EAAP 2010
61st Annual Meeting of the European Association
for Animal Production
AUGUST 23rd-27th, 2010 - HERAKLION, CRETE ISLAND, GREECE

S38.6
jmj@inia.es

Multigenerational estimation of relatedness from molecular markers

Jesús Fernández

Department of Animal Breeding. Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria, Madrid, Spain

INIA
Instituto Nacional de Investigación
y Tecnología Agraria y Alimentaria

Knowledge of relationship between individuals

Parameters estimation

- *heritability*
- *genetic variance*
- *EBV*

Population management

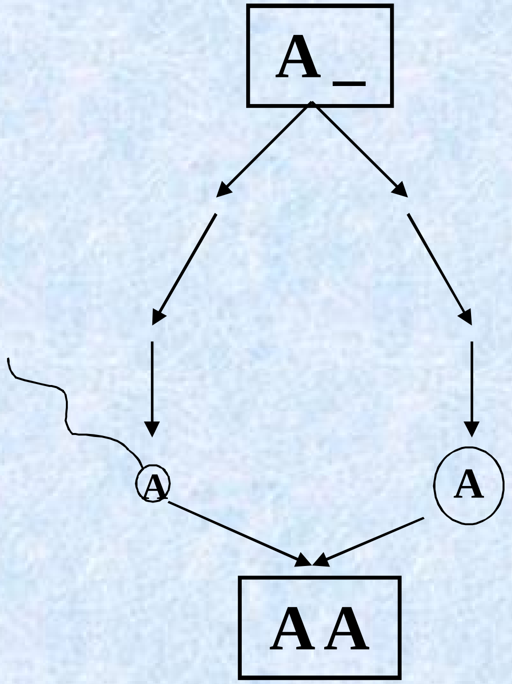
- *inbreeding avoidance*
- *maintenance of variability*

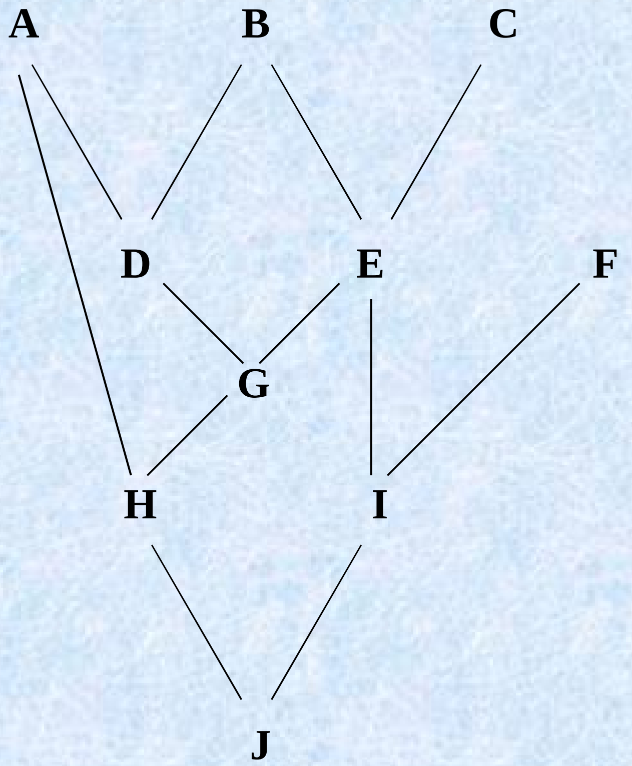
Ecological/behavioural studies

- *reproductive success*
- *dispersal/structure*
- *mating behaviour*

Identical By Descent

PEDIGREE





$$f_{HI} = \frac{1}{4}\left(f_{AE} + f_{AF} + f_{GE} + f_{GF}\right)$$

$$f_{CD} = \frac{1}{2}\left(f_{CA} + f_{CB}\right)$$

$$f_{AA} = \frac{1}{2}\left(1 + F_{A}\right)$$

$\left.\begin{array}{r}\textbf{Almost all natural}\\ \\ \textbf{Many captive}\end{array}\right\}$ **populations** ⇒ ***lack of genealogical records***

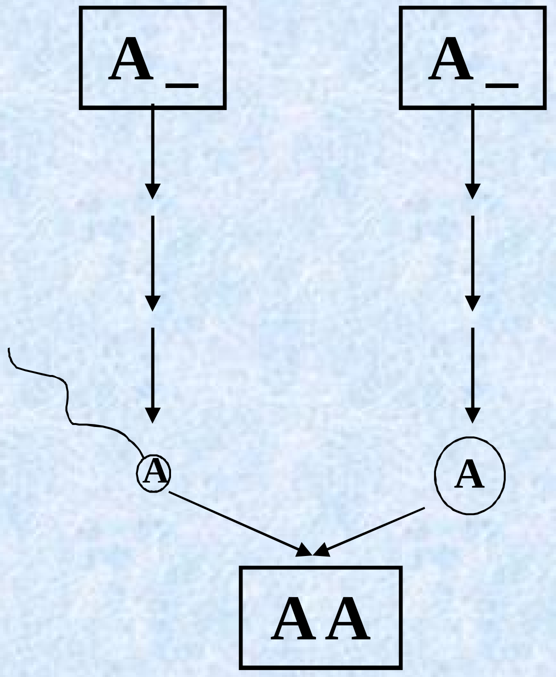
USE OF MOLECULAR MARKERS

(allozymes, DNA markers, etc)

Identical By State

f_{Mij} = molecular coancestry

> *"… probability that two alleles taken at random, one from each individual, are equal …"*



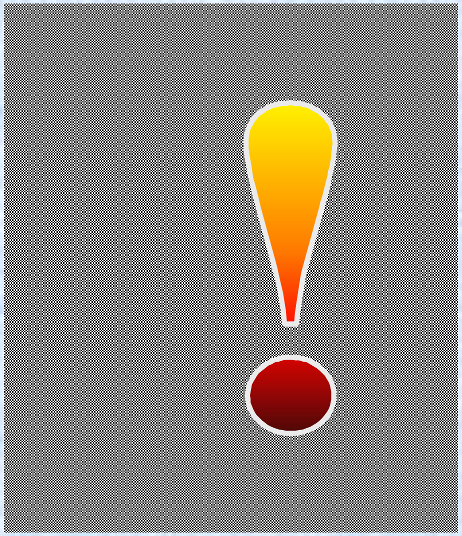
$A_1 A_3$	$A_1 A_3$	$A_2 A_3$	$A_2 A_4$
①	②	③	④

$$1 - 2 \Rightarrow f_M = 0.5$$

$$1 - 3 \Rightarrow f_M = 0.25$$

$$1 - 4 \Rightarrow f_M = 0$$

$$\text{equals} \Rightarrow \left\{ \begin{array}{c} \text{IBD} \\ \\ \text{IBS} \end{array} \right\}$$



- **leading to incompatible FS families**
- **depending on estimated frequencies**
- **familiar structure must be known *a priori***
- **limited number of types of relationship**

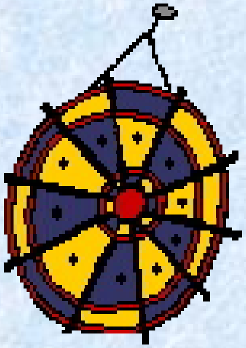
Fernández and Toro (2006)

➢ **correlation between molecular and estimated pedigree coancestry matrices**

- **Only contemporary individuals (assuming no ancestors)**

OBJECTIVE

- to extend the F&T method to allow for individuals from different generations



✓ **‘random’ generation of genealogies**

✓ **‘guided walk’ through solutions’ space**

⇒ ***correlation between molecular and estimated pedigree coancestry matrices***

✓ **markers compatibility**

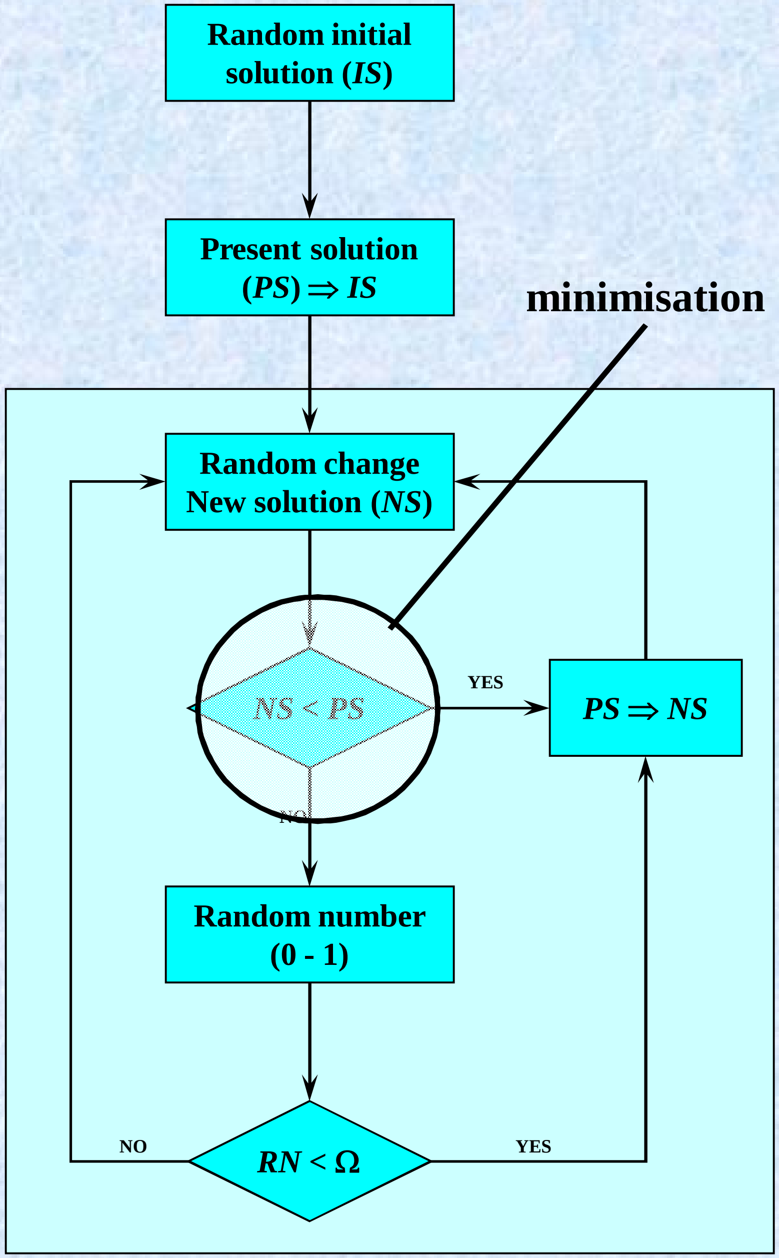
⇒ ***parent - offspring***

⇒ ***FS families***

✓ **age compatibility (if this information available)**

⇒ ***parents older than offspring***

⇒ ***age of sexual maturity***



Random initial solution (*IS*)

Present solution (*PS*) ⇒ *IS*

minimisation

Random change New solution (*NS*)

NS < *PS*

YES

NO

PS* ⇒ *NS

Random number (0 - 1)

NO

$RN < \Omega$

YES

$$\Omega = e^{-\frac{NS-PS}{T}}$$

- several rounds
- a little ‘colder’

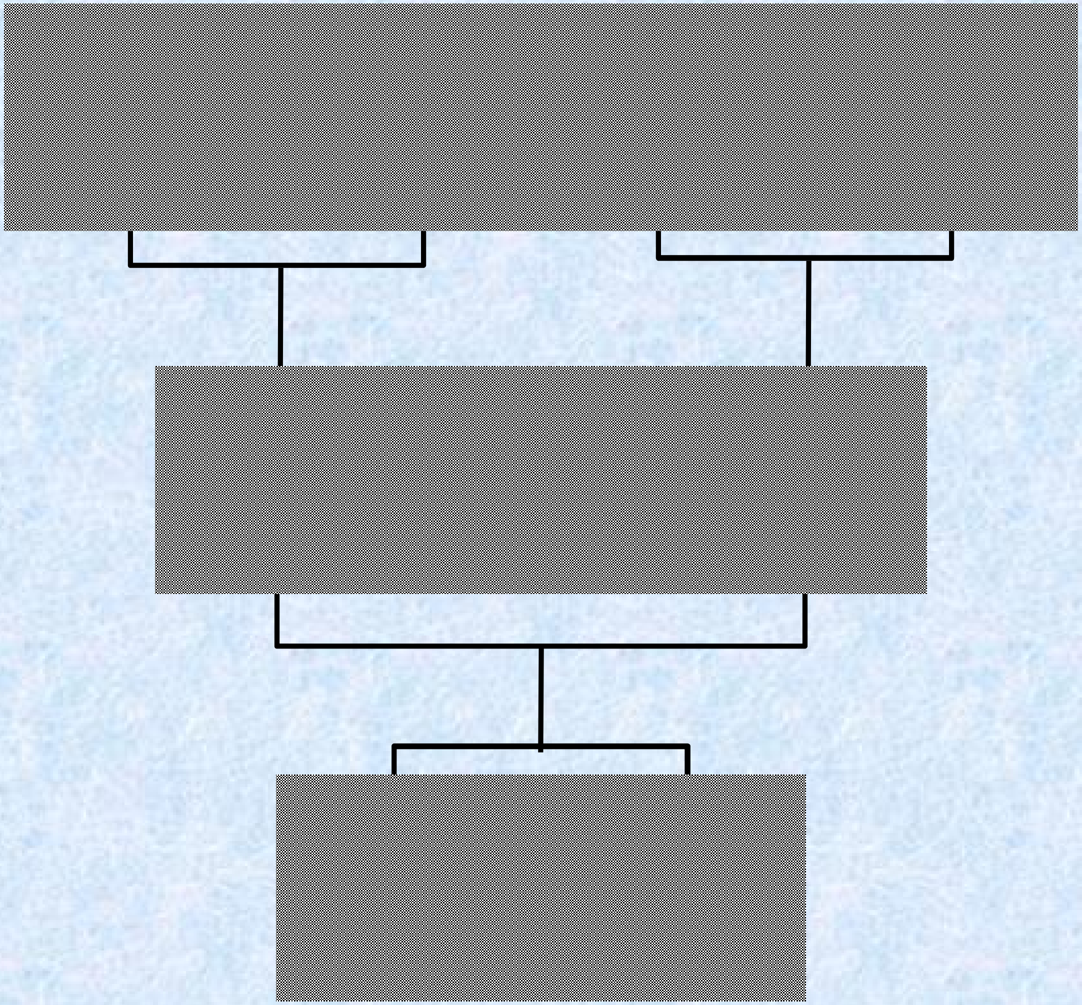
$T\downarrow \ \Omega\downarrow$

- go on with the cycle
- not accepting further changes

MATERIAL AND METHODS

- **Computer simulations**
  - ✓ **Two different pedigrees (small or large)**
  - ✓ **5, 10, 15, 20, 25 or 30 independent codominant markers**
  - ✓ **2, 4, 6, 8 or 10 alleles per marker**
  - ✓ **Averaged over 20 replicates**

Case 1



Age

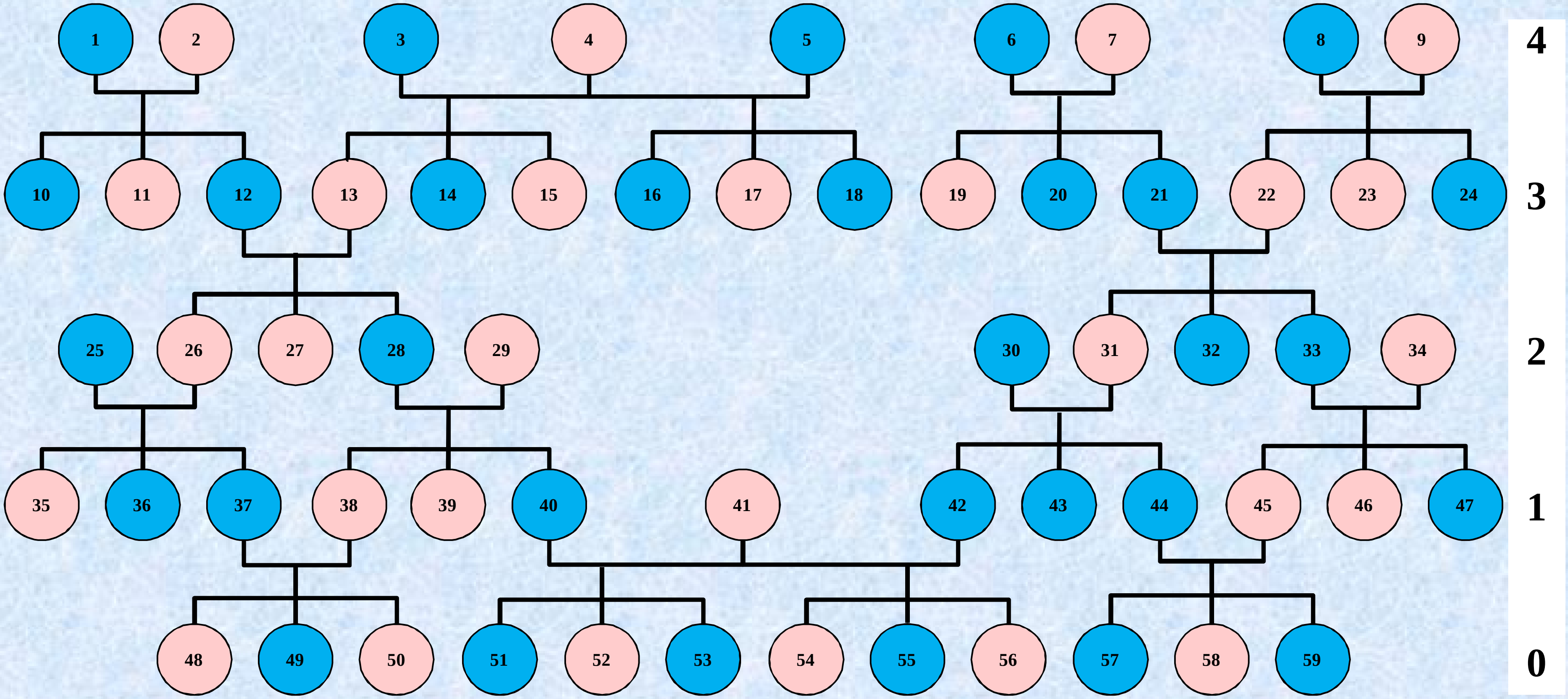
2

1

0

✓ **3 generations**

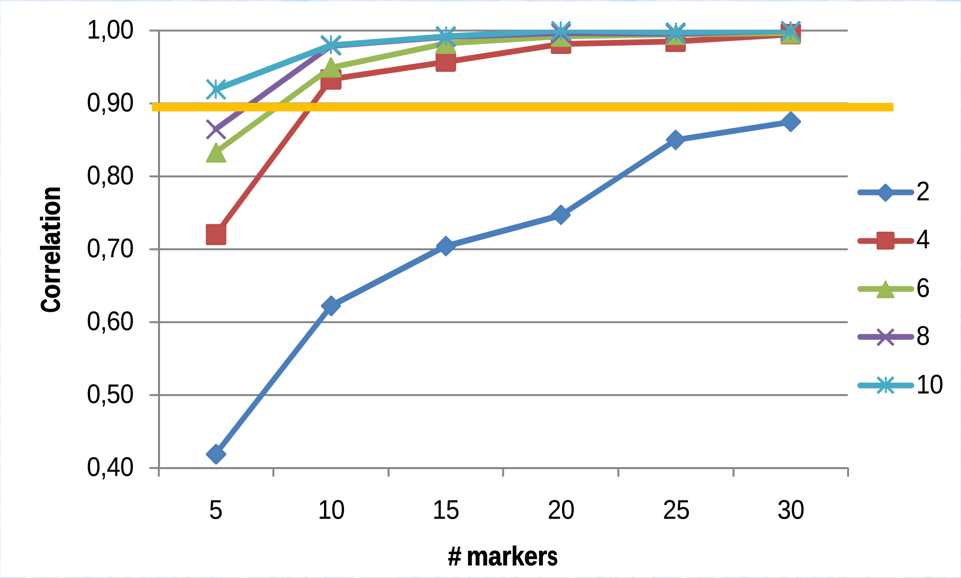
Case 2



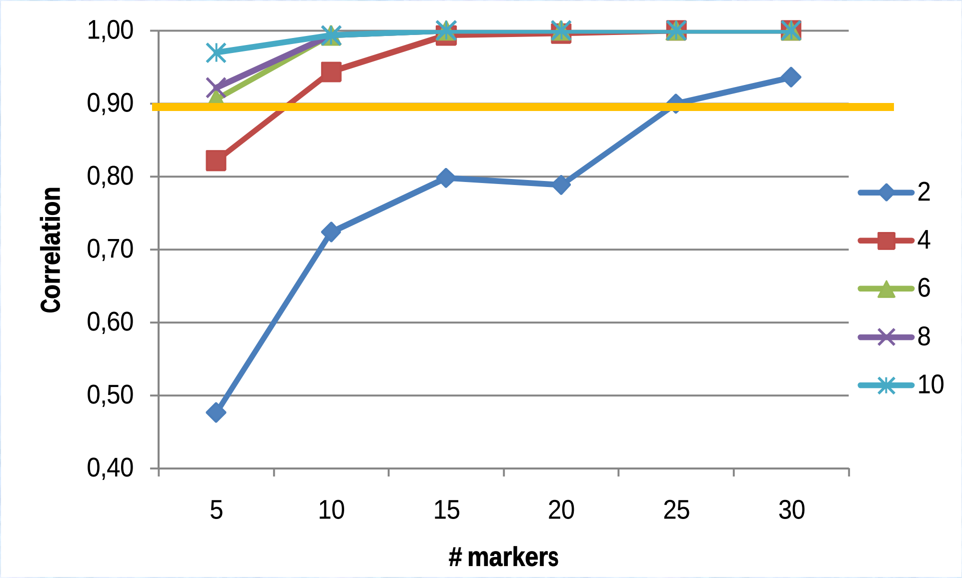
- ✓ 5 generations

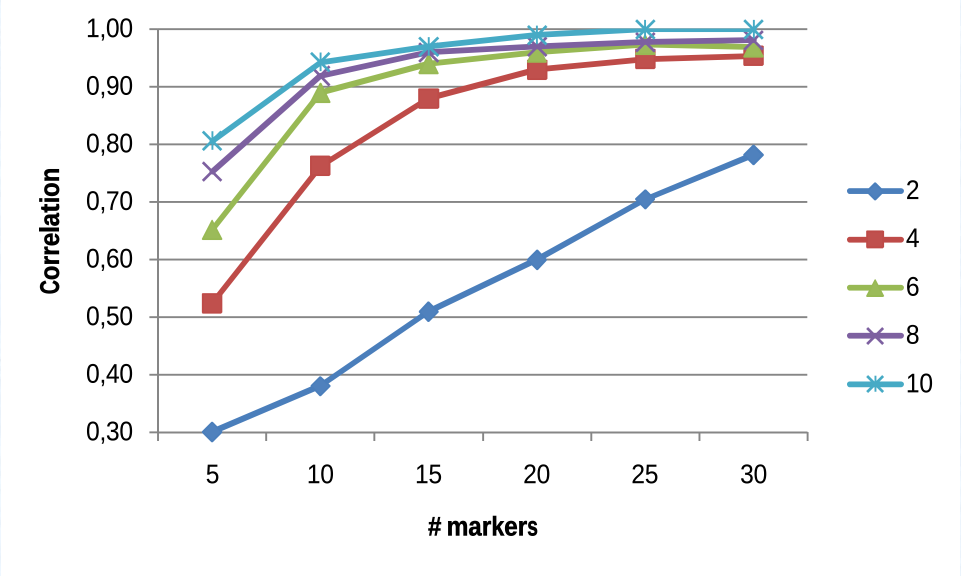
- ➢ **Measures of fit**
  - ✓ **Correlation between real and estimated coancestry matrices**
  - ✓ **Number of errors in triplets**
  - ✓ **Percentage of completely correct pedigrees**

RESULTS

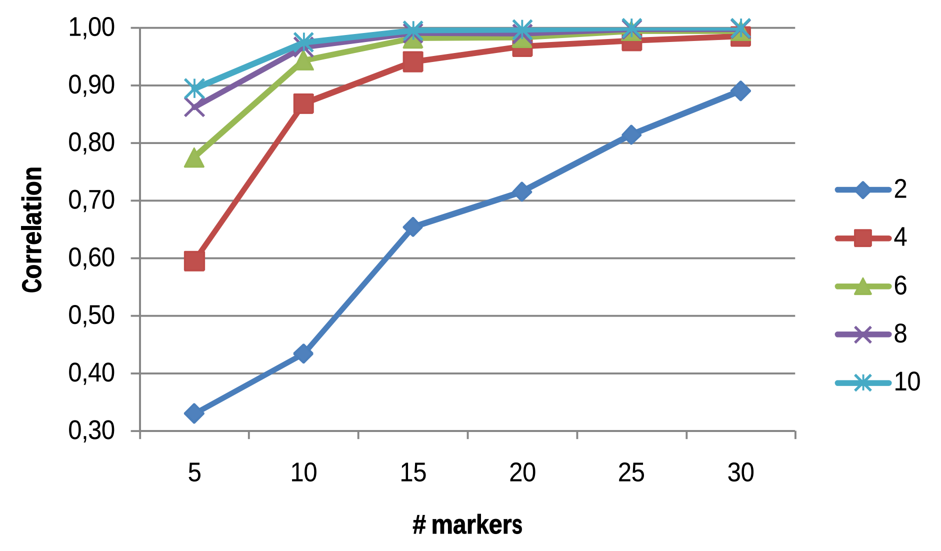


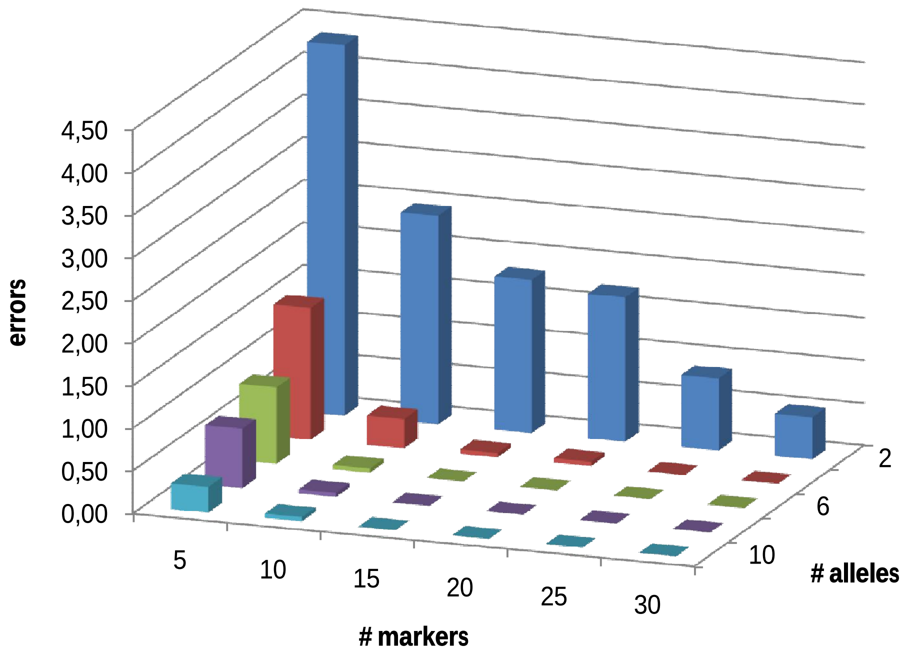
Case 1

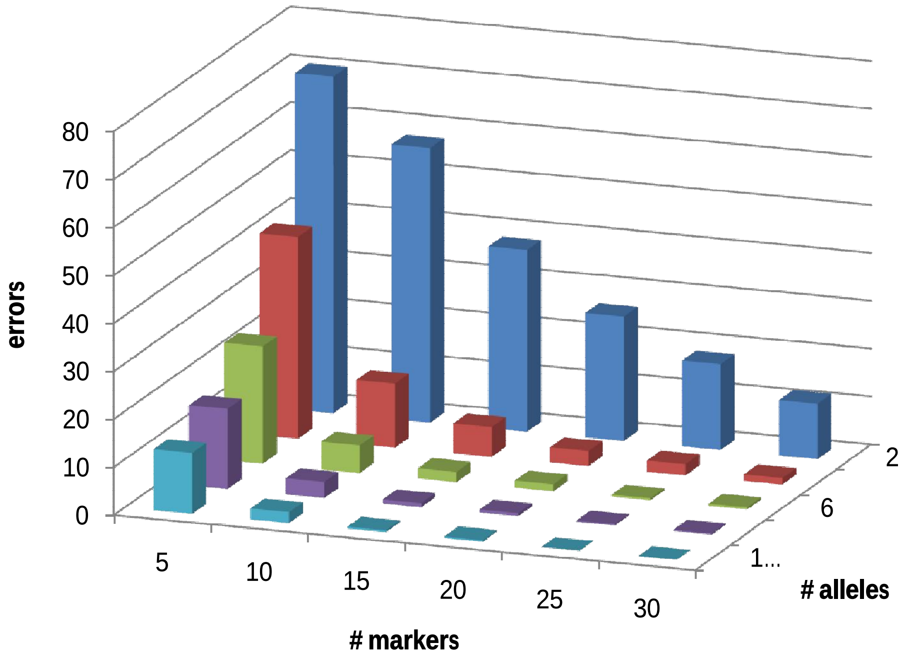


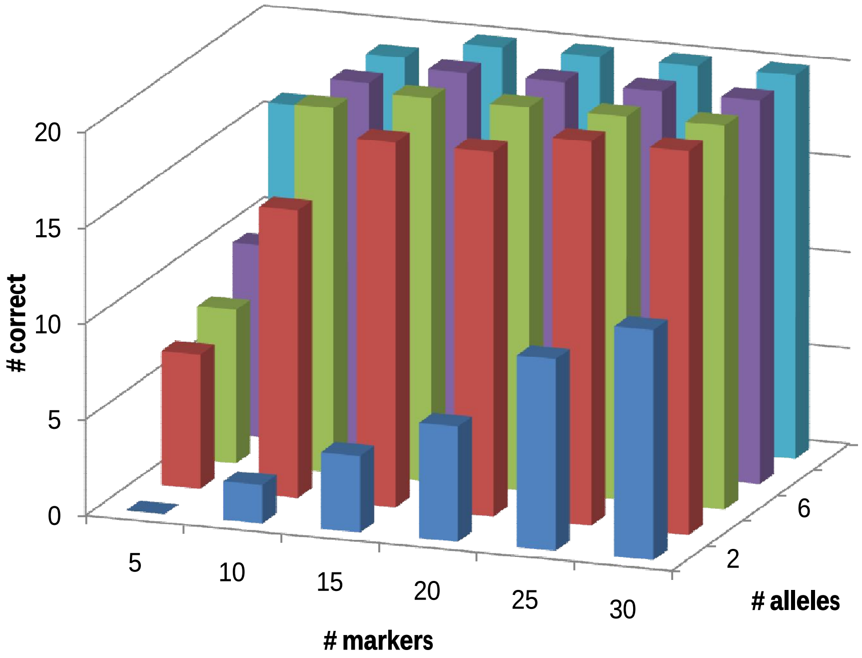

1,00
0,90
0,80
0,70
0,60
0,50
0,40
0,30
Correlation
5
10
15
20
25
30
markers
2
4
6
8
10

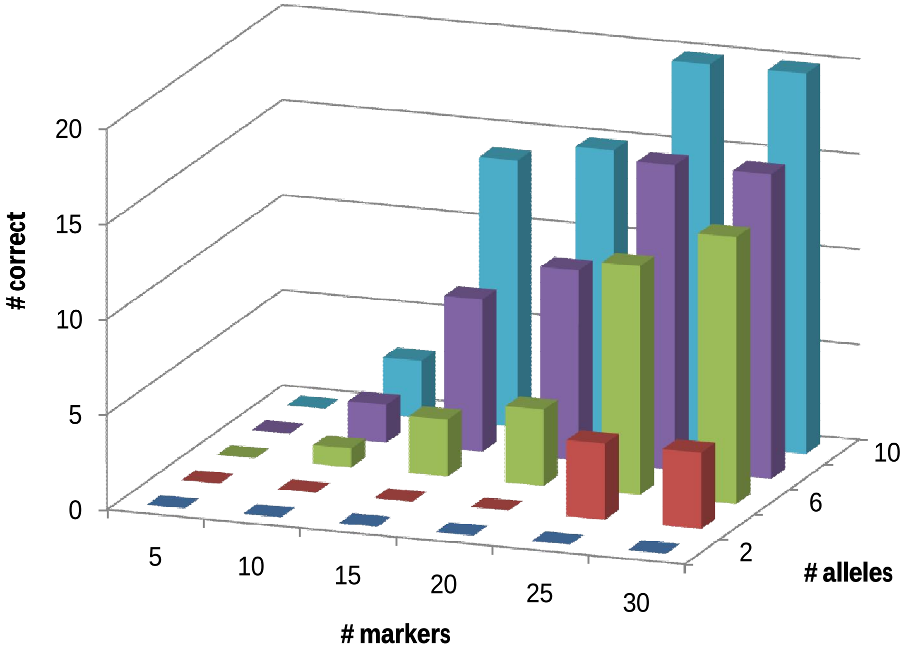
Case 2


1,00
0,90
0,80
0,70
0,60
0,50
0,40
0,30
Correlation
5
10
15
20
25
30
markers
2
4
6
8
10


errors
4,50
4,00
3,50
3,00
2,50
2,00
1,50
1,00
0,50
0,00
5
10
15
20
25
30
markers
2
6
10
alleles


errors
80
70
60
50
40
30
20
10
0
5
10
15
20
25
30
markers
2
6
1...
alleles


correct
20
15
10
5
0
5
10
15
20
25
30
markers
10
6
2
alleles


correct
20
15
10
5
0
5
10
15
20
25
30
markers
10
6
2
alleles

TO TAKE HOME

(or sit-in)

- ✓ **Powerful with a reasonable amount of molecular information**
- ✓ **Improves when ages known**
- ✓ **Can cope with complex situations**

Under development

- **Including ‘virtual’ individuals**
  - ✓ **Complete gaps**
- **Include known relationships**
- **Test under HWD and LD**

THANK YOU!