

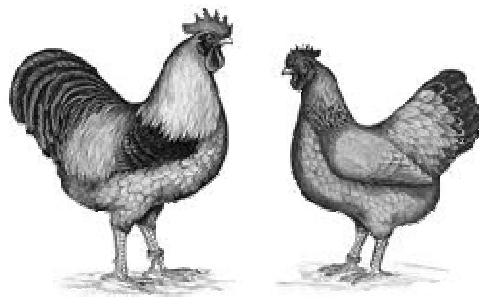


Marker-based estimation of effective population size from one-generation population samples

H. Simianer¹, M. Hansen¹ and S. Weigend²

¹ Institute of Animal Breeding and Genetics,
Georg-August-University Göttingen, Germany

² Institute of Farm Animal Genetics
Friedrich-Loeffler-Institute, Mariensee, Germany

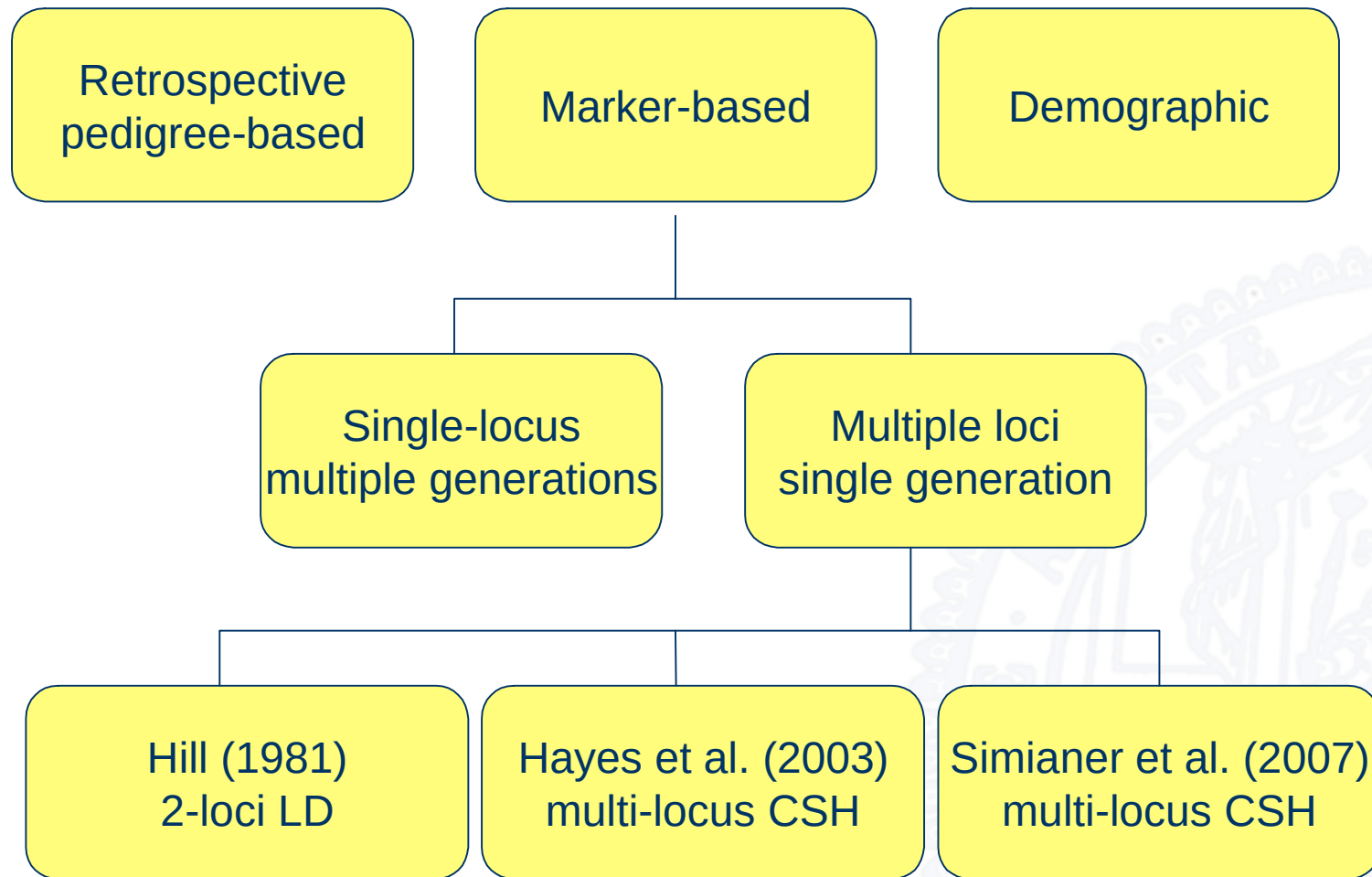




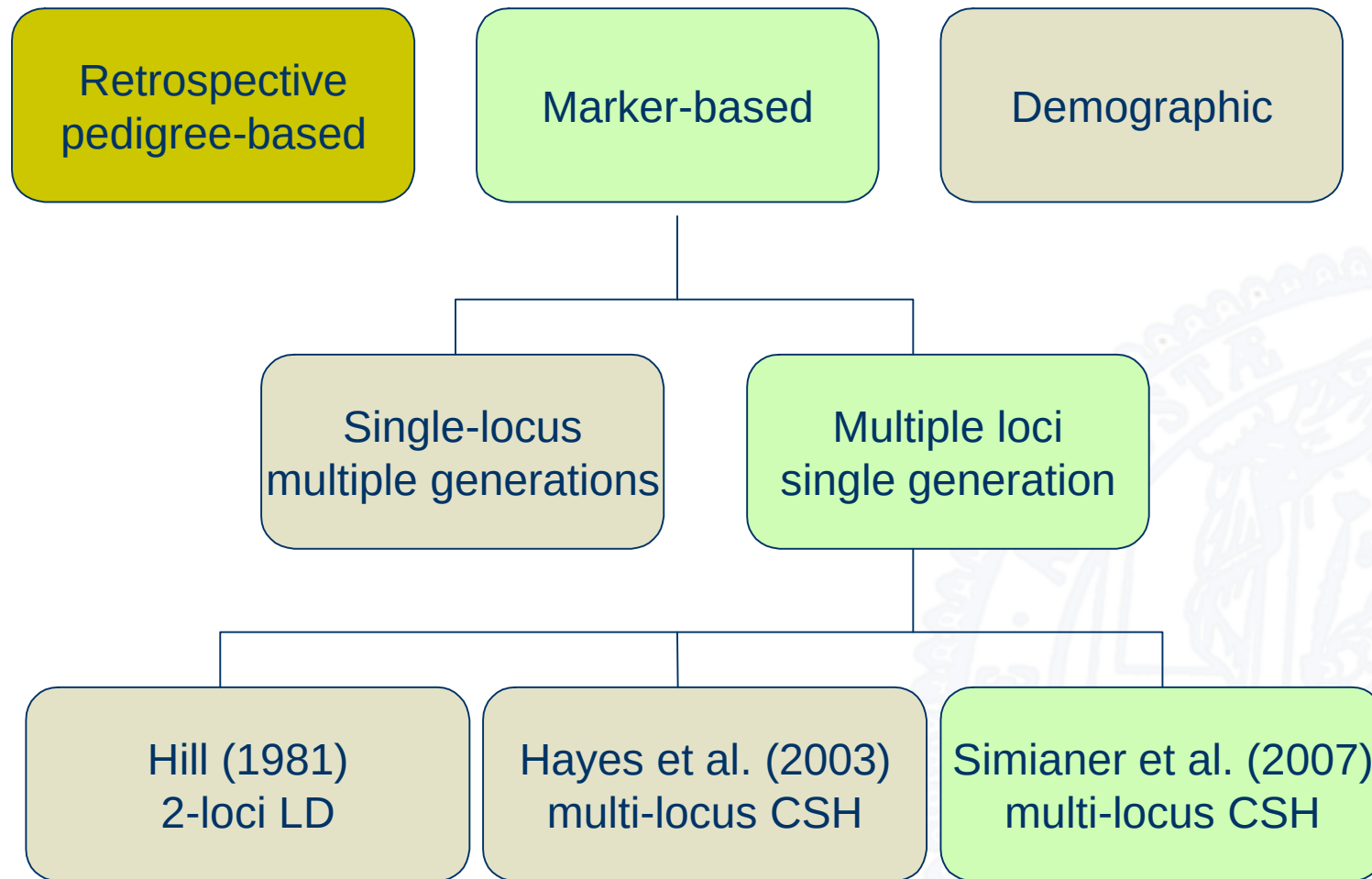
The effective population size N_e of a population P is defined as the size of a **hypothetical ideal population** which has the same increase in inbreeding as the **existing true population P**

N_e is a parameter of central importance in many areas of population genetics and breeding, but is **'notoriously difficult to estimate'** (Frankham, 1995)

Approaches to estimate N_e

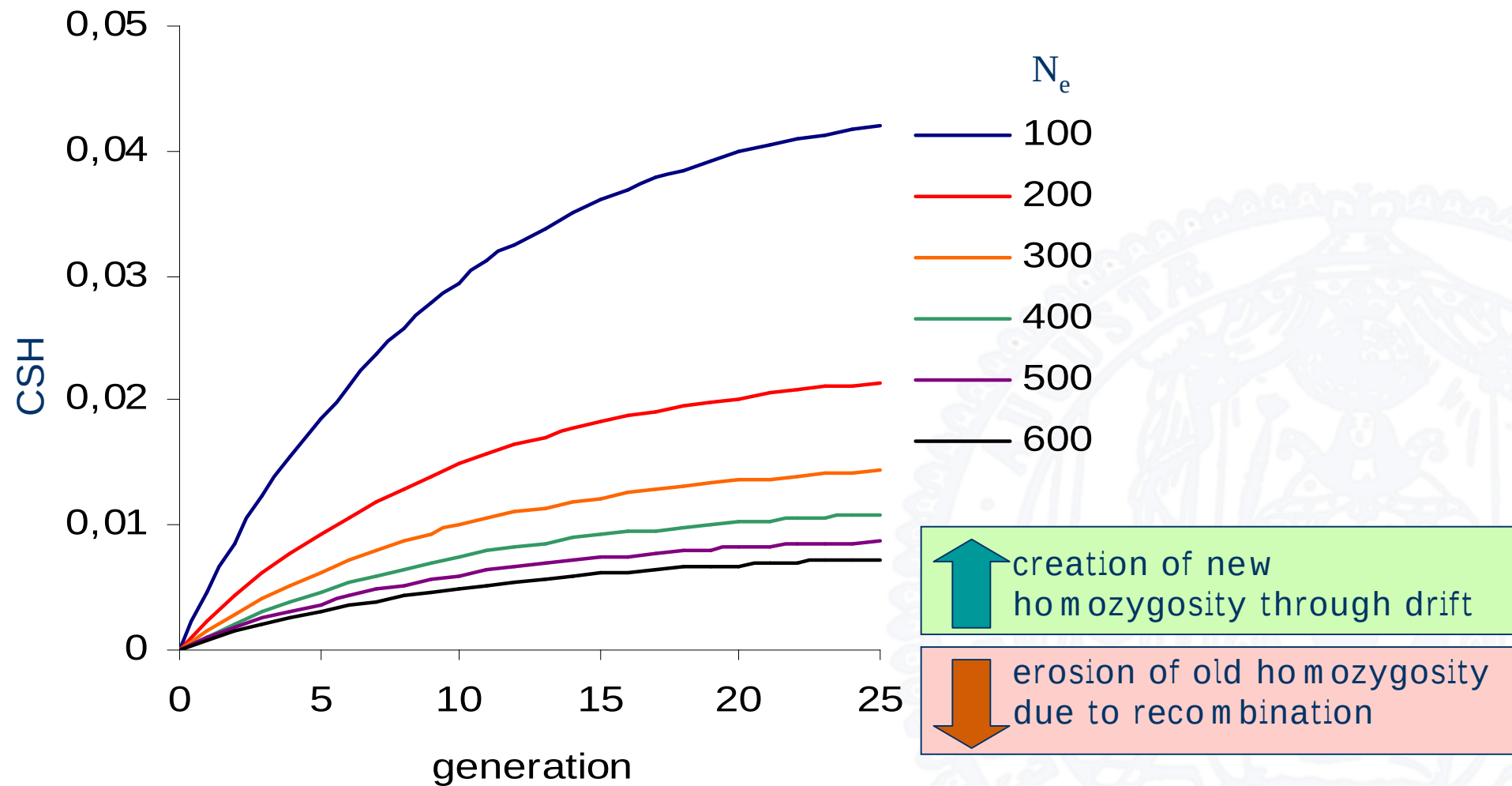


Approaches to estimate N_e





Expected CSH as function of effective population size and number of generations with a segment length $x = 0,05 M$ (Flury et al., 2006)





The equilibrium value of CSH in a closed population is a function of segment length x (in Morgan) and effective population size

$$CSH_{\infty} = \frac{e^{-2x}}{e^{-2x} + 2N_e(1 - e^{-2x})}$$



With known segment length we can estimate the effective population size from the observed CSH, assuming the equilibrium state is reached

$$N_e = \frac{e^{-2x}(1 - CSH_{\infty})}{2(1 - e^{-2x})CSH_{\infty}}$$

Chicken populations studied



Lohmann Tierzucht GmbH

2 commercial lines

LSL-A ○ , LB-A ●



Institute of Farm Animal Genetics

3 experimental lines

L68 ● , R11 & R22 ○



- per line 10 roosters and 30 mother-daughter pairs
- complete pedigrees over 11 generations
- DNA samples from generation 2006/07



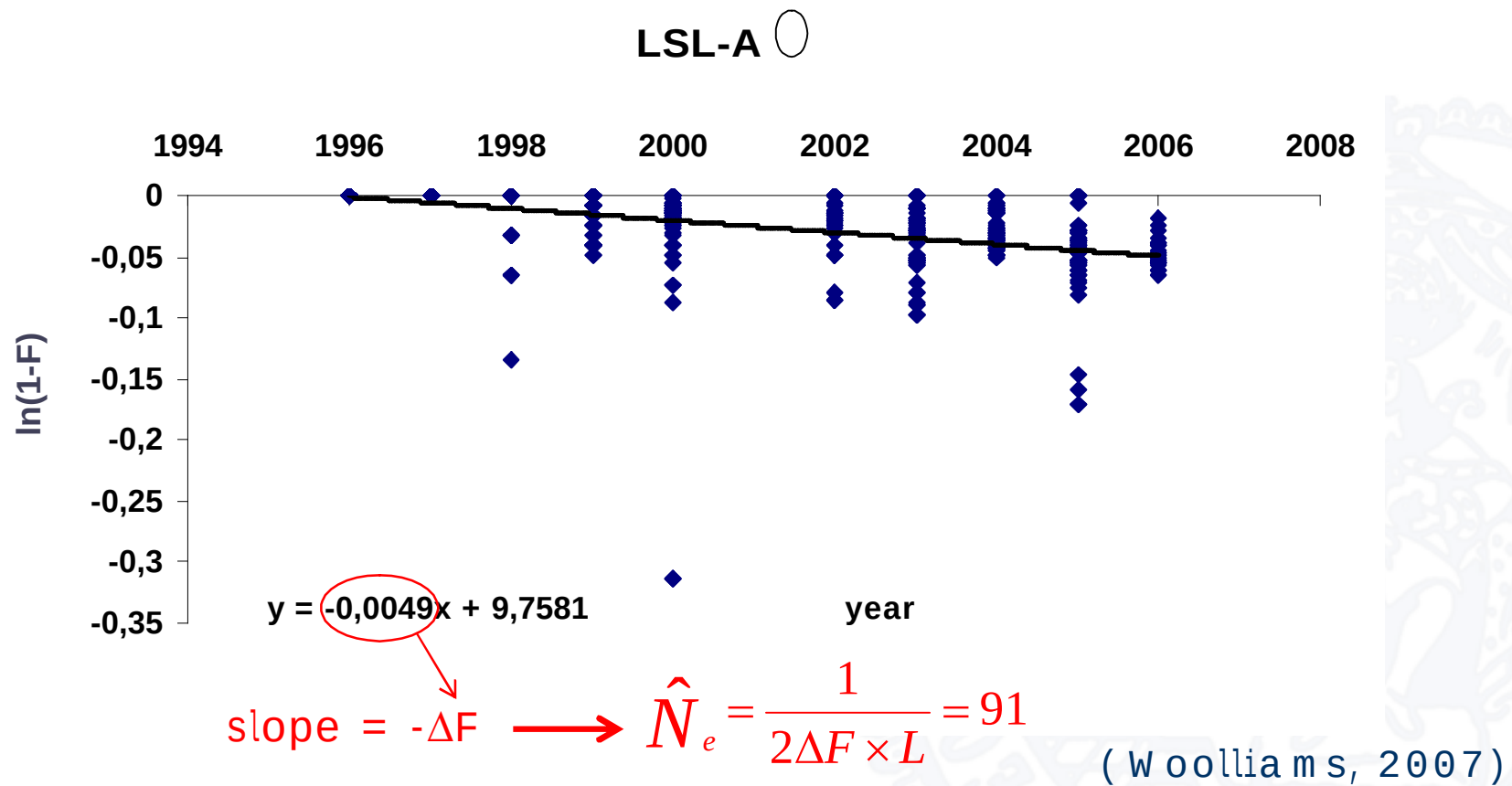


Methods

- **509 animals** from all lines inclusive pedigree information (11 generations)
- **Pedigree-based estimation** of $\Delta F \rightarrow N_e$ with the method of Woolliams (2007)
- **Genotyping** of 6 segments with 4-6 (S30) microsatellites (15'270 single data points)
- **Haplotype reconstruction** with PedPhase (Li & Jiang, 2003)
- Calculation of **observed CSH**, correction for CSH expected under Hardy-Weinberg and linkage equilibrium
- **Marker-based estimation** of N_e from corrected CSH value with the method of Simianer et al. (2007)

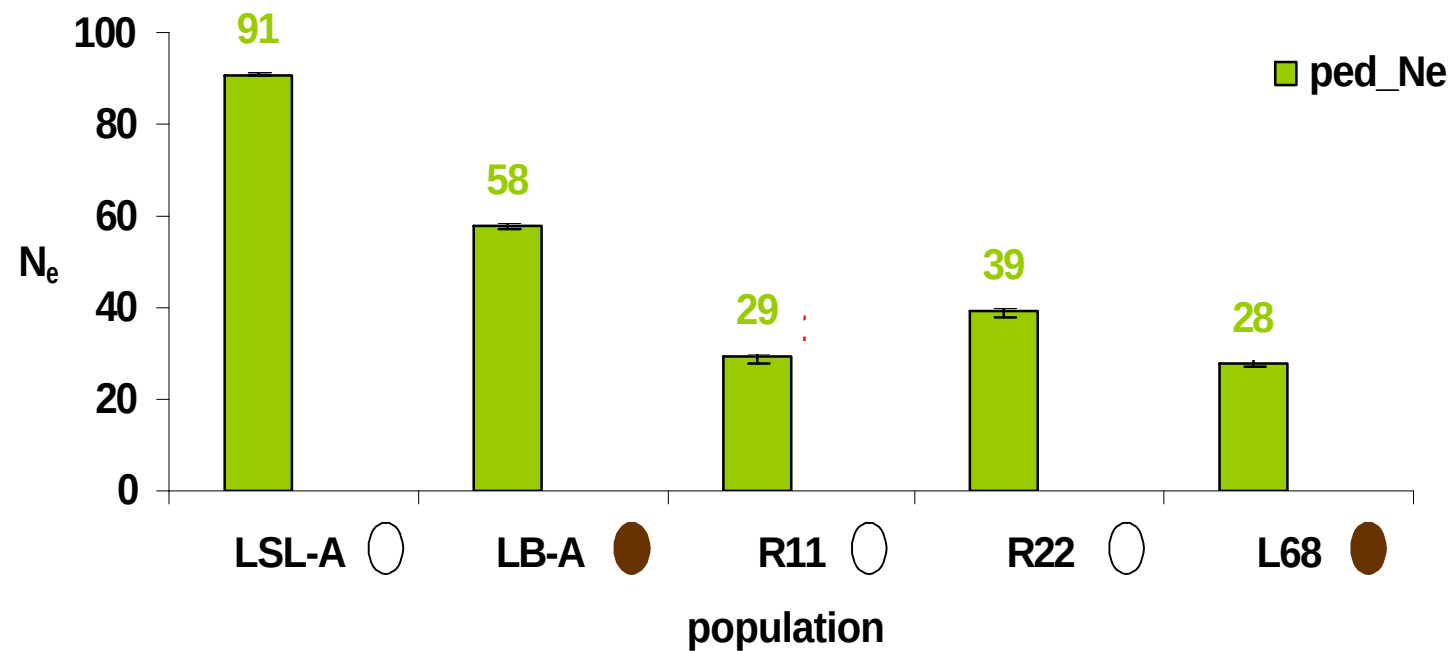


Pedigree-based N_e for chicken lines



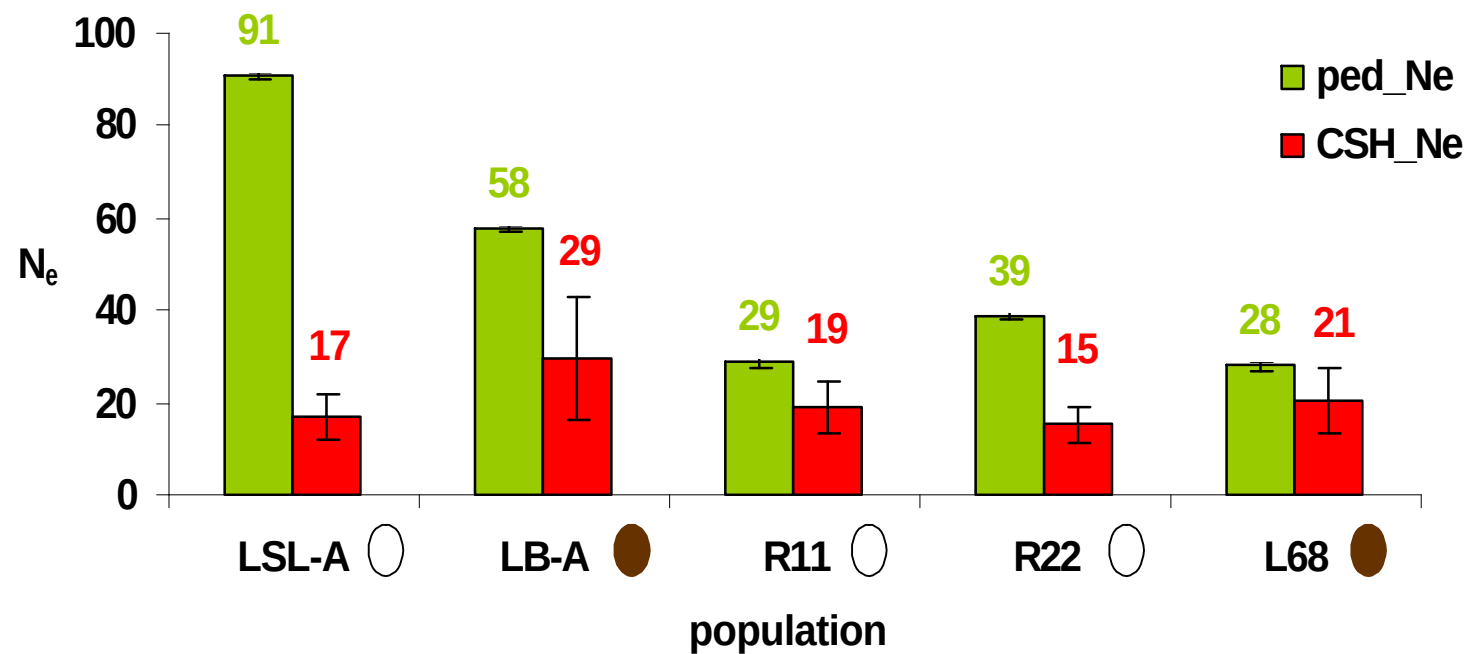


Comparison of pedigree- and marker-based estimators of N_e



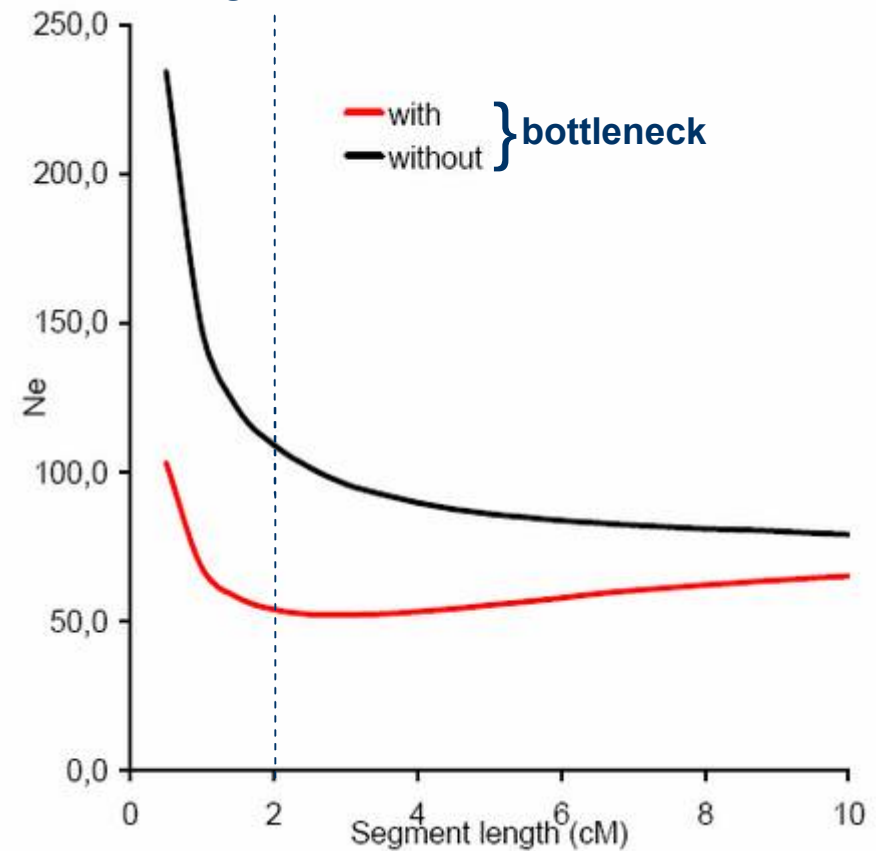
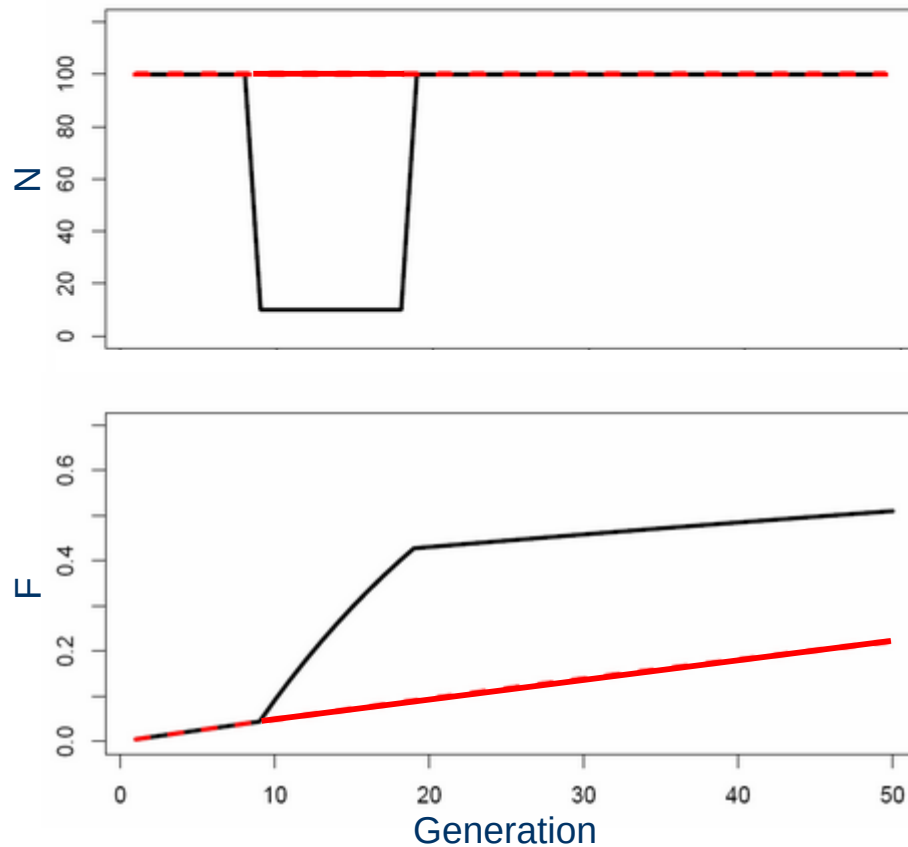


Comparison of pedigree- and marker-based estimators of N_e



Simulation:

- 50 generations random mating, ♂:♀ = 1:1
- reference scenario: 100 animals per generation
- **bottleneck scenario**: 10 animals per generation in generation 10 - 19



Pedigree-based estimates of N_e : ~ 100
(with 10 or 50 generations, w/w'out bottleneck)

Summary



- Pedigree-based and marker-based approaches yield **different results**
- Marker-based estimates of N_e are **consistently smaller** than pedigree-based estimates, especially so in (commercial) white layer lines
- Obviously much **more homozygosity** occurred on the molecular level than expected from the available (11 generation) pedigree
- Marker-based methods may be affected by **historic events** (e.g. population bottlenecks) and appear to produce systematically different results in defined scenarios



How much genetic information is needed to detect differences in N_e

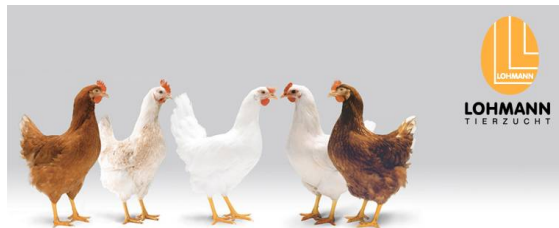
Reference population with $N_e = 30$; 90 % power on the $\alpha = 0.05$ level

Difference in $SE(CSH)$	Difference in CSH	lower critical N_e	upper critical N_e	Required information units
$\pm 0,25$	± 0.00645	26.9	33.9	336
$\pm 0,50$	± 0.01290	24.3	38.8	84
$\pm 0,75$	± 0.01935	22.2	45.3	37
$\pm 1,00$	± 0.02580	20.4	54.3	21

1 information unit: 6 segments with on average 5 microsatellites

1 microsatellite $\sim$ 3 SNPs $\rightarrow$ 84 information units = **7560 SNPs needed**

High density SNP genotyping will soon provide a sufficient basis for marker-based estimation of N_e



We would like to thank the
German Research Foundation
(DFG) for funding the project
and
Lohmann Tierzucht GmbH for
providing samples
and pedigree data

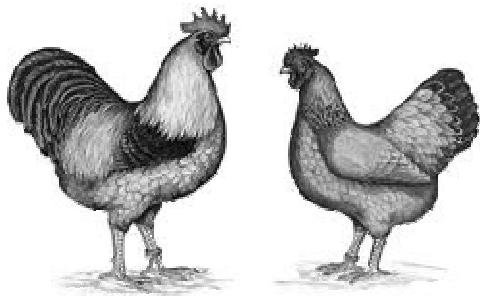


Marker-based estimation of effective population size from one-generation population samples

H. Simianer¹, M. Hansen¹ and S. Weigend²

¹ Institute of Animal Breeding and Genetics,
Georg-August-University Göttingen, Germany

² Institute of Farm Animal Genetics
Friedrich-Loeffler-Institute, Mariensee, Germany

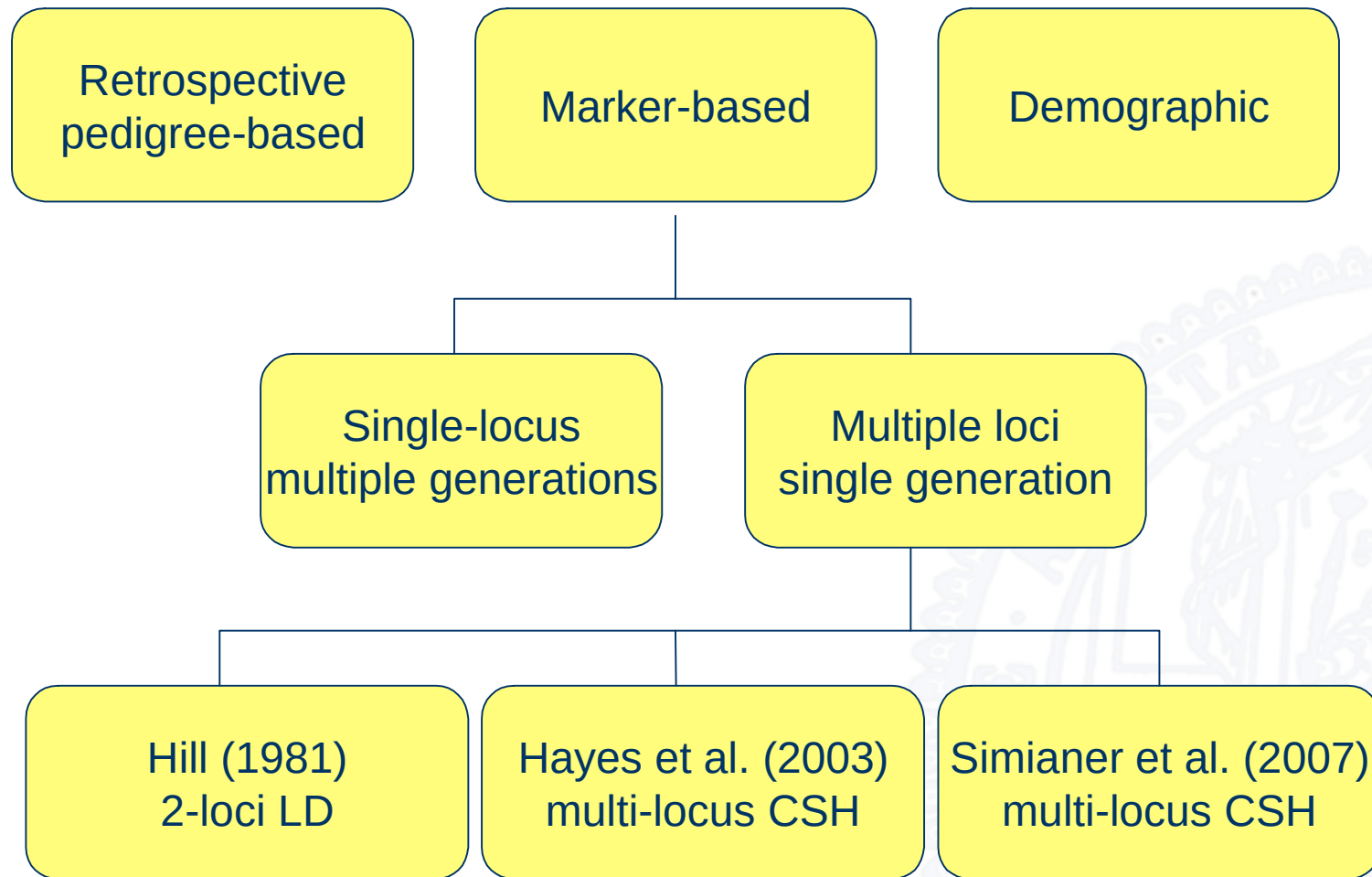




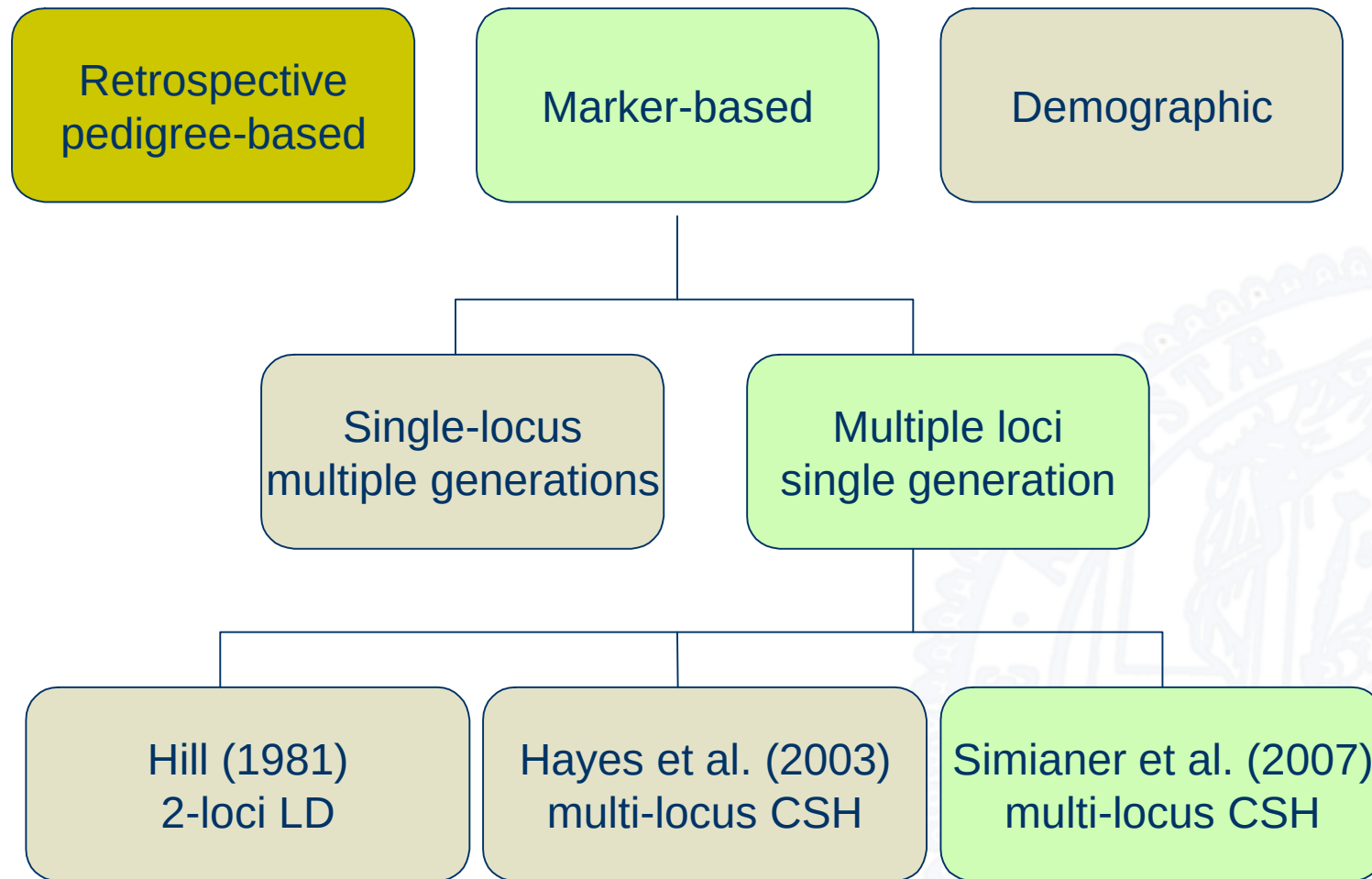
The effective population size N_e of a population P is defined as the size of a **hypothetical ideal population** which has the same increase in inbreeding as the **existing true population P**

N_e is a parameter of central importance in many areas of population genetics and breeding, but is **'notoriously difficult to estimate'** (Frankham, 1995)

Approaches to estimate N_e

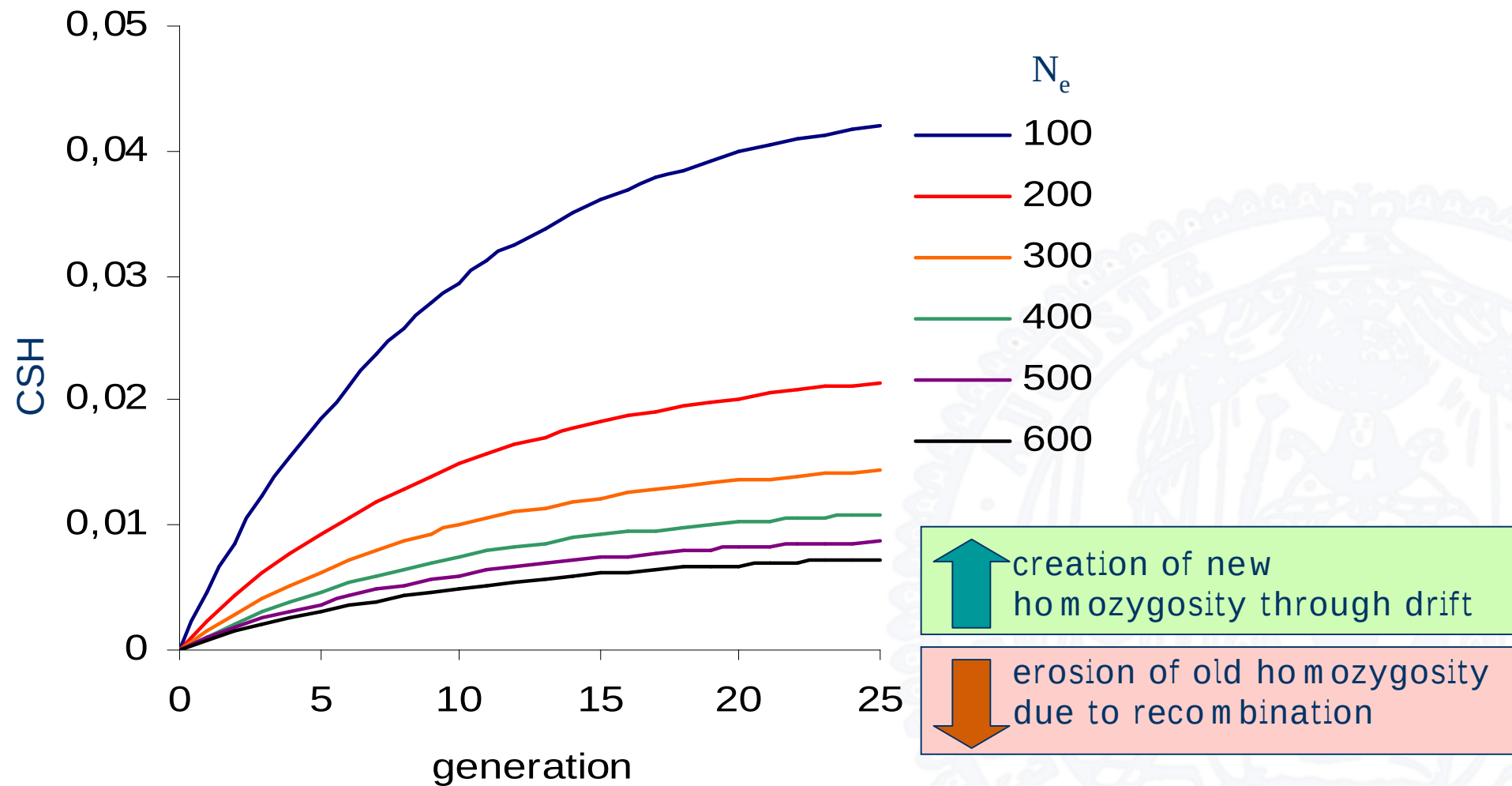


Approaches to estimate N_e





Expected CSH as function of effective population size and number of generations with a segment length $x = 0,05 M$ (Flury et al., 2006)





The equilibrium value of CSH in a closed population is a function of segment length x (in Morgan) and effective population size

$$CSH_{\infty} = \frac{e^{-2x}}{e^{-2x} + 2N_e(1 - e^{-2x})}$$



With known segment length we can estimate the effective population size from the observed CSH, assuming the equilibrium state is reached

$$N_e = \frac{e^{-2x}(1 - CSH_{\infty})}{2(1 - e^{-2x})CSH_{\infty}}$$

Chicken populations studied



Lohmann Tierzucht GmbH

2 commercial lines

LSL-A ○ , LB-A ●



Institute of Farm Animal Genetics

3 experimental lines

L68 ● , R11 & R22 ○



- per line 10 roosters and 30 mother-daughter pairs
- complete pedigrees over 11 generations
- DNA samples from generation 2006/07

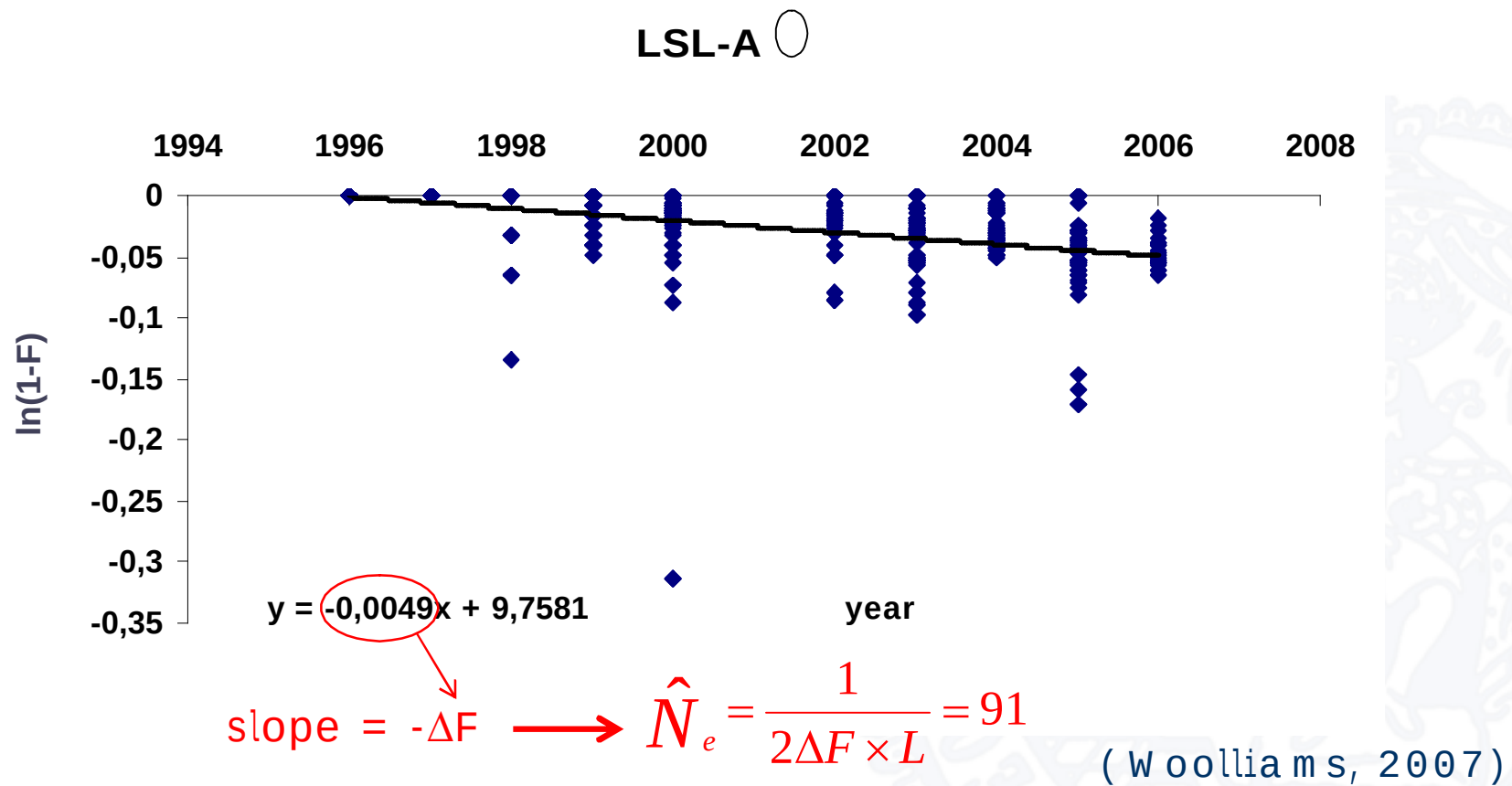


Methods

- **509 animals** from all lines inclusive pedigree information (11 generations)
- **Pedigree-based estimation** of $\Delta F \rightarrow N_e$ with the method of Woolliams (2007)
- **Genotyping** of 6 segments with 4-6 (S30) microsatellites (15'270 single data points)
- **Haplotype reconstruction** with PedPhase (Li & Jiang, 2003)
- Calculation of **observed CSH**, correction for CSH expected under Hardy-Weinberg and linkage equilibrium
- **Marker-based estimation** of N_e from corrected CSH value with the method of Simianer et al. (2007)

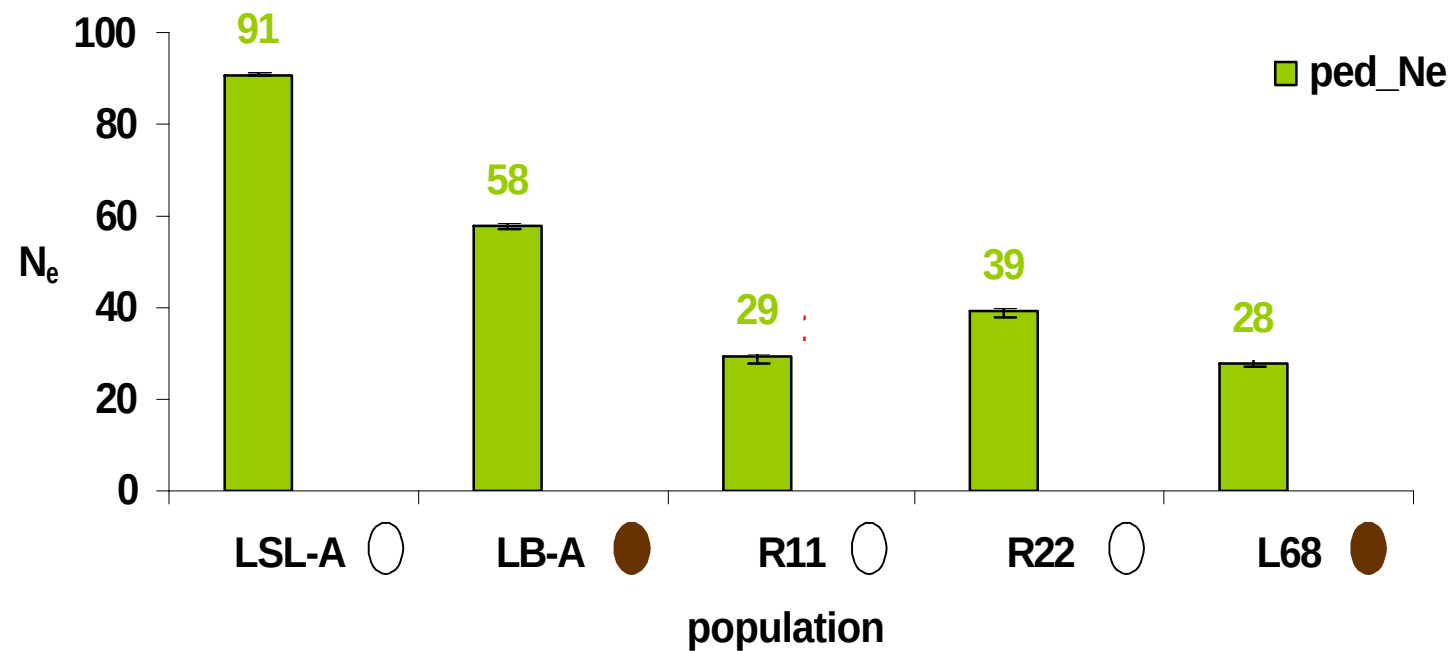


Pedigree-based N_e for chicken lines



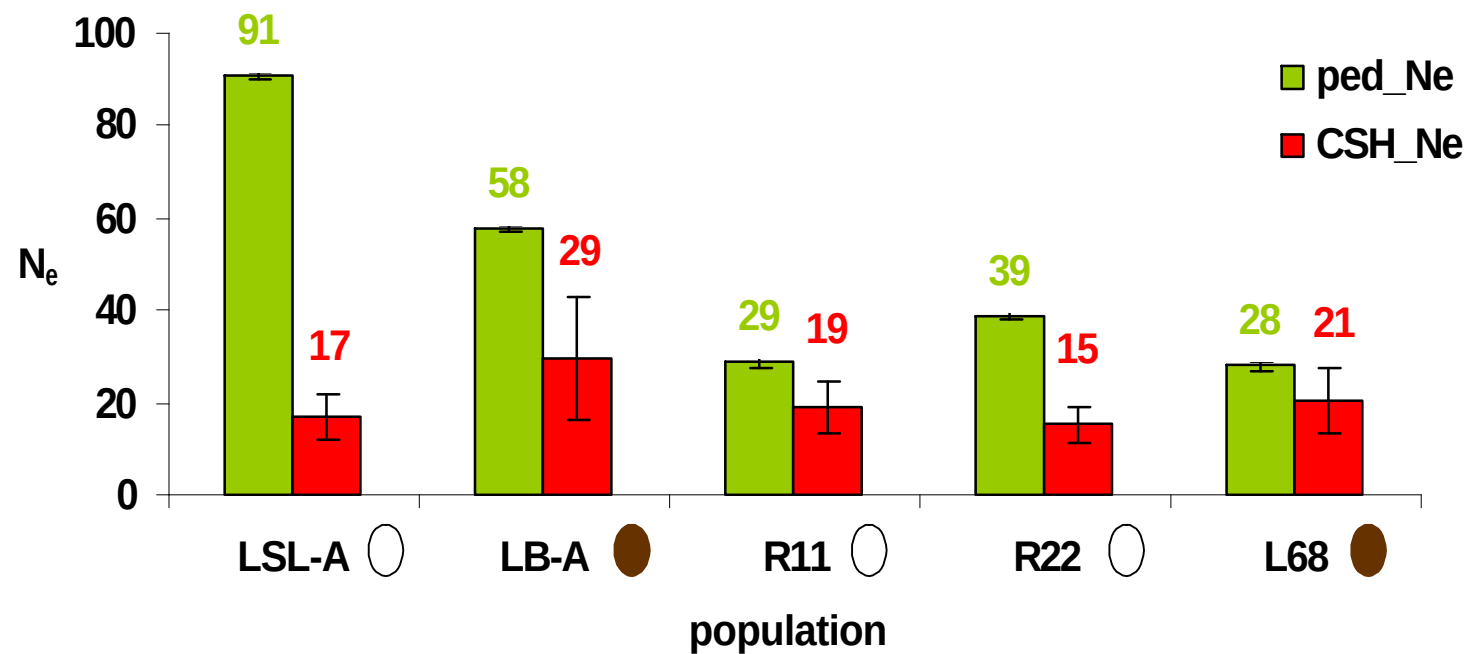


Comparison of pedigree- and marker-based estimators of N_e



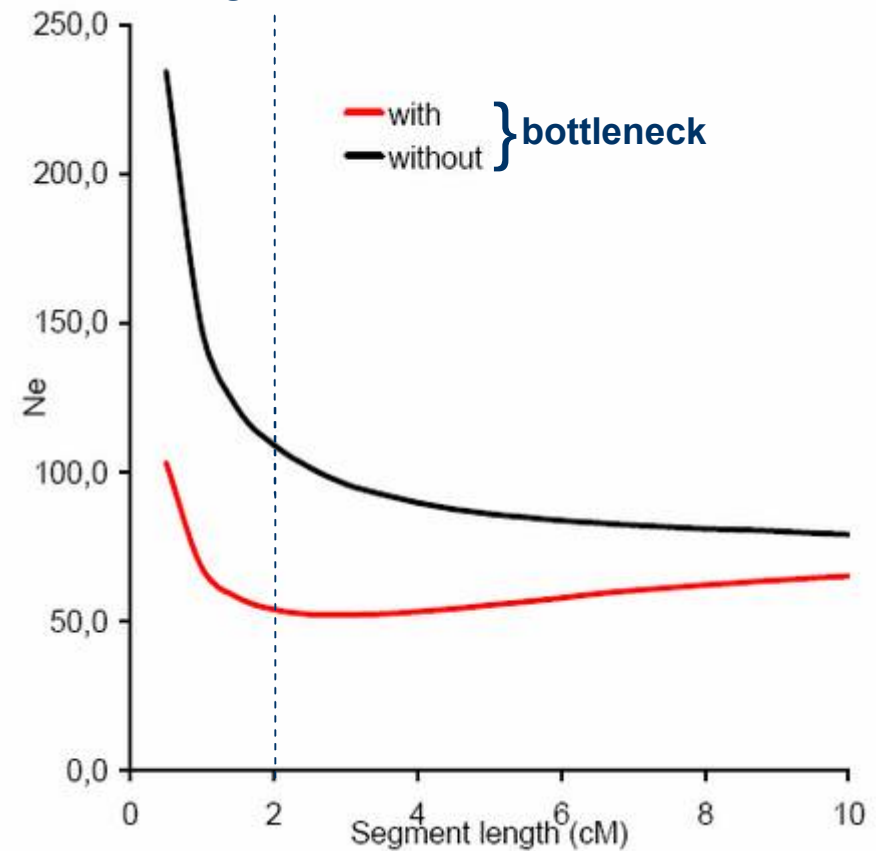
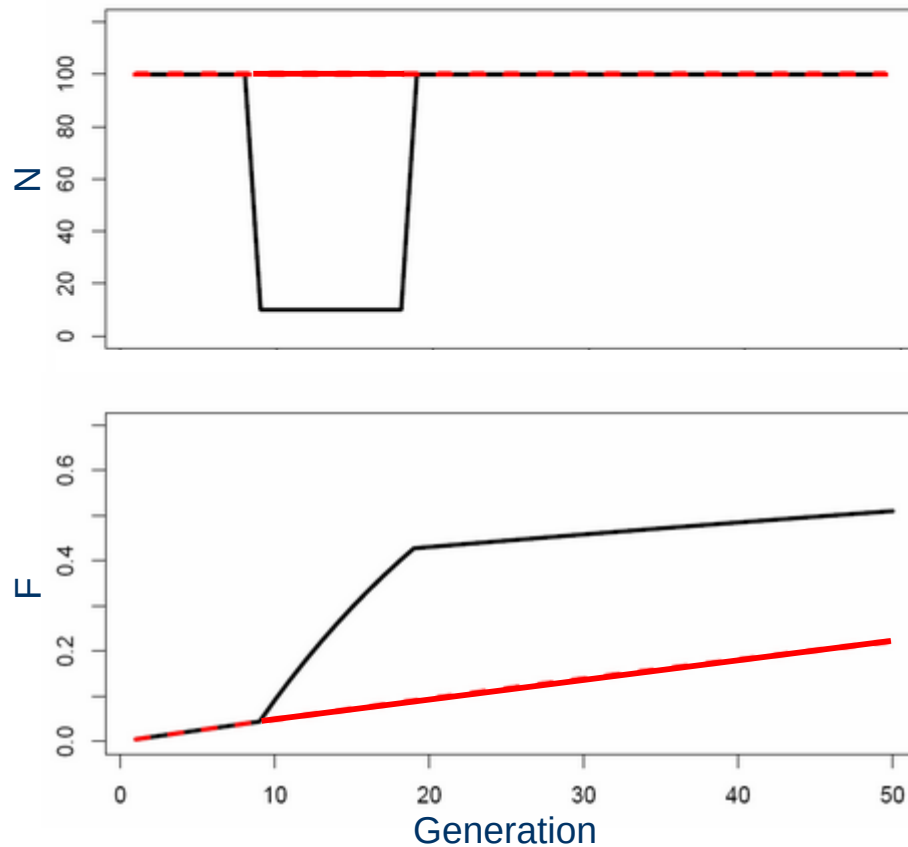


Comparison of pedigree- and marker-based estimators of N_e



Simulation:

- 50 generations random mating, ♂:♀ = 1:1
- reference scenario: 100 animals per generation
- **bottleneck scenario**: 10 animals per generation in generation 10 - 19



Pedigree-based estimates of N_e : ~ 100
(with 10 or 50 generations, w/w'out bottleneck)

Summary



- Pedigree-based and marker-based approaches yield **different results**
- Marker-based estimates of N_e are **consistently smaller** than pedigree-based estimates, especially so in (commercial) white layer lines
- Obviously much **more homozygosity** occurred on the molecular level than expected from the available (11 generation) pedigree
- Marker-based methods may be affected by **historic events** (e.g. population bottlenecks) and appear to produce systematically different results in defined scenarios



How much genetic information is needed to detect differences in N_e

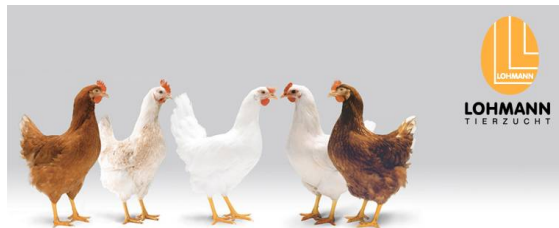
Reference population with $N_e = 30$; 90 % power on the $\alpha = 0.05$ level

Difference in $SE(CSH)$	Difference in CSH	lower critical N_e	upper critical N_e	Required information units
$\pm 0,25$	± 0.00645	26.9	33.9	336
$\pm 0,50$	± 0.01290	24.3	38.8	84
$\pm 0,75$	± 0.01935	22.2	45.3	37
$\pm 1,00$	± 0.02580	20.4	54.3	21

1 information unit: 6 segments with on average 5 microsatellites

1 microsatellite $\sim$ 3 SNPs $\rightarrow$ 84 information units = **7560 SNPs needed**

High density SNP genotyping will soon provide a sufficient basis for marker-based estimation of N_e



We would like to thank the
German Research Foundation
(DFG) for funding the project
and
Lohmann Tierzucht GmbH for
providing samples
and pedigree data