8 23 15 5 13
```

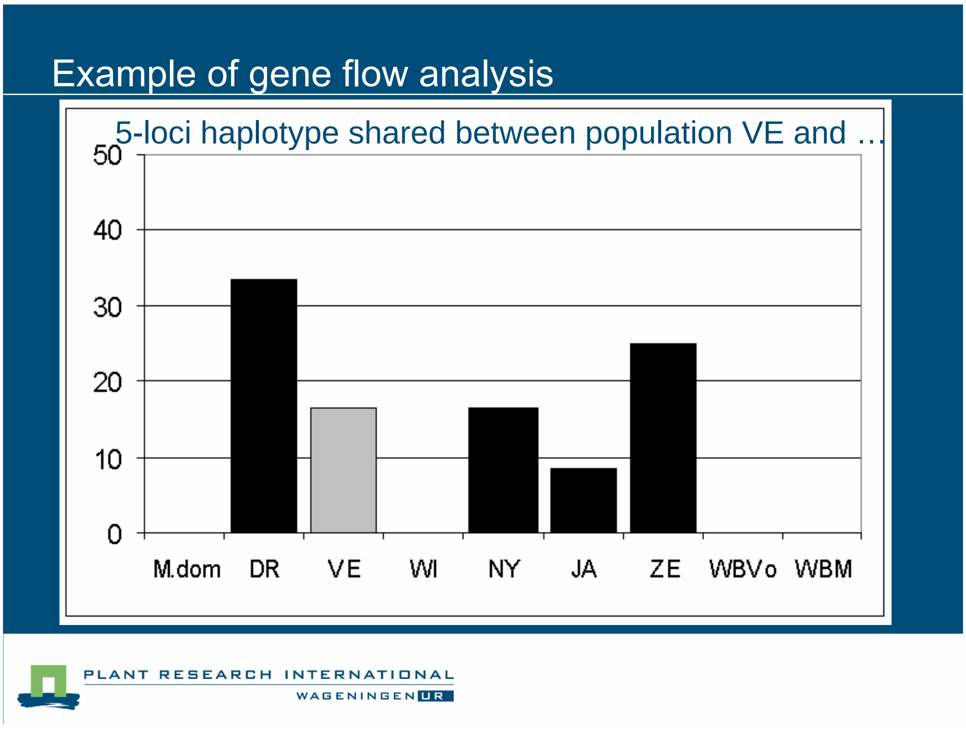
```
BEGIN PHASEPROBS

0.74 = = 0.68 = 1.00 1.00 1.00

0.90 0.89 1.00 = 1.00 = 1.00 0.96

= = 0.99 = 0.99 = = 1.00

1.00 1.00 1.00 1.00 = 1.00 1.00 1.00

0.94 = 0.97 = 1.00 1.00 1.00 1.00
```

```
0 #NY-15
= (18) ? 13 26 11 11 9
= (19) ? 8 9 19 21 17
0 #NY-16
(7) (14) (11) (8) (17) = ? =
(11) (10) (16) (7) (13) = ? =
0 #NY-18
(8) = (18) (6) 17 ? 14 7
(9) = (17) (8) 8 ? 8 12
0 #NY-19
11 11 ? = = ? ? =
13 8 ? = = ? ? =
0 #NY-22
= (11) = = (1) (11) (13) (16)
= (22) = = (22) (1) (8) (12)
0 #NY-24
(8) (10) (18) (13) = (1) (14) (7)
(13) (15) (13) (8) = (11) (8) (9)
0 #NY-25
(3) (11) = = = = ? =
(15) (8) = = = = ? =
END BESTPAIRS2
```

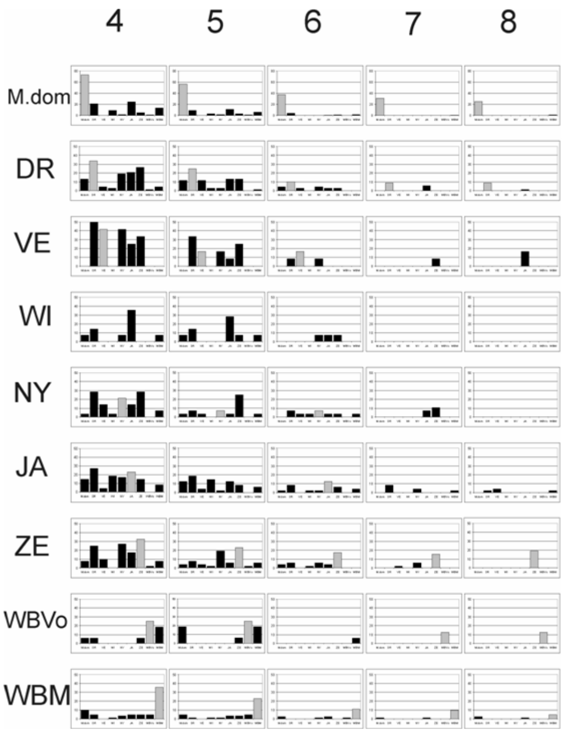
```
= 0.82 ? 0.96 1.00 1.00 0.93 1.00

0.77 0.72 0.68 0.79 0.77 = ? =

0.66 = 0.54 0.66 1.00 ? 1.00 1.00

0.97 0.97 ? = = ? ? =

= 0.60 = = 0.78 0.76 0.82 0.82

0.81 0.80 0.79 0.75 = 0.80 0.81 0.81

0.73 0.86 = = = = ? =

END PHASEPROBS
```

Example of introgression analysis


5-loci haplotypes of cultivated apple (GEN) found in wild apple



Example of gene flow analysis



Introgression: unlinked vs linked haplotypes

	Presence of *M. domestica* haplotypes inferred by PHASE		
Classification of *M. sylvestris* and putative hybrid plants based on STRUCTURE groups	**Total number of plants**	**Yes (hybrid)**	**No (*M. sylvestris*)**
Total number of plants	**161**	**24**	**137**
Hybrid	**28**	**7**	**21**
M. sylvestris	**133**	**17***	**116**

*** of these 17 plants, 3 were homozygous for the cultivar haplotype**

PLANT RESEARCH INTERNATIONAL
WAGENINGEN UR

Gene flow: unlinked Fst vs haplotype sharing

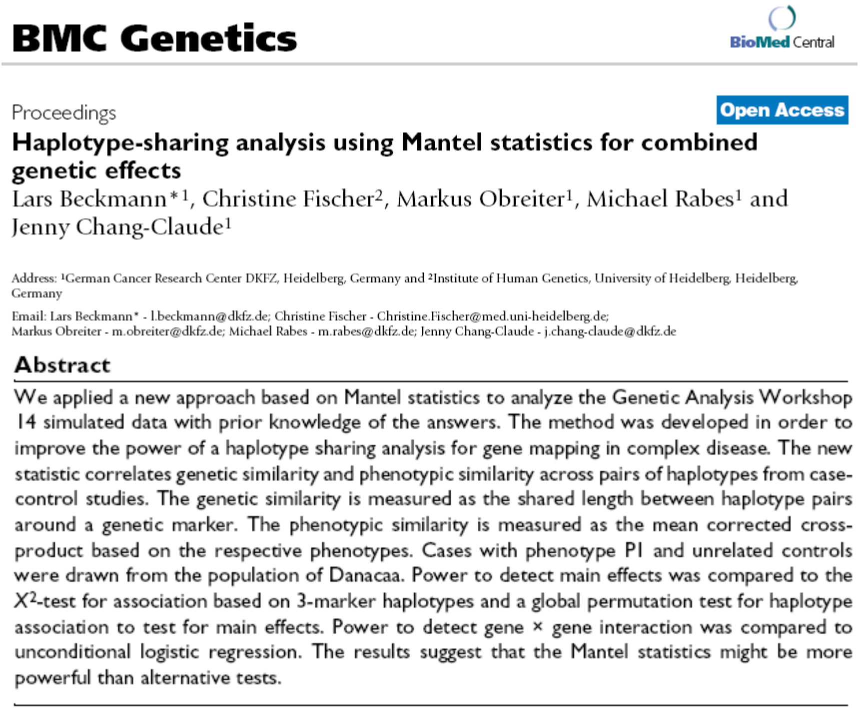
	DR	VE	WI	NY	JA	ZE	WBVo	WBM
DR	8	6	5	6	8	6	4	5
VE	0.016	6	no	6	8	7	no	no
WI	0.037	0.02	no	6	6	6	no	5
NY	0.023	0.033	0.013	6	7	7	no	6
JA	0.042*	0.038	0.015	-0.007	7	6	no	7
ZE	0.037*	0.039	0.018	-0.006	0.029	8	5	5
WBVo	0.056*	0.061	0.066	0.041	0.042	0.056*	5	6
WBM	0.047*	0.058*	0.043*	0.026	0.033*	0.032*	0.033	8

PLANT RESEARCH INTERNATIONAL
WAGENINGEN UR

- Introgression
  - One long haplotype shared > recent introgression
  - ~ 15% of wild apple plants have short haplotypes > remnants of old introgression
- Gene flow
  - Haplotype sharing largely consistent with Fst analysis (low genetic distance > more haplotypes shared)
  - A long haplotype shared from a small population is more proof of gene flow than high but not significant Fst
  - Within some small populations plants share longer haplotypes, in others they do not > different history?

Conclusions on the use of haplotype sharing

- Additional information obtained:
  - On the temporal scale of processes (recent vs old dispersal, recent vs old hybridisation events)
  - In populations with small sample size (no conclusive Fst) one long haplotype shared between two populations indicates gene flow in a qualitative way
  - On possible differences in population history of populations now comparable in size



BMC Genetics

Proceedings

Open Access

Haplotype-sharing analysis using Mantel statistics for combined genetic effects

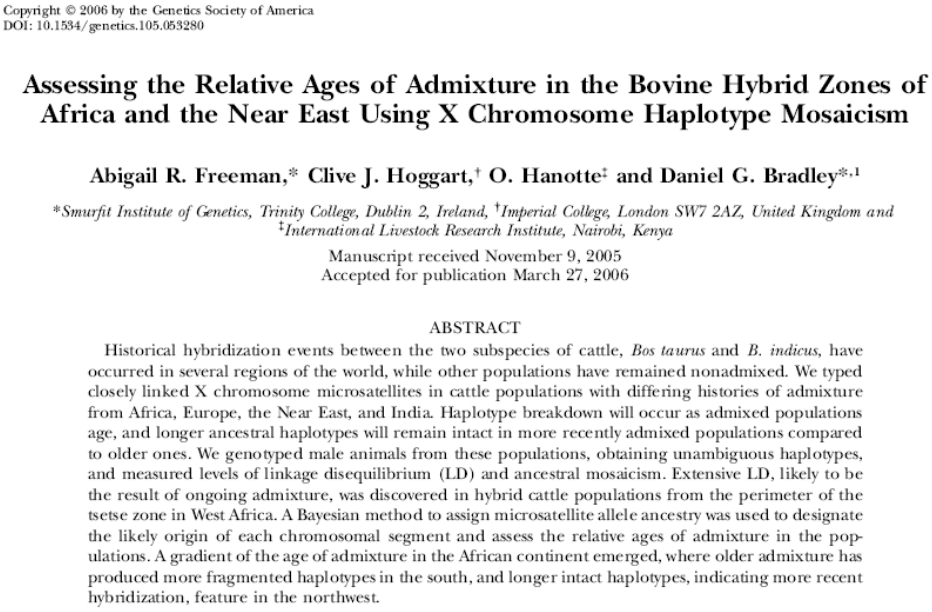
Lars Beckmann*[1], Christine Fischer[2], Markus Obreiter[1], Michael Rabes[1] and Jenny Chang-Claude[1]

Address: [1]German Cancer Research Center DKFZ, Heidelberg, Germany and [2]Institute of Human Genetics, University of Heidelberg, Heidelberg, Germany

Email: Lars Beckmann* - l.beckmann@dkfz.de; Christine Fischer - Christine.Fischer@med.uni-heidelberg.de;
Markus Obreiter - m.obreiter@dkfz.de; Michael Rabes - m.rabes@dkfz.de; Jenny Chang-Claude - j.chang-claude@dkfz.de

Abstract

We applied a new approach based on Mantel statistics to analyze the Genetic Analysis Workshop 14 simulated data with prior knowledge of the answers. The method was developed in order to improve the power of a haplotype sharing analysis for gene mapping in complex disease. The new statistic correlates genetic similarity and phenotypic similarity across pairs of haplotypes from case-control studies. The genetic similarity is measured as the shared length between haplotype pairs around a genetic marker. The phenotypic similarity is measured as the mean corrected cross-product based on the respective phenotypes. Cases with phenotype P1 and unrelated controls were drawn from the population of Danacaa. Power to detect main effects was compared to the X^2-test for association based on 3-marker haplotypes and a global permutation test for haplotype association to test for main effects. Power to detect gene × gene interaction was compared to unconditional logistic regression. The results suggest that the Mantel statistics might be more powerful than alternative tests.



DOI: 10.1534/genetics.105.053280

Assessing the Relative Ages of Admixture in the Bovine Hybrid Zones of Africa and the Near East Using X Chromosome Haplotype Mosaicism

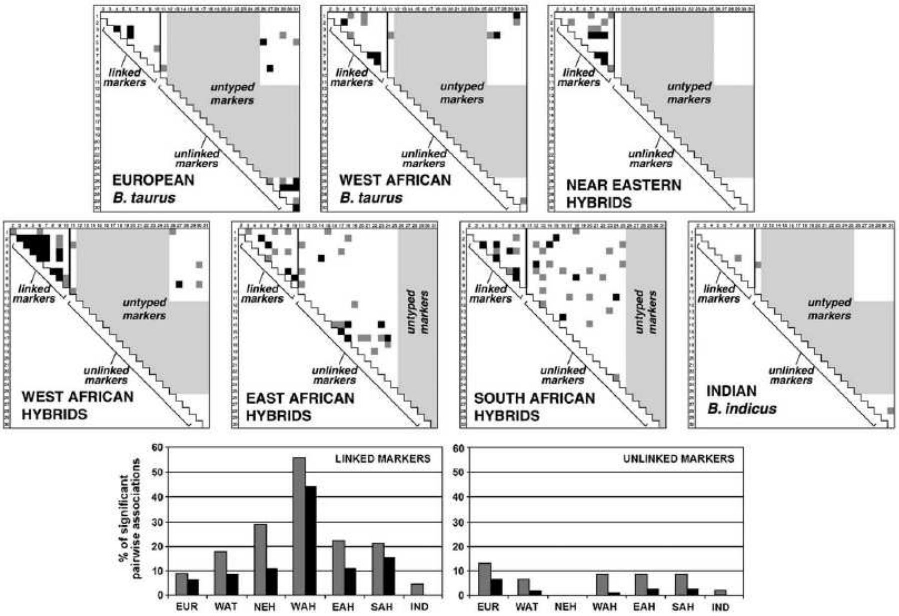
Abigail R. Freeman,* Clive J. Hoggart,† O. Hanotte‡ and Daniel G. Bradley*,[1]

**Smurfit Institute of Genetics, Trinity College, Dublin 2, Ireland, †Imperial College, London SW7 2AZ, United Kingdom and*
‡International Livestock Research Institute, Nairobi, Kenya

Manuscript received November 9, 2005
Accepted for publication March 27, 2006

ABSTRACT

Historical hybridization events between the two subspecies of cattle, *Bos taurus* and *B. indicus*, have occurred in several regions of the world, while other populations have remained nonadmixed. We typed closely linked X chromosome microsatellites in cattle populations with differing histories of admixture from Africa, Europe, the Near East, and India. Haplotype breakdown will occur as admixed populations age, and longer ancestral haplotypes will remain intact in more recently admixed populations compared to older ones. We genotyped male animals from these populations, obtaining unambiguous haplotypes, and measured levels of linkage disequilibrium (LD) and ancestral mosaicism. Extensive LD, likely to be the result of ongoing admixture, was discovered in hybrid cattle populations from the perimeter of the tsetse zone in West Africa. A Bayesian method to assign microsatellite allele ancestry was used to designate the likely origin of each chromosomal segment and assess the relative ages of admixture in the populations. A gradient of the age of admixture in the African continent emerged, where older admixture has produced more fragmented haplotypes in the south, and longer intact haplotypes, indicating more recent hybridization, feature in the northwest.



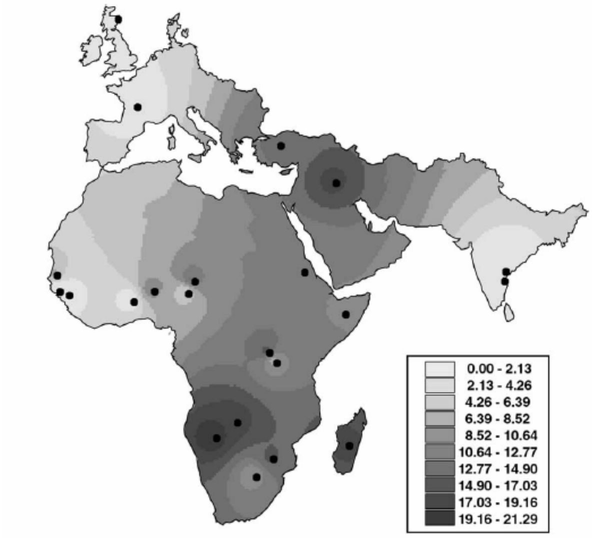


FIGURE 3.—Map showing values of τ. τ is the number of generations since a single admixture event. In these cattle populations admixture probably occurred in a continuous manner; therefore, τ is interpreted as a relative measure of the age of admixture among the populations studied (see MATERIALS AND METHODS). Higher values of τ have darker shading and suggest older admixture in those areas.

Thank you for your attention

PLANT RESEARCH INTERNATIONAL
WAGENINGEN UR

26