



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

Principi base della selezione genomica in acquacoltura

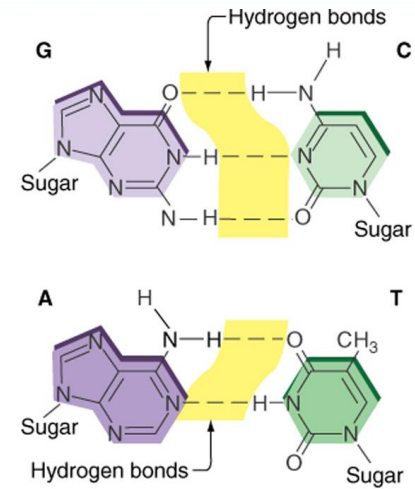
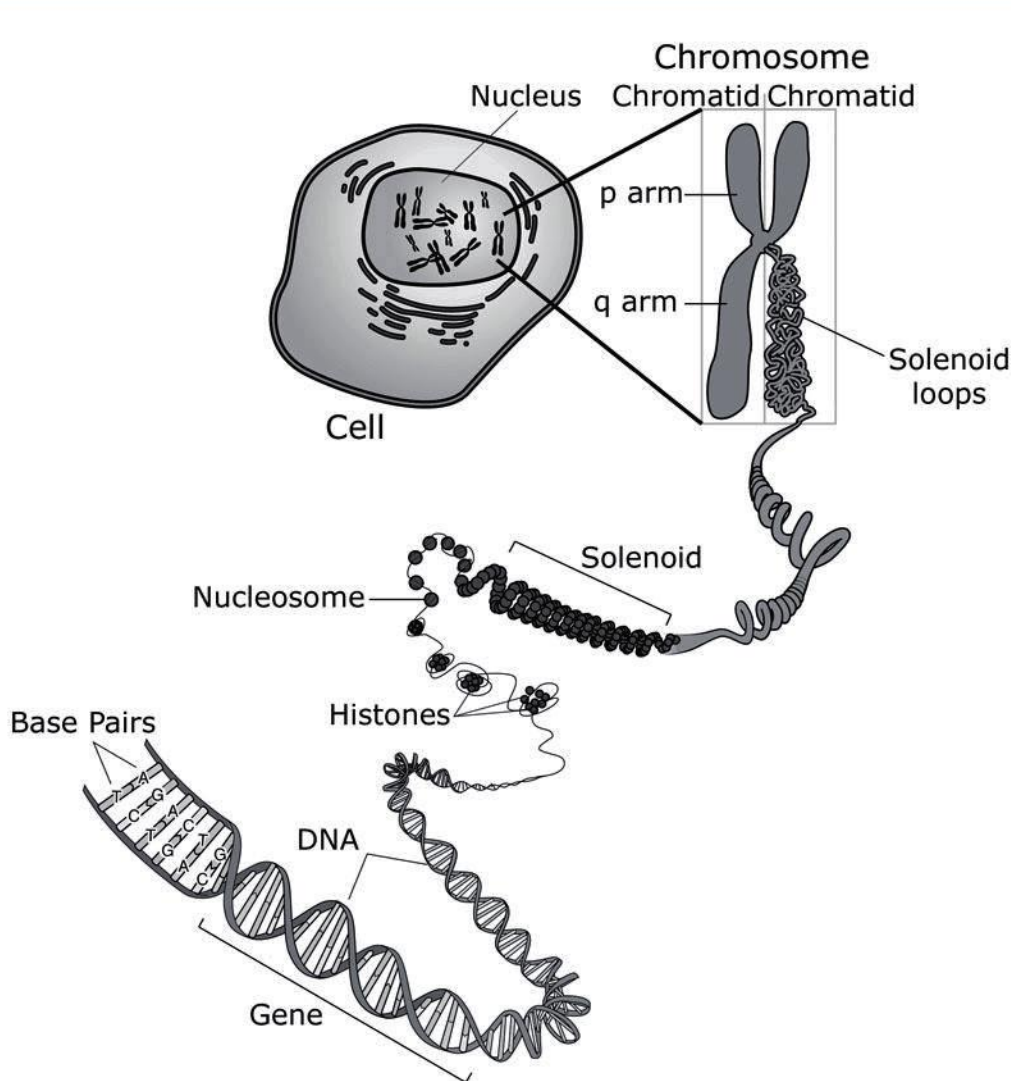
Paolo Ajmone Marsan e Licia Colli

Dipartimento di Scienze Animali della Nutrizione
e degli Alimenti – DIANA

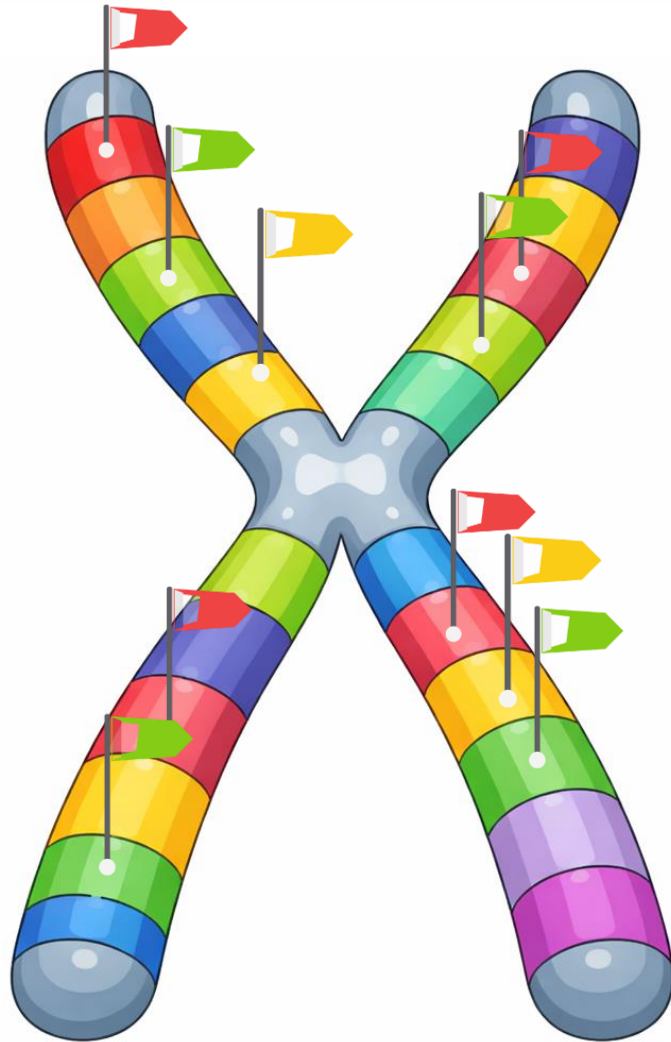
Università Cattolica del S. Cuore

paolo.ajmone@unicatt.it licia.colli@unicatt.it

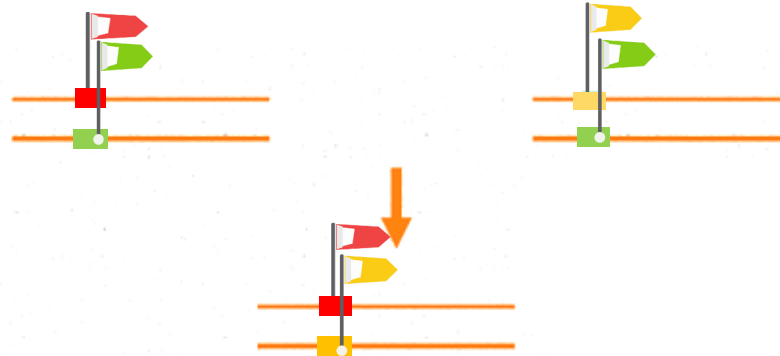
Il DNA e i genomi



I Marcatori Genetici



- Identificano delle mutazioni che si sono accumulate nel corso della storia evolutiva delle popolazioni;
- Sono facilmente visibili (dal vivo o in laboratorio);
- Hanno comportamento Mendeliano



- Diversità
- Consanguineità
- Struttura delle popolazioni
- Origine degli individui
- Regioni sotto selezione
- **Selezione assistita**

I marcatori molecolari SNP



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

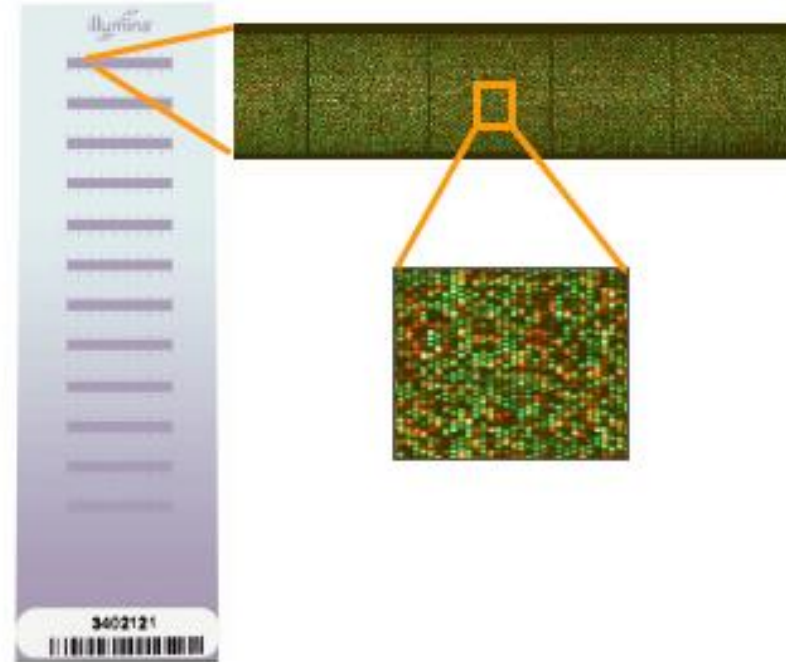
CCGTGACGGGA**G**CTGATGATCGTGT
...GGCACTGCCCT**C**GACTACTAGCACA...

CCGTGACGGGA**T**CTGATGATCGTGT
...GGCACTGCCCT**A**GACTACTAGCACA...



frontie

BeadChip



Development of a High-Density 665 K SNP Array for Rainbow Trout Genome-Wide Genotyping

Maria Bernard^{1,2}, Audrey Dehaillon¹, Guangtu Gao³, Katy Paul¹, Henri Lagarde¹,
Mathieu Charles^{1,2}, Martin Prchal⁴, Jeanne Danon⁵, Lydia Jaffrelo⁵, Charles Poncet⁵,
Pierre Patrice⁶, Pierrick Haffray⁶, Edwige Quillet¹, Mathilde Dupont-Nivet¹, Yniv Palti³,
Delphine Lallias¹ and Florence Phocas^{1*}

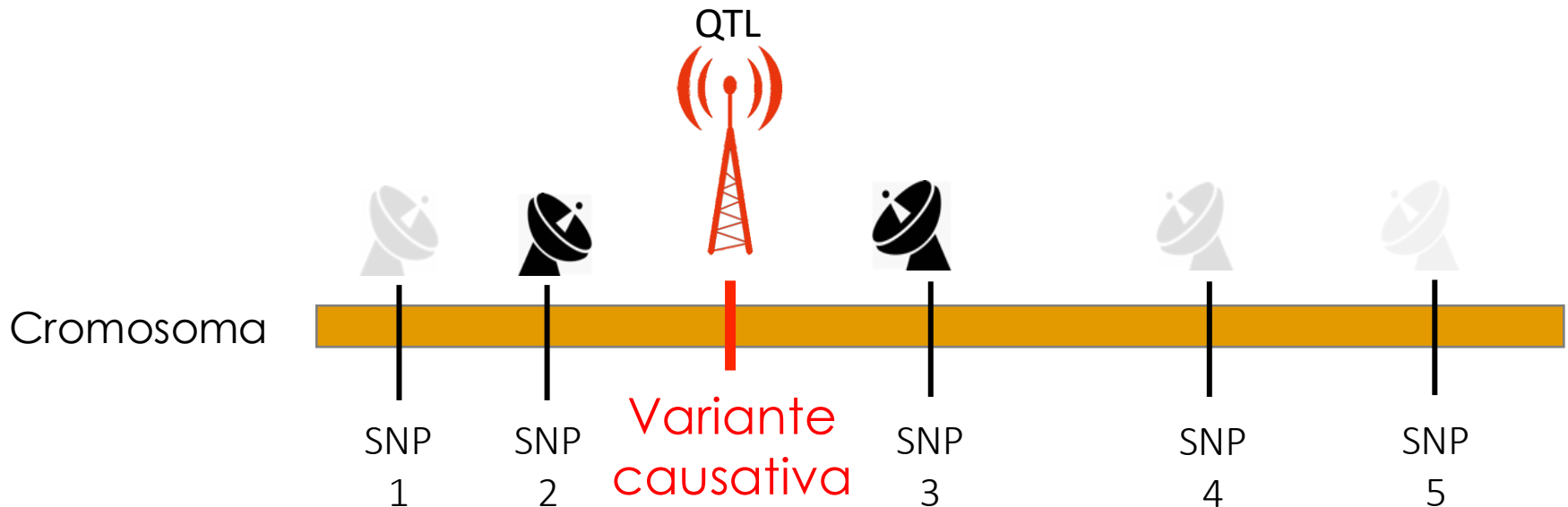
Trovati più di 3 milioni di SNP, scelti 665.000

Analisi di associazione Genome-Wide (GWAS)



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

Il fatto che il genoma sia 'a blocchi' permette la rilevazione del segnale da parte dei marcatori nelle vicinanze

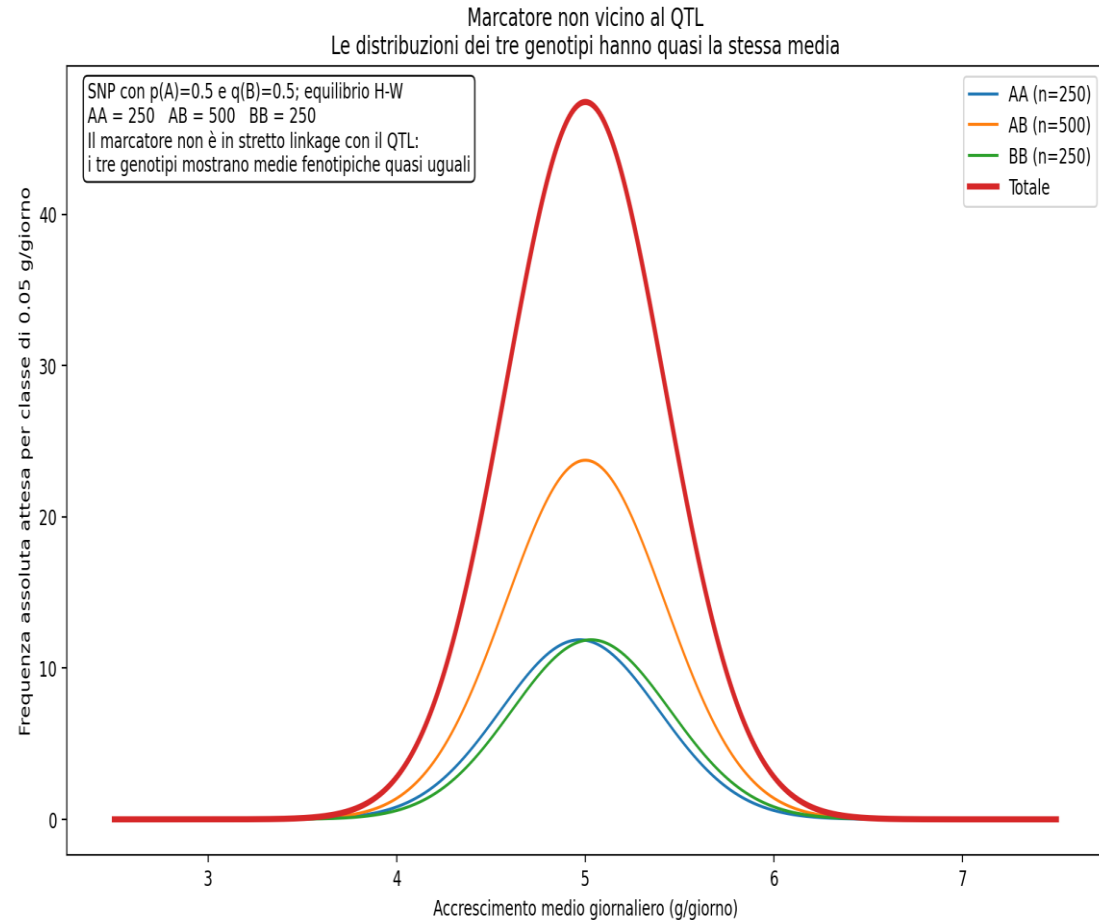


QTL =Quantitative Trait Locus

Analisi di associazione Genome-Wide (GWAS)



- Inizio dallo SNP1, che ha due alleli (A e B)
- Divido gli individui in gruppi per genotipo (AA, AB e BB)
- Guardo se le medie dei gruppi sono significativamente diverse
- Non lo sono, procedo

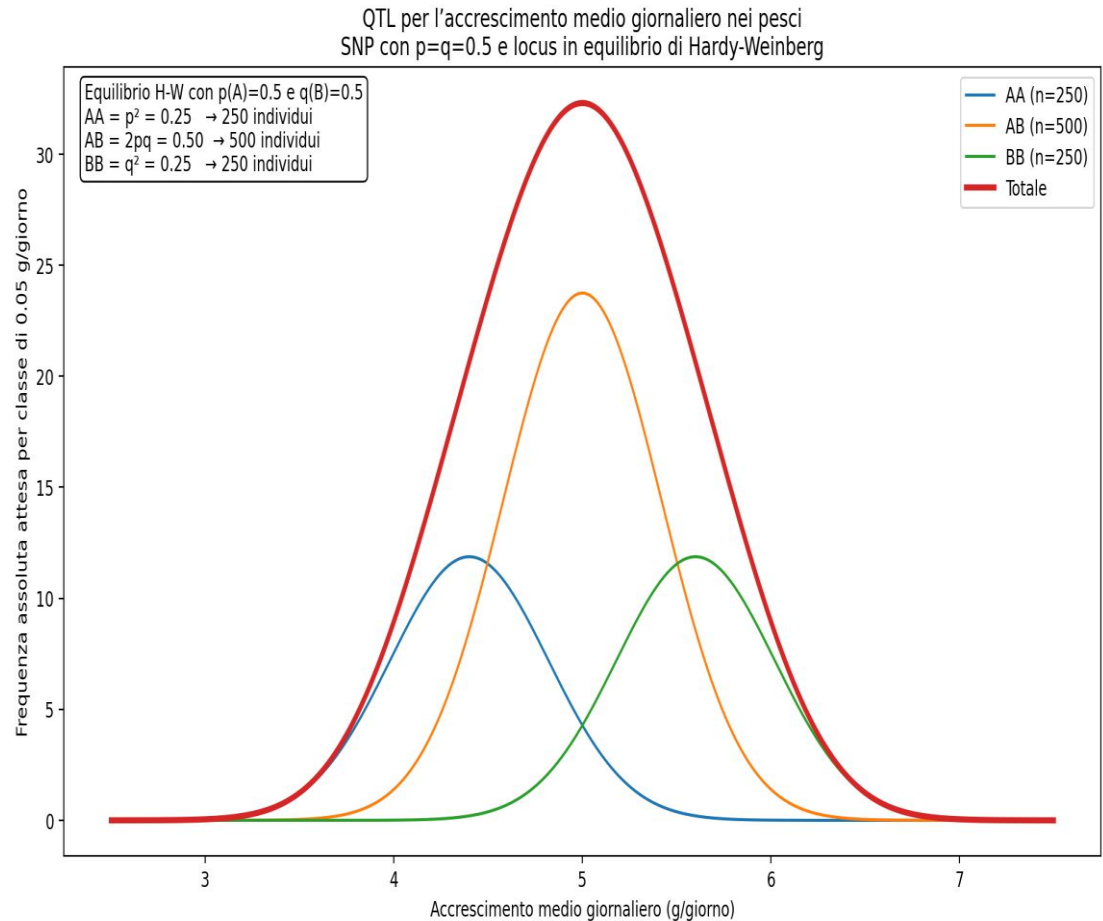


Analisi di associazione Genome-Wide (GWAS)



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

- Procedo con lo SNP2, ecc.
- Divido gli individui in gruppi per genotipo e controllo le medie
- Arrivo allo SNP2045
- Le medie dei gruppi sono significativamente diverse
- C'è un QTL nelle vicinanze
- Noto che B è l'allele migliore
- **Uso l'allele B dello SNP2045 nella selezione.**



Limiti GWAS



- Bassa potenza statistica
- Identifica solo gli effetti più rilevanti → successo dipende dall'architettura genetica del carattere
- Influenzata dalla struttura della popolazione (es. suddivisione in famiglie)
- Deve tenere conto del multiple testing (molte migliaia di test indipendenti)
- **Per caratteri con architettura genetica complessa → Genomic selection**

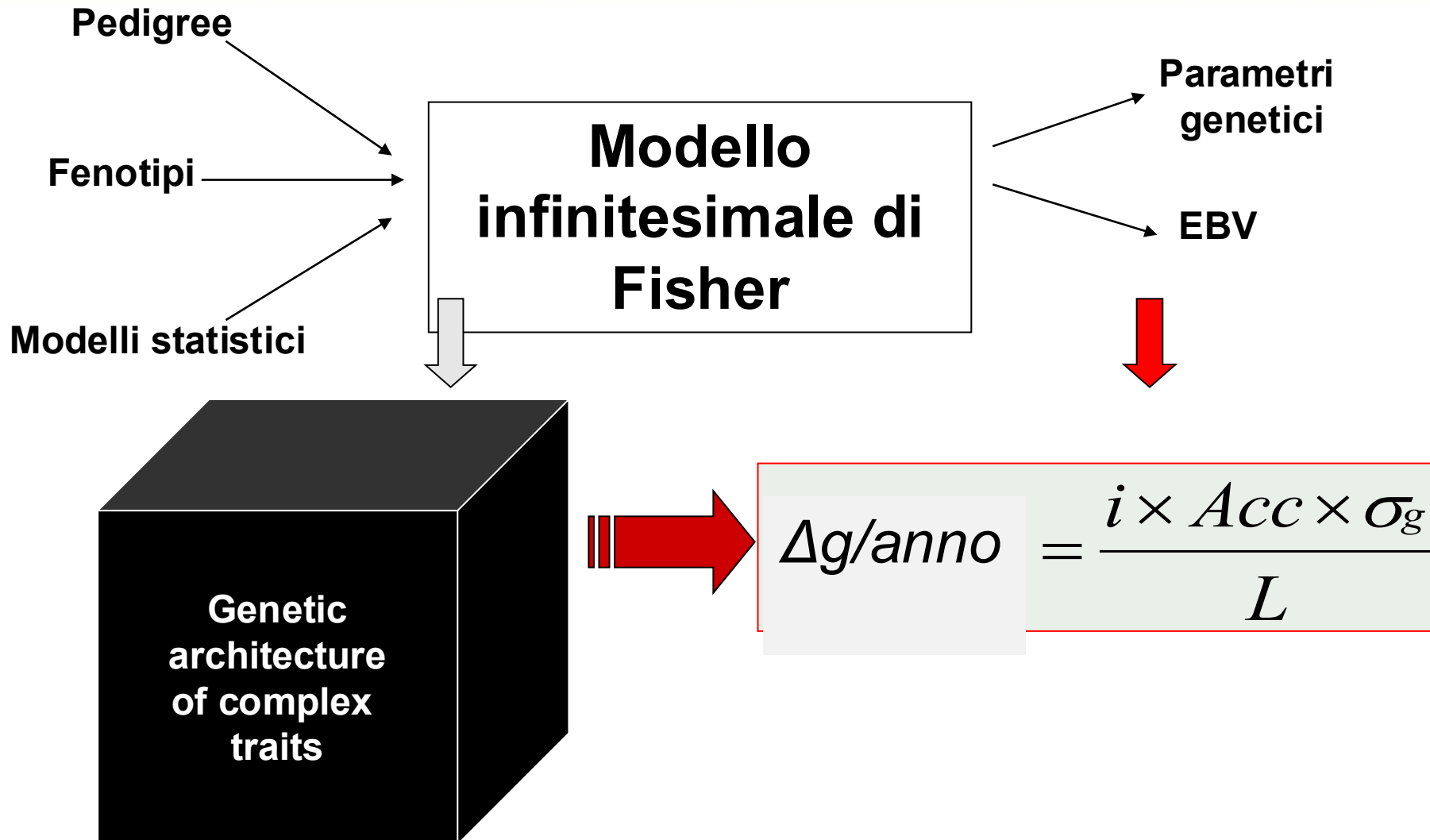
La selezione genetica



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

$$P = G + E$$

Selezione tradizionale



Parentela

Rassomiglianza fra parenti

- > grado di parentela
- > rassomiglianza

$$\mathbf{A} = \begin{bmatrix} 1+f_1 & a_{12} & a_{13} & a_{14} & a_{15} \\ a_{21} & 1+f_2 & a_{23} & a_{24} & a_{25} \\ a_{31} & a_{32} & 1+f_3 & a_{34} & a_{35} \\ a_{41} & a_{42} & a_{43} & 1+f_4 & a_{45} \\ a_{51} & a_{52} & a_{53} & a_{54} & 1+f_5 \end{bmatrix}$$

Esiste una covarianza genetica fra individui

Assunzioni:

Popolazione di base

- individui non parenti
- individui non consanguinei

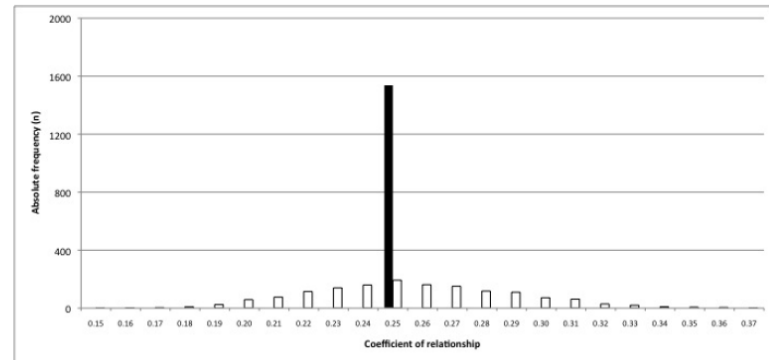
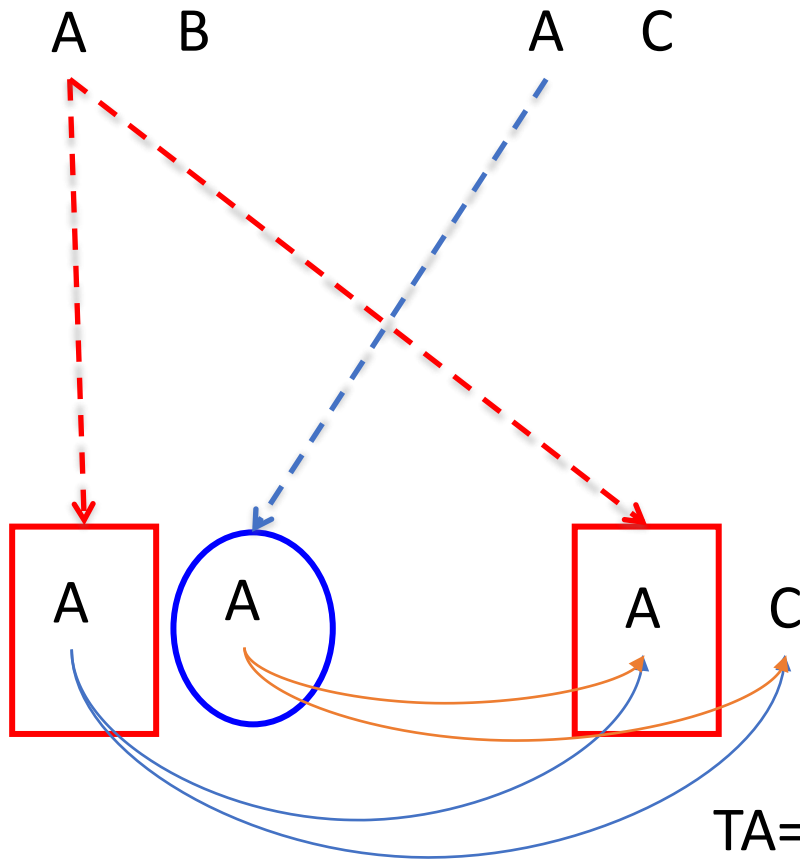
Trasmissione nelle generazioni

- probabilità non osservazione
- corrette in media
- tutte sbagliate in pratica

Parentela genomica



La relazione allelica tra due individui x e y ad un determinato locus si ottiene confrontando ogni allele di x con i due alleli di y



The black bar indicates the average relationship (i.e. currently used in 'traditional' genetic evaluations), whereas the white bars indicate the genomic relationship for the same bulls. Unpublished data.

$$TA_1 = \frac{\sum_{i=1}^2 \sum_{j=1}^2 I_{ij}}{2}$$

Selezione genomica



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

Copyright © 2001 by the Genetics Society of America

Prediction of Total Genetic Value Using Genome-Wide Dense Marker Maps

T. H. E. Meuwissen,* B. J. Hayes[†] and M. E. Goddard^{†,‡}

**Research Institute of Animal Science and Health, 8200 AB Lelystad, The Netherlands,* [†]*Victorian Institute of Animal Science, Attwood 3049, Victoria, Australia*
and [‡]*Institute of Land and Food Resources, University of Melbourne, Parkville 3052, Victoria, Australia*

Manuscript received August 17, 2000 Accepted for publication January 17, 2001

ABSTRACT

Recent advances in molecular genetic techniques will make dense marker maps available and genotyping many individuals for these markers feasible. Here we attempted to estimate the effects of ~50,000 marker haplotypes simultaneously from a limited number of phenotypic records. A genome of 1000 cM was simulated with a marker spacing of 1 cM. The markers surrounding every 1-cM region were combined into marker haplotypes. Due to finite population size ($N_e = 100$), the marker haplotypes were in linkage disequilibrium with the QTL located between the markers. Using least squares, all haplotype effects could not be estimated simultaneously. When only the biggest effects were included, they were overestimated and the accuracy of predicting genetic values of the offspring of the recorded animals was only 0.32. Best linear unbiased prediction of haplotype effects assumed equal variances associated to each 1-cM chromosomal segment, which yielded an accuracy of 0.73, although this assumption was far from true. Bayesian methods that assumed a prior distribution of the variance associated with each chromosome segment increased this accuracy to 0.85, even when the prior was not correct. It was concluded that selection on genetic values predicted from markers could substantially increase the rate of genetic gain in animals and plants, especially if combined with reproductive techniques to shorten the generation interval.

Popolazioni di training e validazione

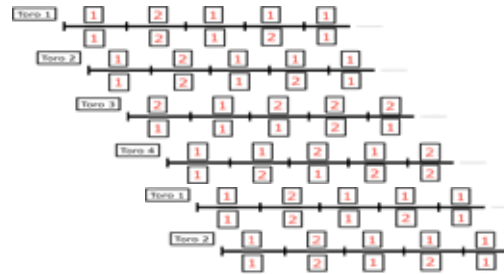


Training

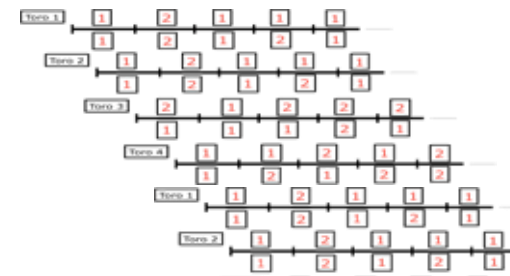
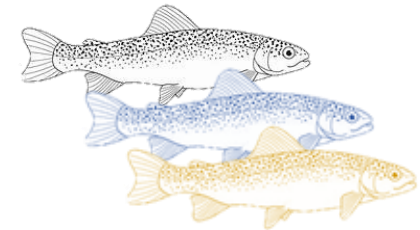
Fenotipi



Genotipi



Validazione



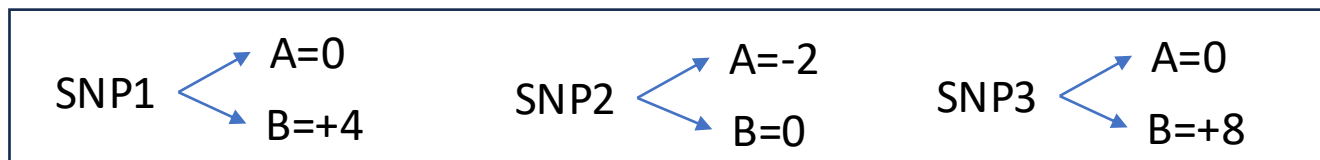
$$\text{Fenotipo}_1 = \text{effetto}_{\text{SNP1}} + \text{effetto}_{\text{SNP2}} + \dots + \text{effetto}_{\text{SNP54k}} + e_i$$

$$\text{Fenotipo}_2 = \text{effetto}_{\text{SNP1}} + \text{effetto}_{\text{SNP2}} + \dots + \text{effetto}_{\text{SNP54k}} + e_i$$

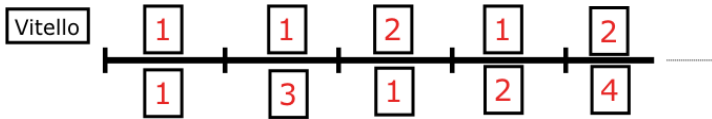
.....

$$\text{Fenotipo}_n = \text{effetto}_{\text{SNP1}} + \text{effetto}_{\text{SNP2}} + \dots + \text{effetto}_{\text{SNP54k}} + e_i$$

Si risolve il sistema di equazioni e si stimano gli effetti degli alleli a ciascuno SNP



La selezione genomica in azione



$$EBV_i = \text{effect}_{\text{SNP1}} + \text{effect}_{\text{SNP2}} + \dots + \text{effect}_{\text{SNP54k}} + e_i$$

$$100 = (+1) + (-4) + \dots + (-2)$$

$$DGV_i = g_{\text{SNP1}} * \text{effect}_{\text{SNP1}} + g_{\text{SNP2}} * \text{effect}_{\text{SNP2}} + \dots + g_{\text{SNP54k}} * \text{effect}_{\text{SNP54k}} + e_i$$

$$DGV_i = 0 * (+1) + 2 * (-4) + \dots + 1 * (-2)$$

DGV



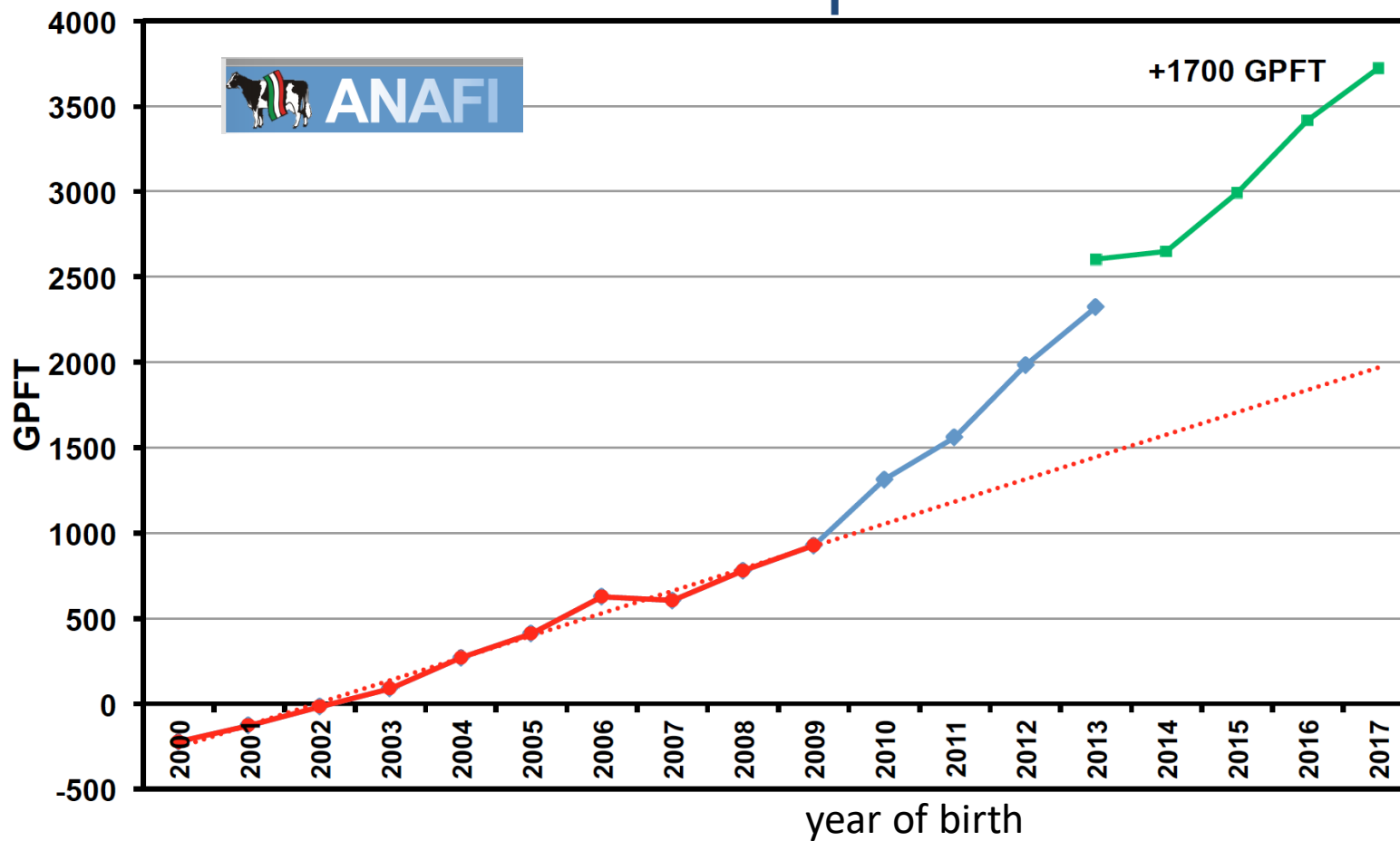
G-EBV

Vantaggi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

$$\Delta g / \text{anno} = \frac{i \times Acc \times \sigma_g}{L}$$



Raffaella
Finocchiaro



Jan-Thijs
Van Kaam

Great impact in (dairy cattle) breeding