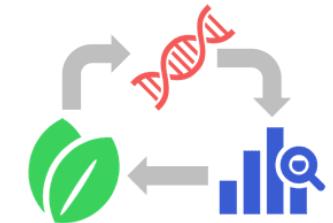


REGIONAL TRAINING COURSE

MUTATION BREEDING AND MOLECULAR TECHNIQUE FOR CROP IMPROVEMENT

Genomic Prediction



29/04/2024 ~ 03/05/2024

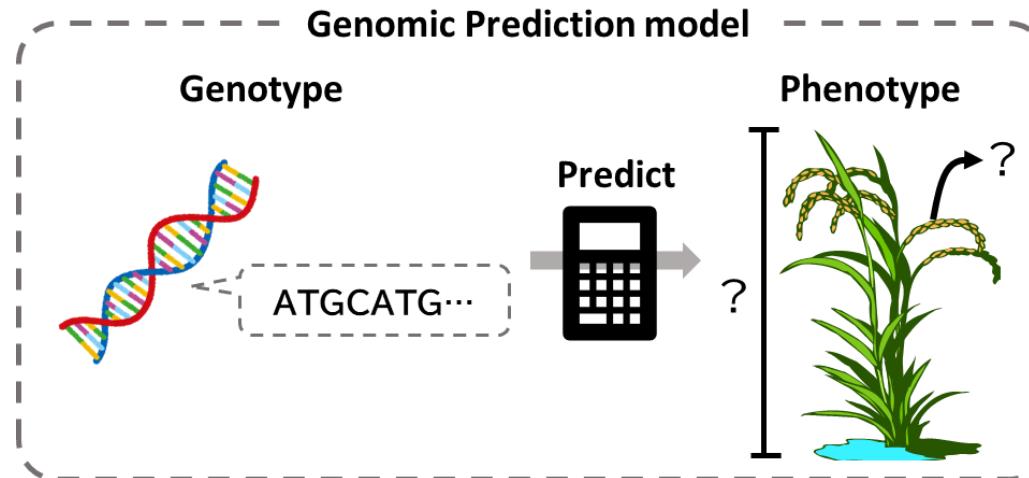
Toshiyuki Sakai
Crop evolution laboratory in Kyoto University

Today's contents

Genomic Prediction

- Basic concept & example
- Practice Session
- Ideas for application of genomic prediction

Genomic Prediction is used for Genomic Selection

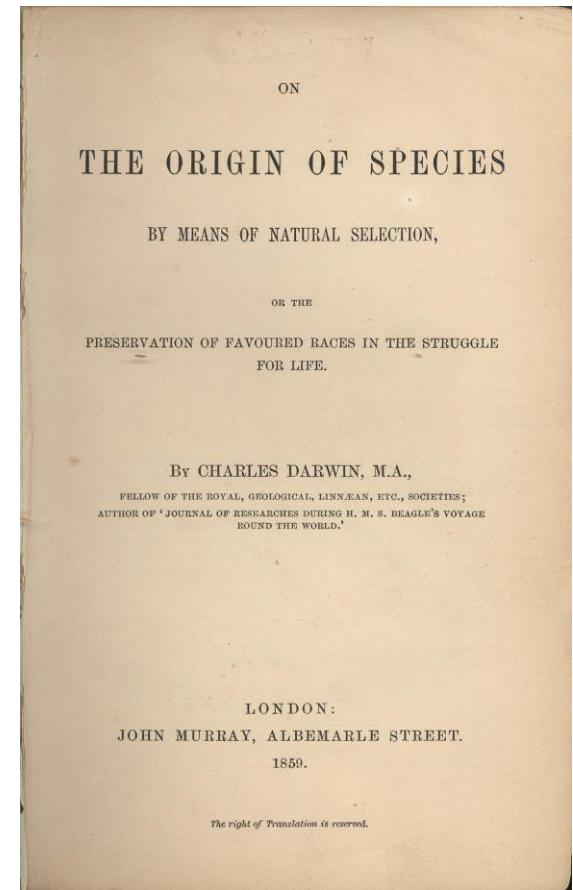
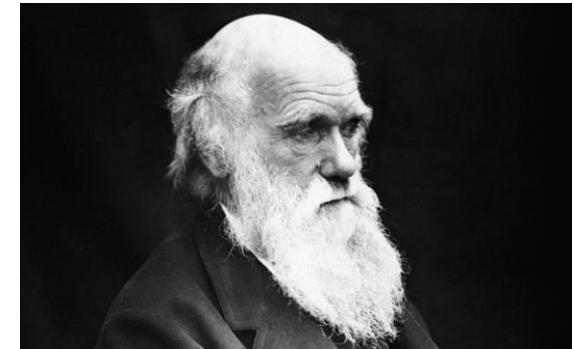


Genomic Selection

“Selection” is important for breeding

“Selective breeding” (Charles Darwin, 1859)

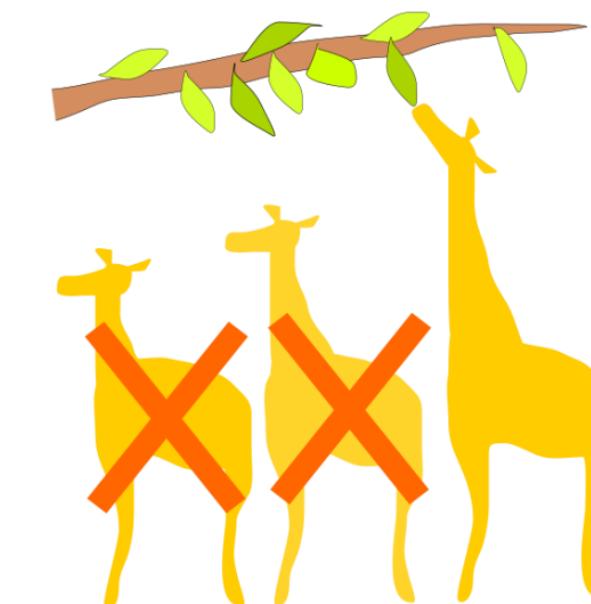
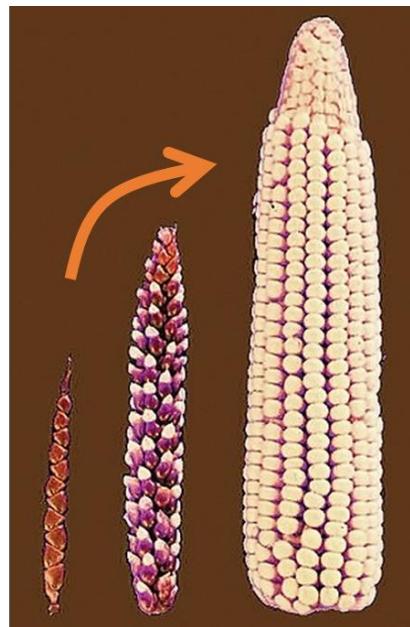
Slow though the process of selection may be, if feeble man can do much by his powers of **artificial selection**, I can see no limit to the amount of change, to the beauty and infinite complexity of the co-adaptations between all organic beings, one with another and with their physical conditions of life, which may be effected in the long course of time by nature's power of selection.



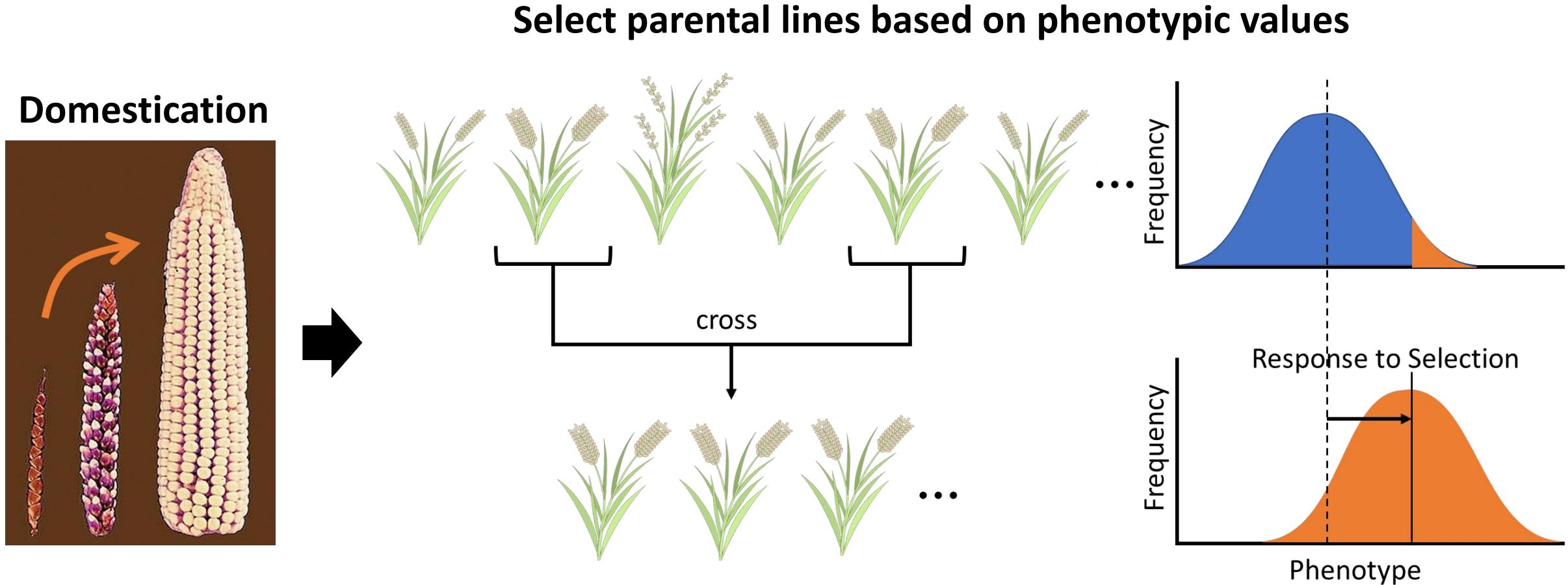
Artificial selection



Natural selection

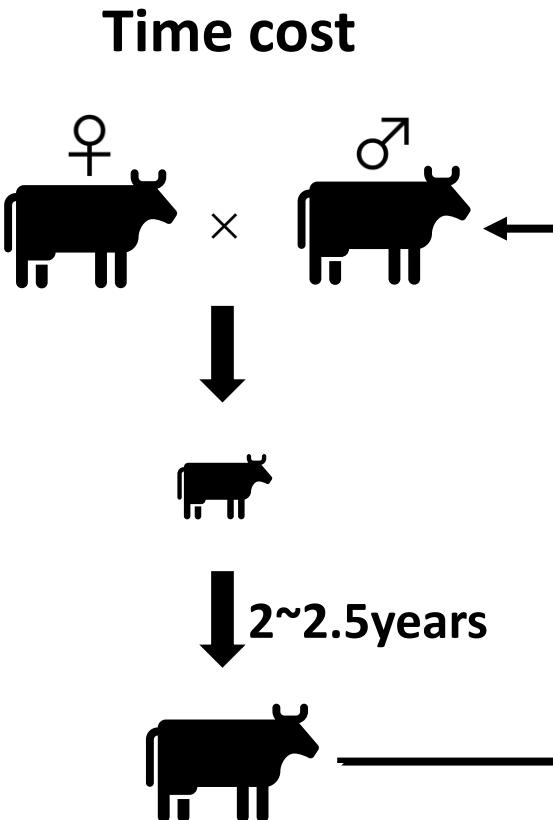


Traditional breeding is based on phenotypic selection



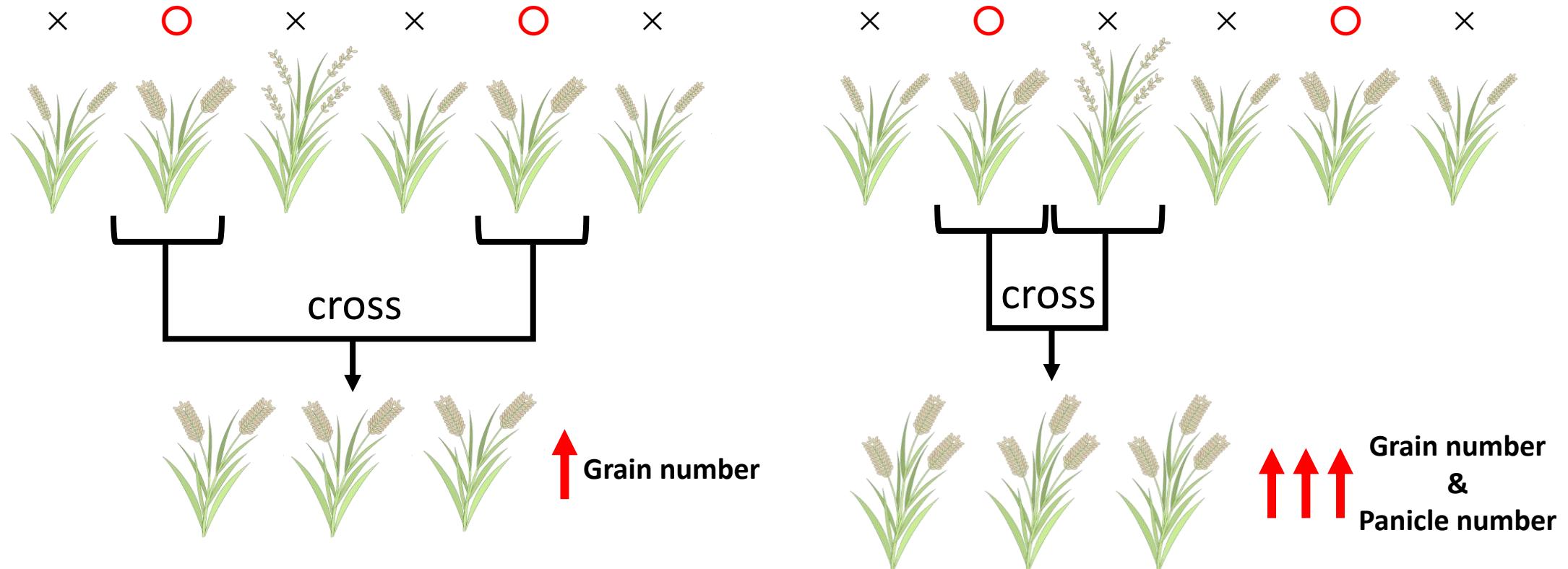
Phenotypic selection has limitations

1. Time cost



Phenotypic selection has limitations

2. Ignore possibility



→ Method to accurately evaluate potential of cultivars is required

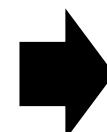
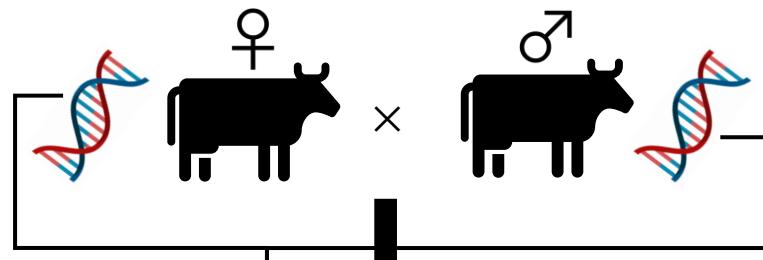
We must evaluate the potential from Genotype to solve limitations

- Phenotypic selection has limitations

Time cost & can't evaluate potential



- Genotype was inherited from parents. Phenotype is result.

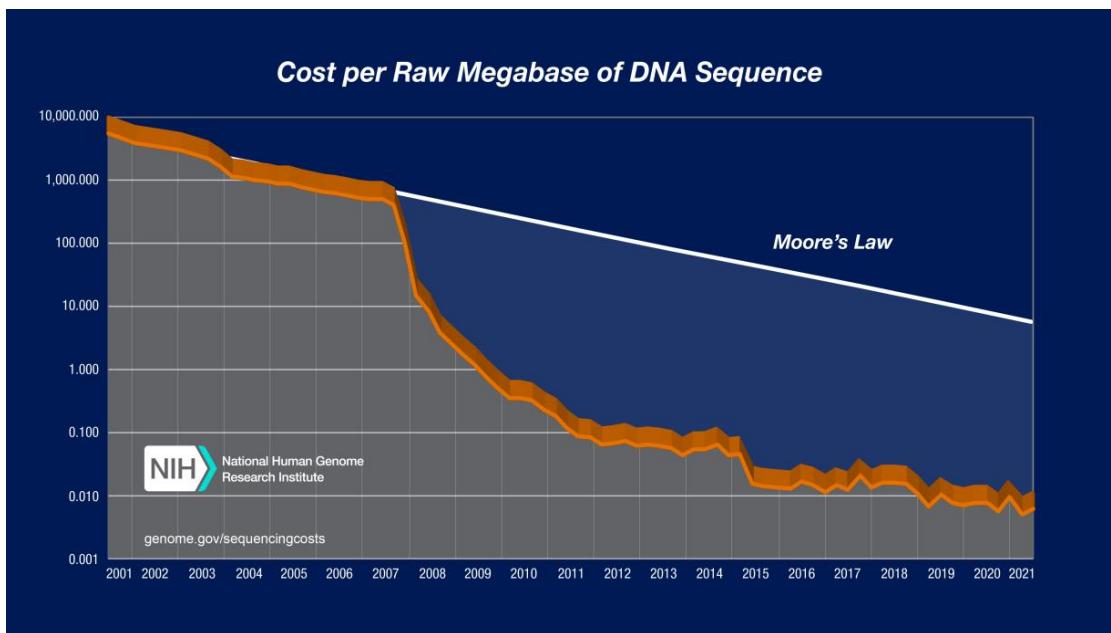


Selection method based on Genotype is better?

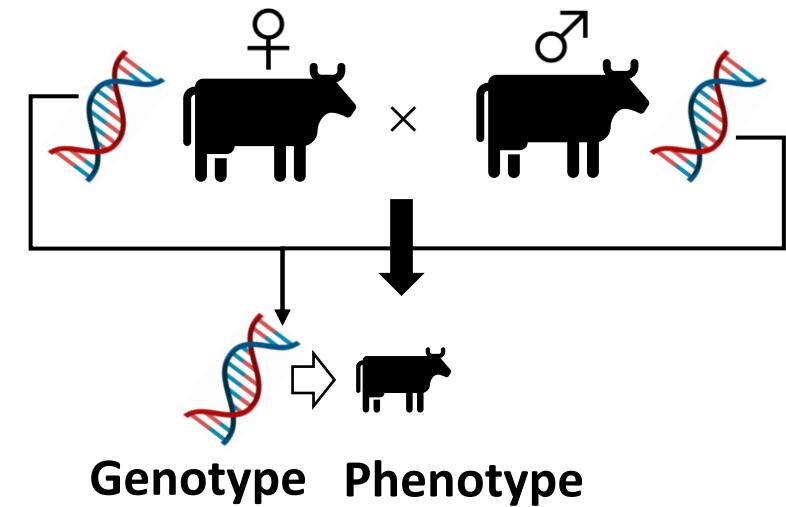
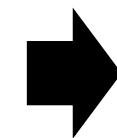
Genotype Phenotype

Low sequencing cost enable us to do genomic approach

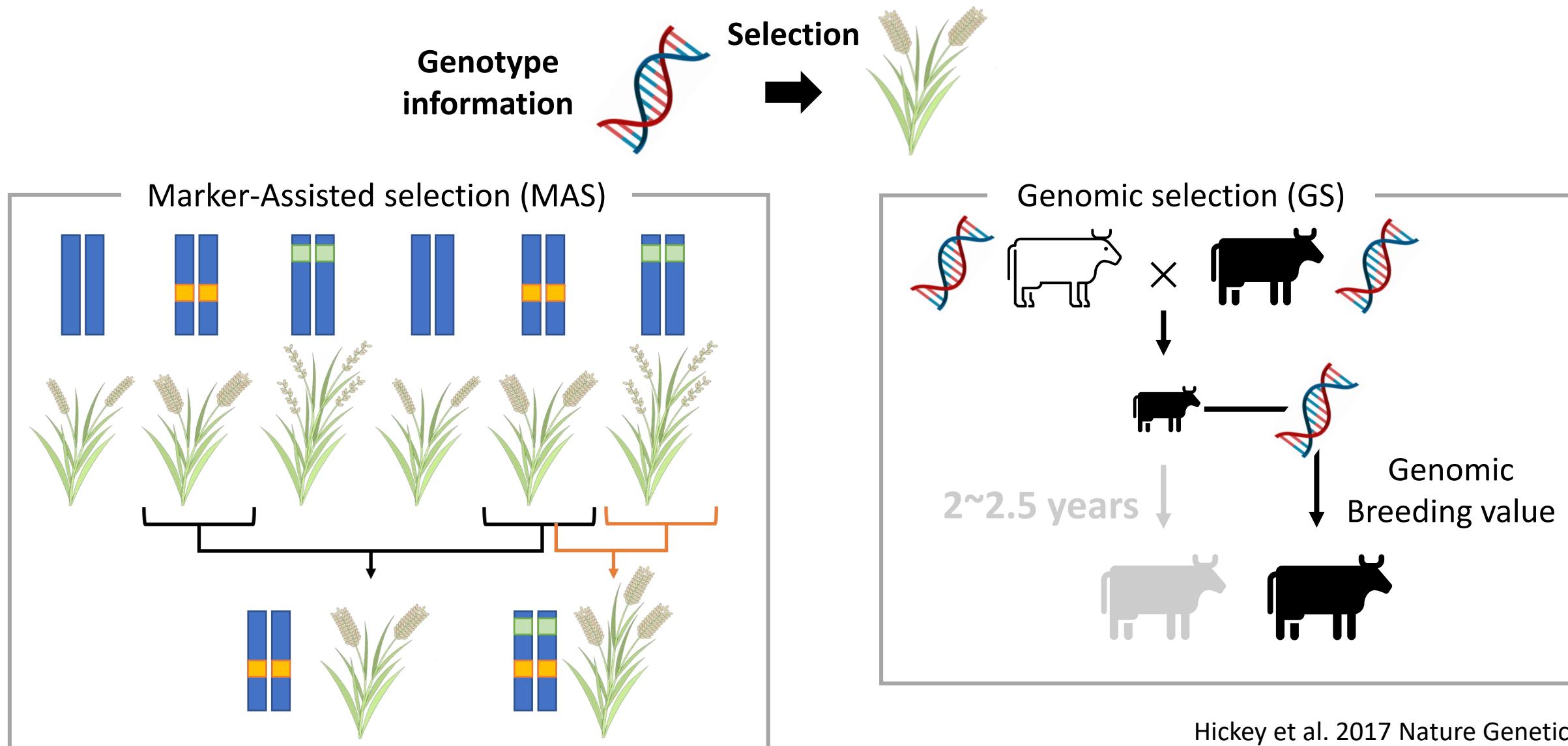
Sequencing cost is decreasing



We can use genotype information in selection

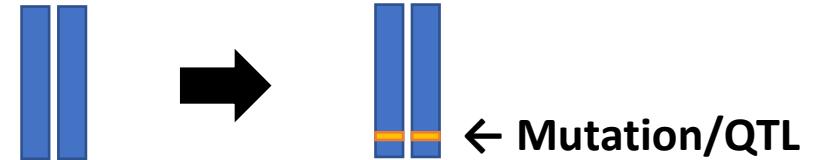


Recently, genotype information is utilized in selective breeding

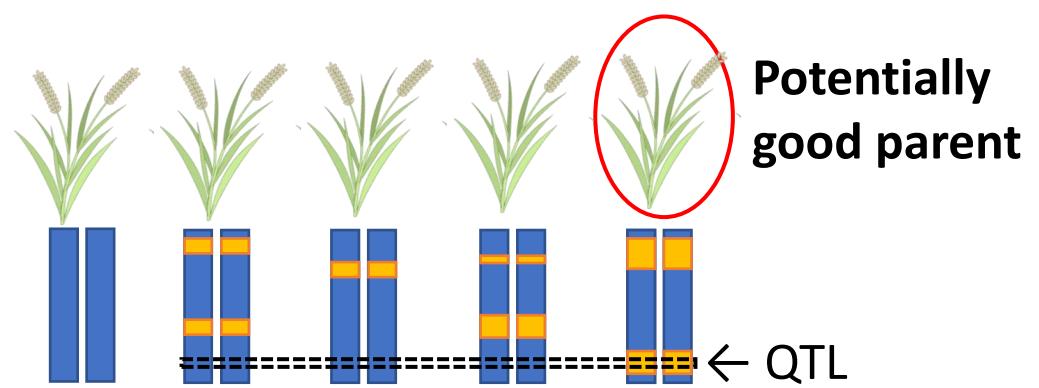


Marker-Assisted selection (MAS)

1. Identify QTLs of target traits



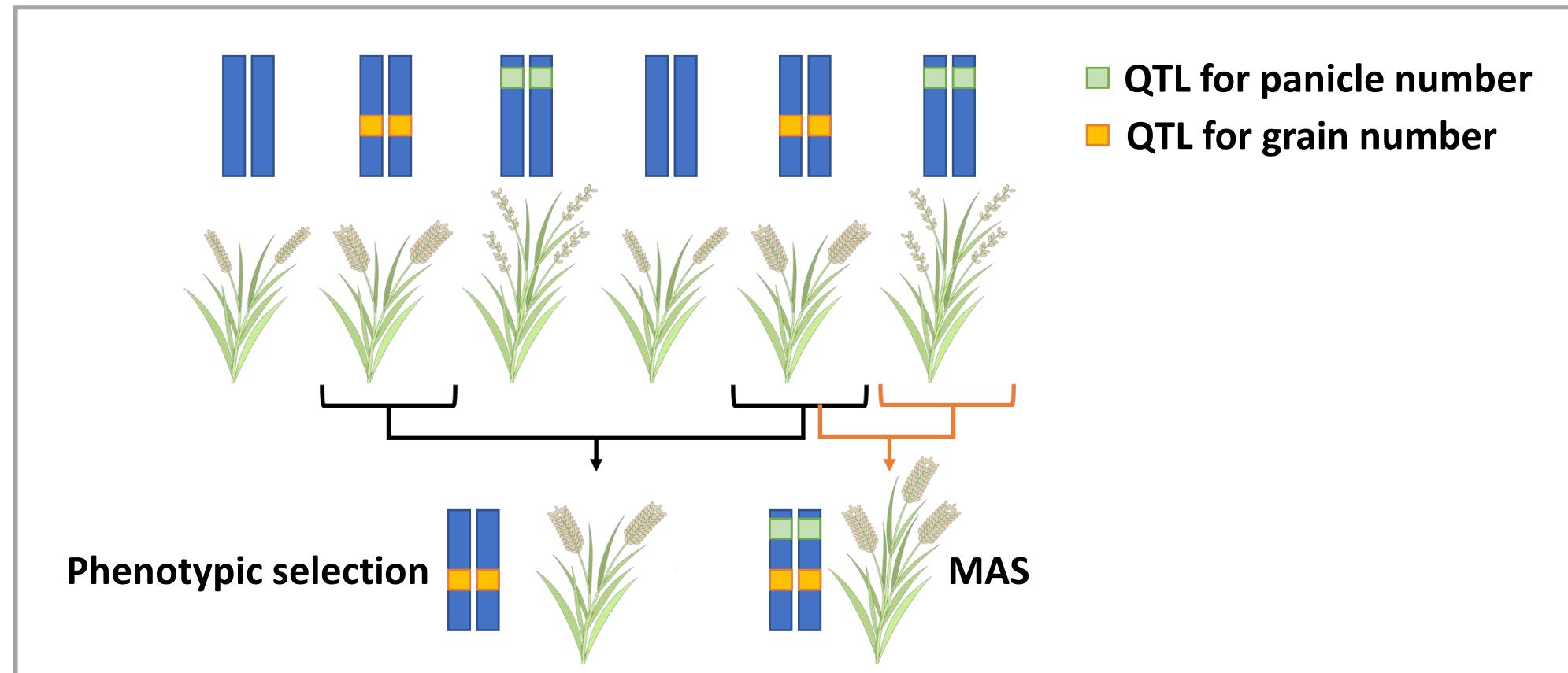
2. Select line based on genotype of QTL



GWAS, QTL-mapping, ... etc

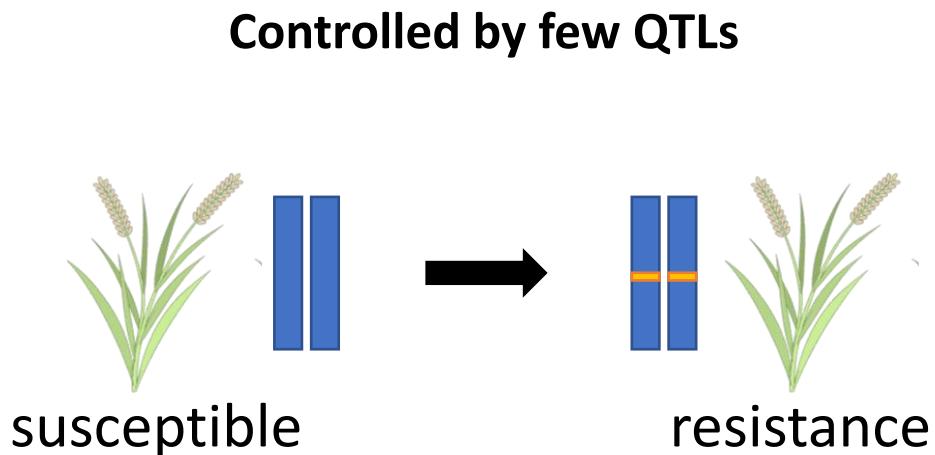
Ribaut and Hoisington. 1998 Trends in Plant Science
Lande and Thompson. 1990 Genetics

We can select good combination of parents based on MAS (ex. Gene pyramiding)

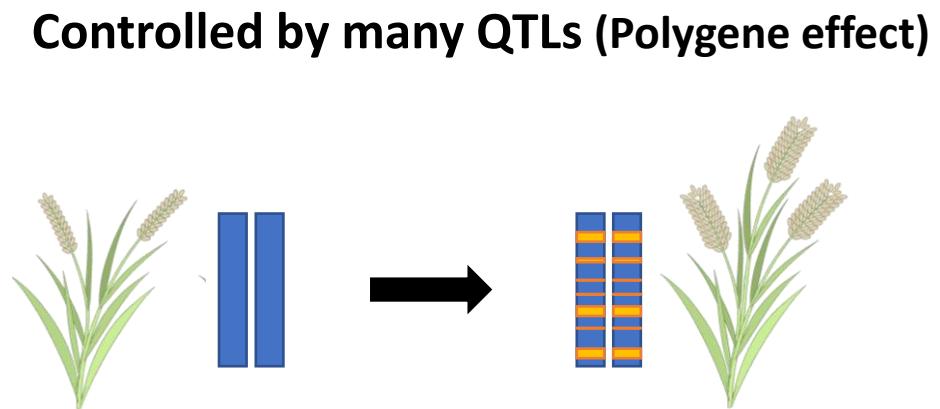


Marker Assisted Selection has limitations for traits with many QTLs

- **Simple traits** (ex. Resistant or Susceptible)



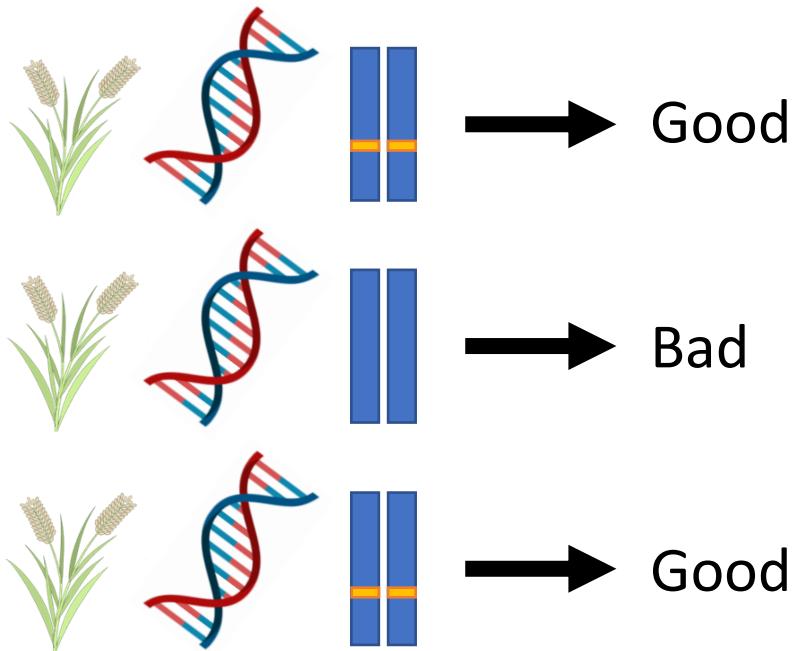
- **Complex traits** (ex. Yields)



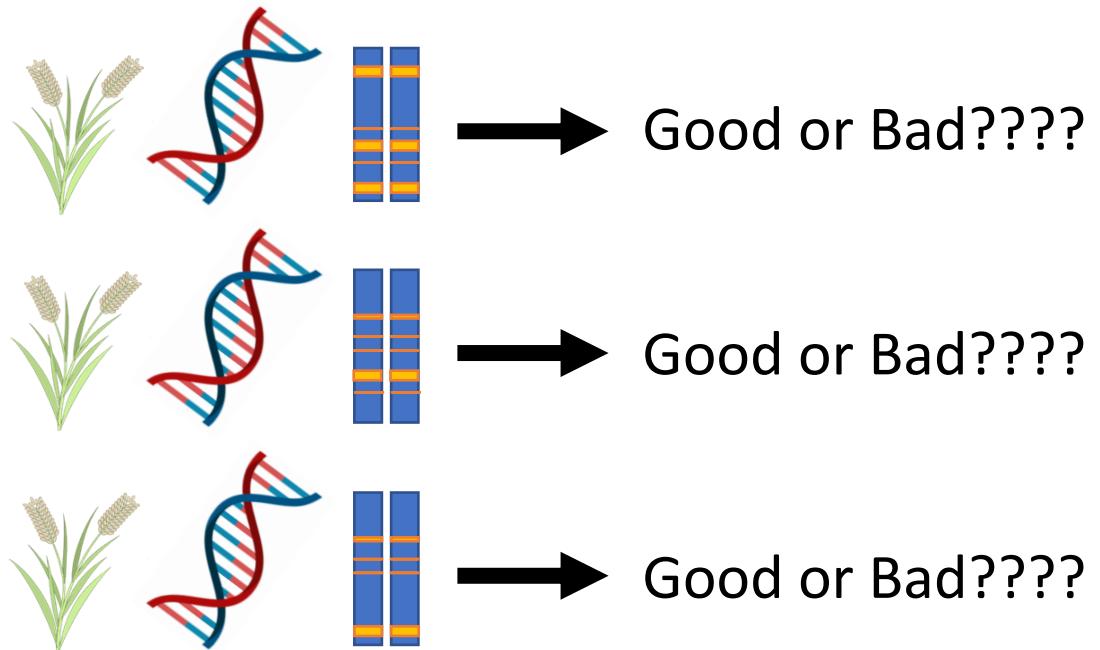
We have to focus only on strong QTL.
(miss other QTLs.)

Selection methods based on whole genotype are required

- Simple traits

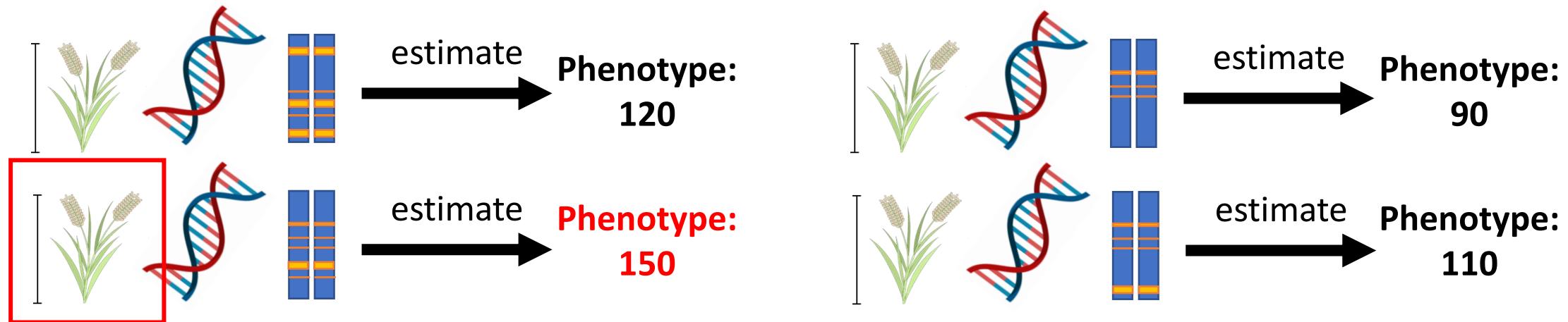


- Complex traits (ex. Yields)

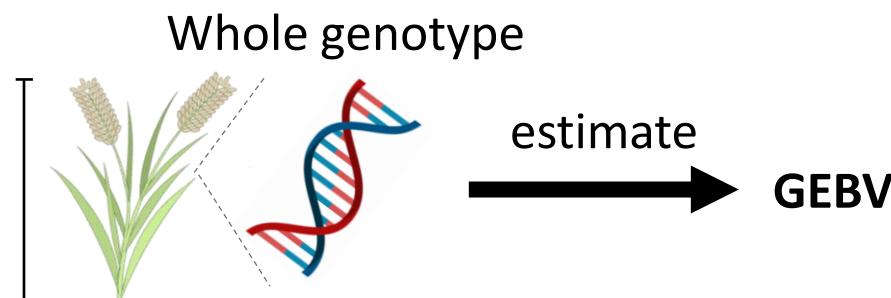


If we can evaluate potential from whole genotype, we can select best cultivar

- Genomic Selection



- Genomic estimated breeding value (GEBV)



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

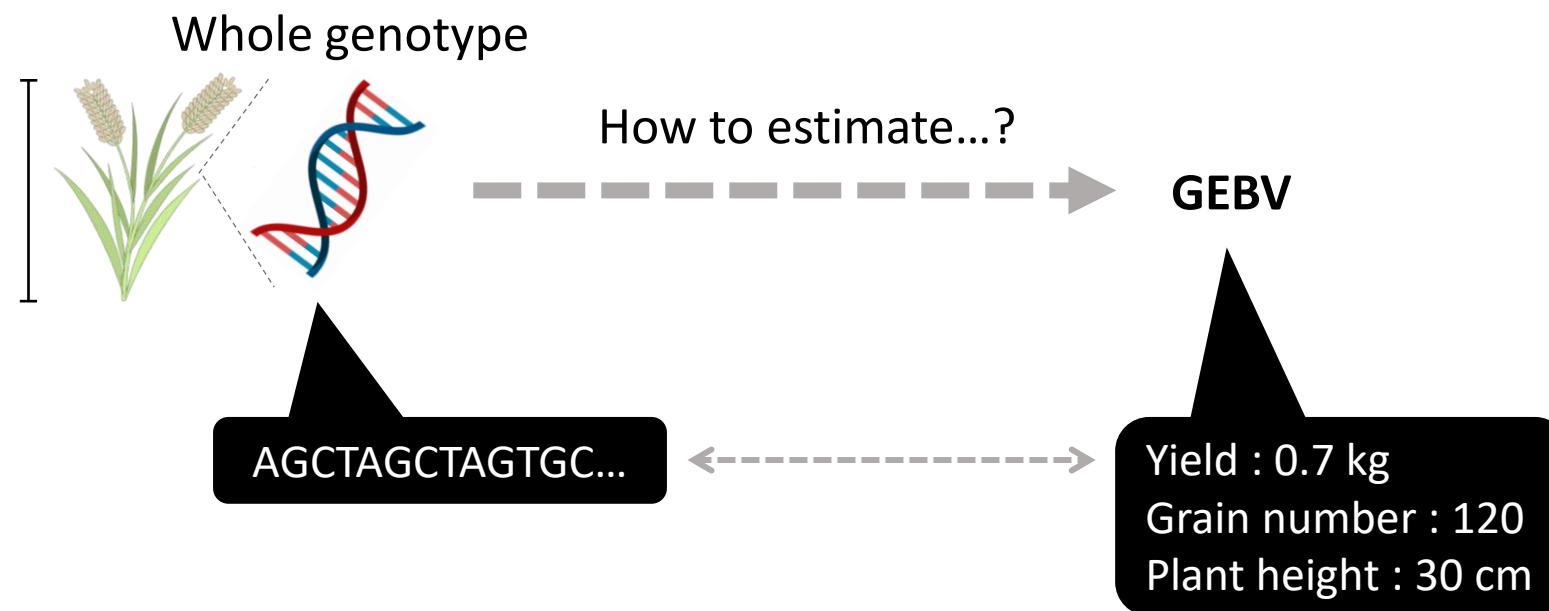
Manuscript received August 17, 2000

Accepted for publication January 17, 2001

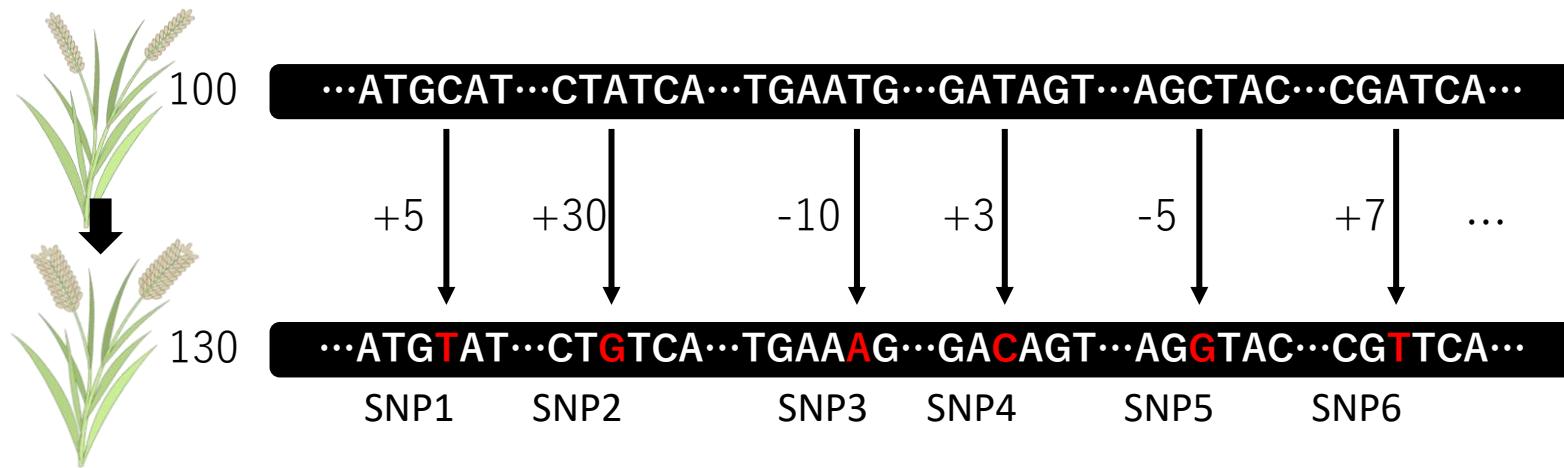
Meuwissen et al., 2001 Genetics

So, How to estimate breeding values?

- For Genomic Selection, we have to estimate Genomic estimated breeding value (GEBV)



To estimate GEBV, genomic prediction(GP) model is required

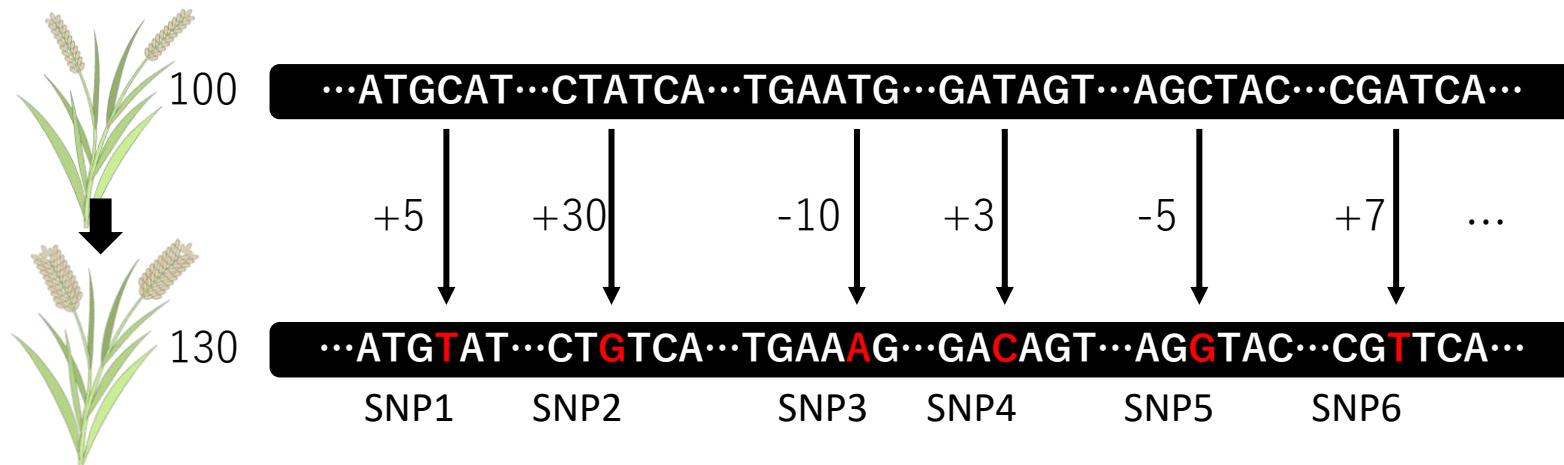


Genomic prediction model



ex) Phenotype = 100 + SNP1 effect + SNP2 effect + ... + SNP6 effect

To estimate GEBV, genomic prediction(GP) model is required



Genomic prediction model



ex) Phenotype = 100 + SNP1 effect + SNP2 effect + ... + SNP6 effect

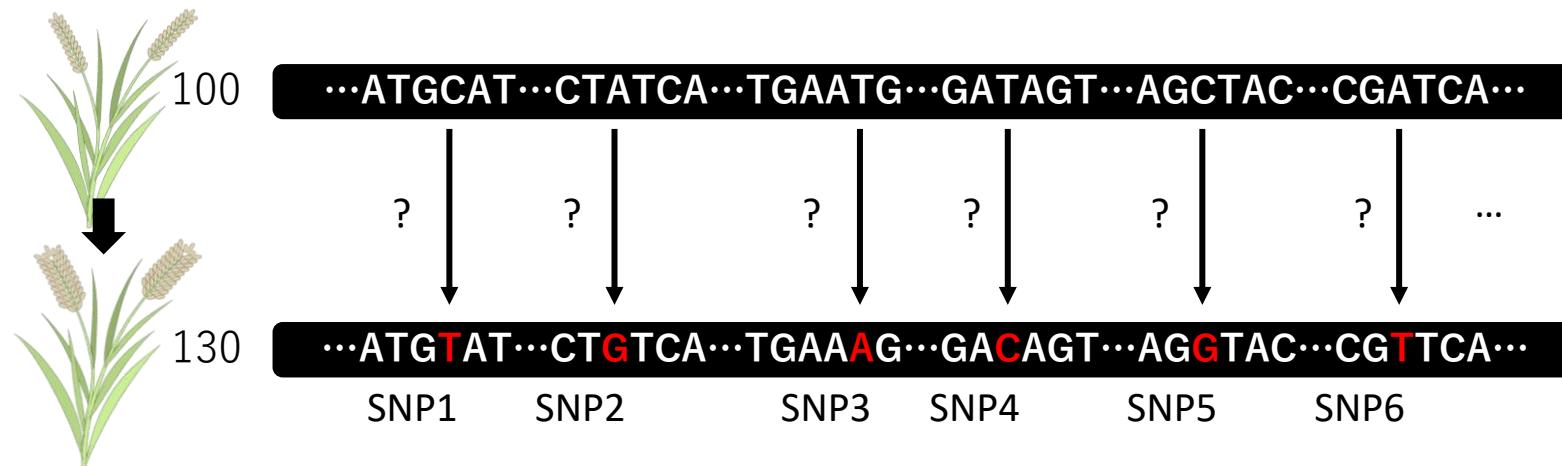


SNP3 SNP5

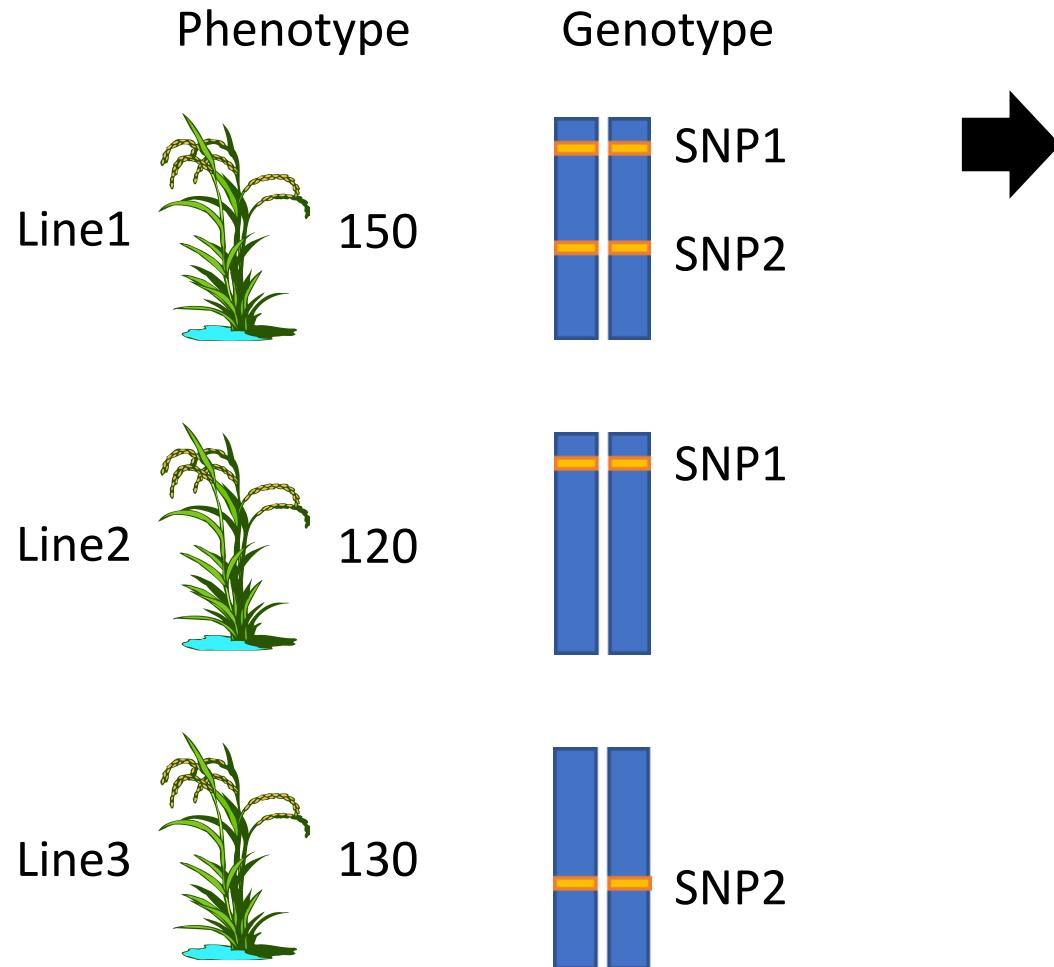
...ATGCAT...CTATCA...TGAA**AG**...GATAGT...AG**G**TAC...CGATCA...

$$\text{Phenotype} = 100 - 10 - 5 = 85$$

Next question is ... How to estimate SNP effect ... ?

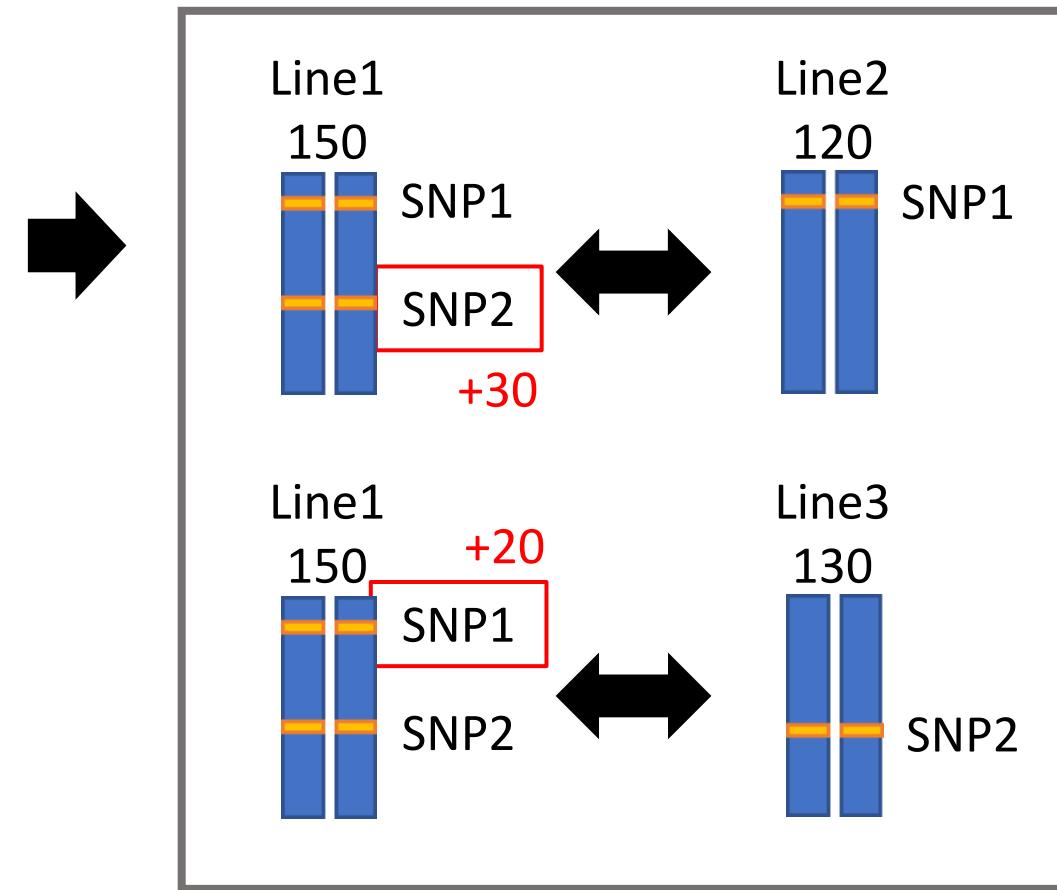


Very simple example: How to make genomic prediction model

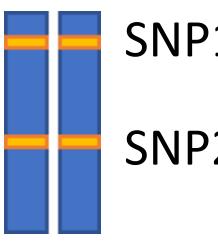
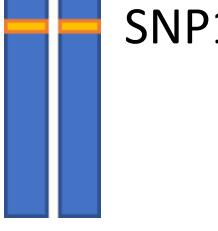
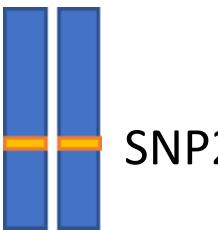


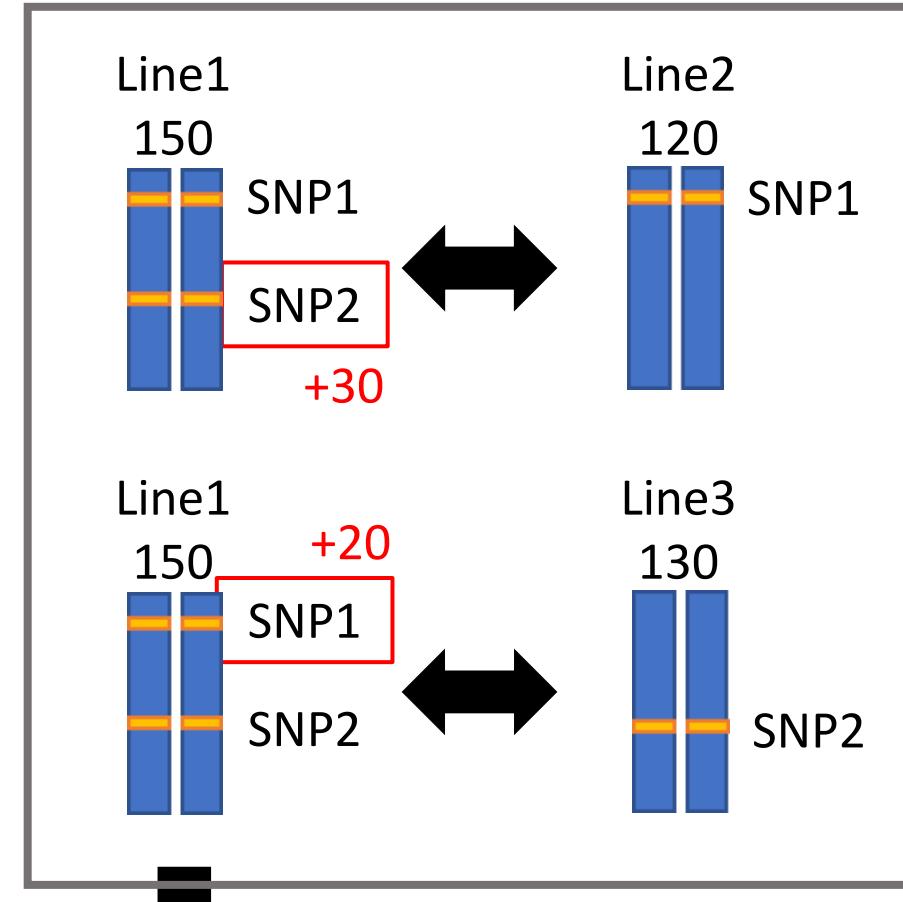
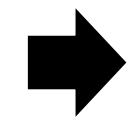
Very simple example: How to make genomic prediction model

	Phenotype	Genotype
Line1	150	 SNP1 SNP2
Line2	120	 SNP1
Line3	130	 SNP2



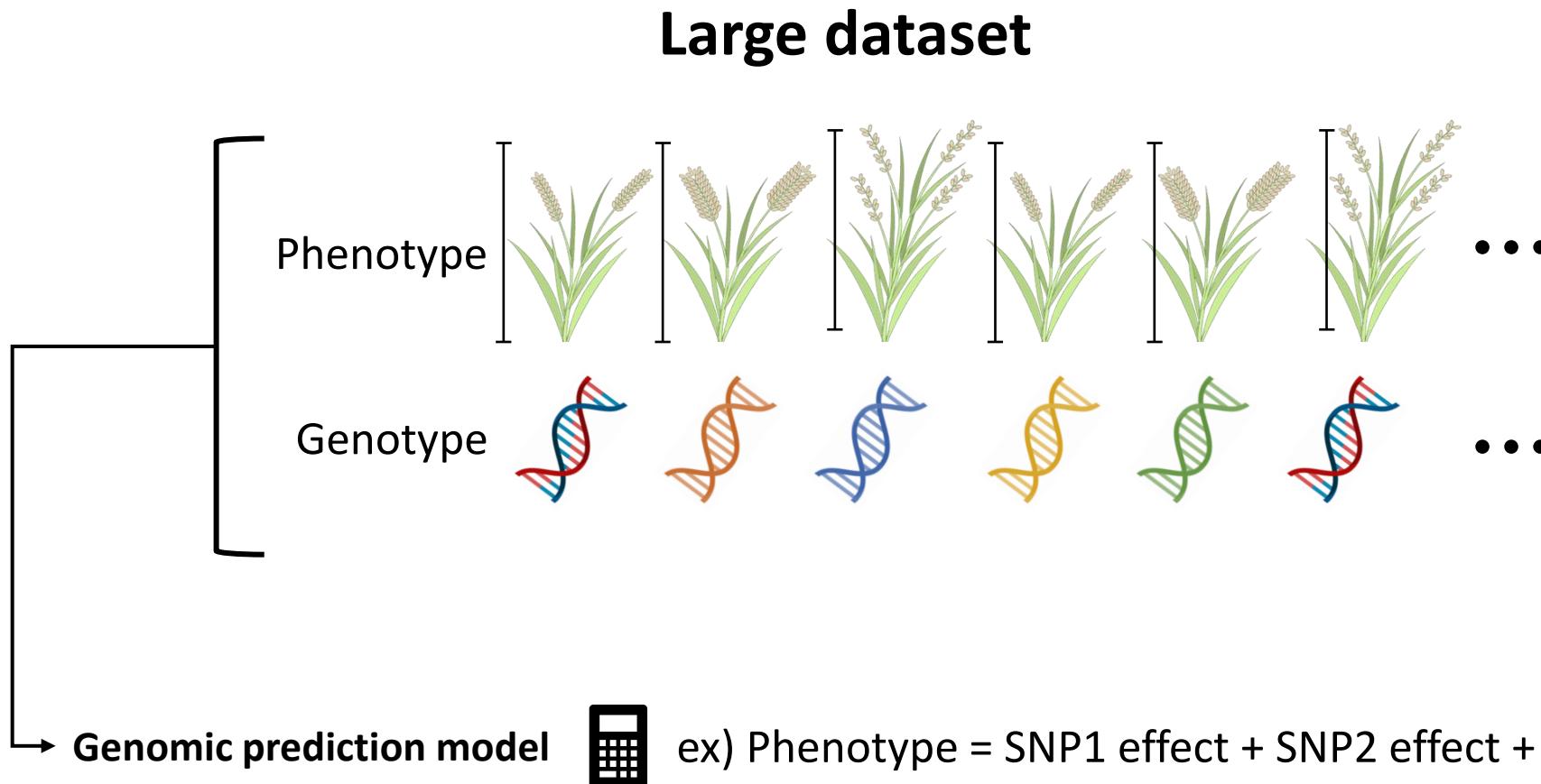
Very simple example: How to make genomic prediction model

	Phenotype	Genotype
Line1	150	
Line2	120	
Line3	130	



Genomic prediction model:
 $\text{Phenotype} = 100 + \text{SNP1} \times 20 + \text{SNP2} \times 30$

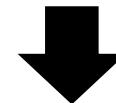
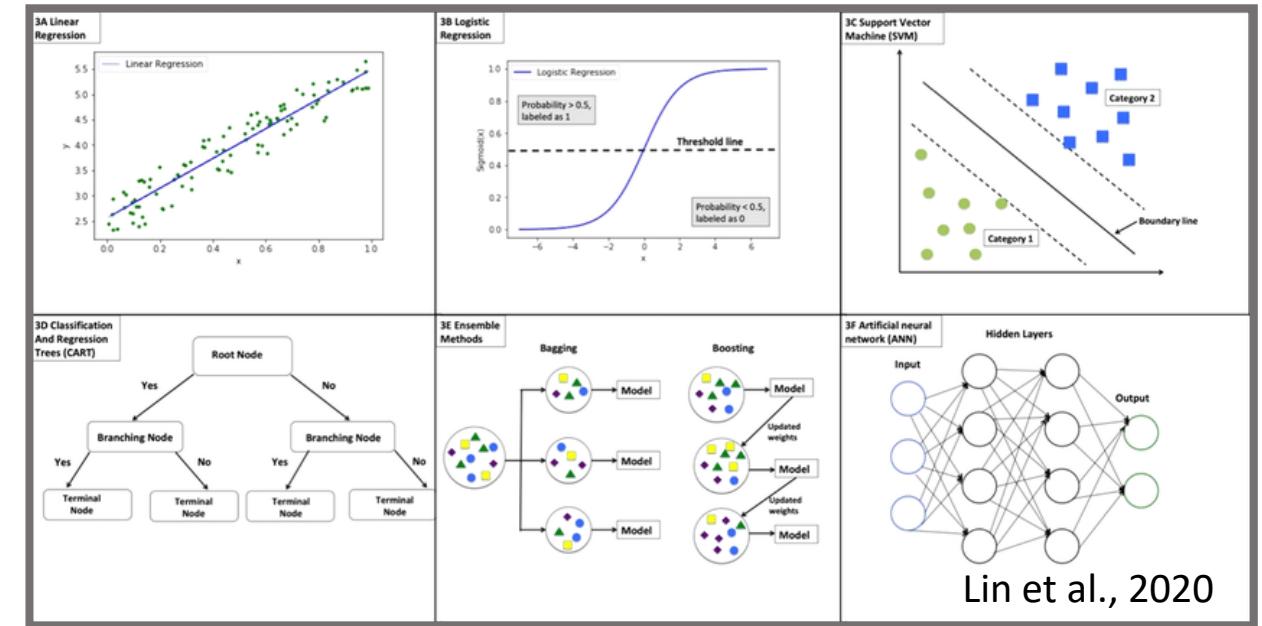
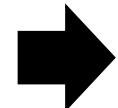
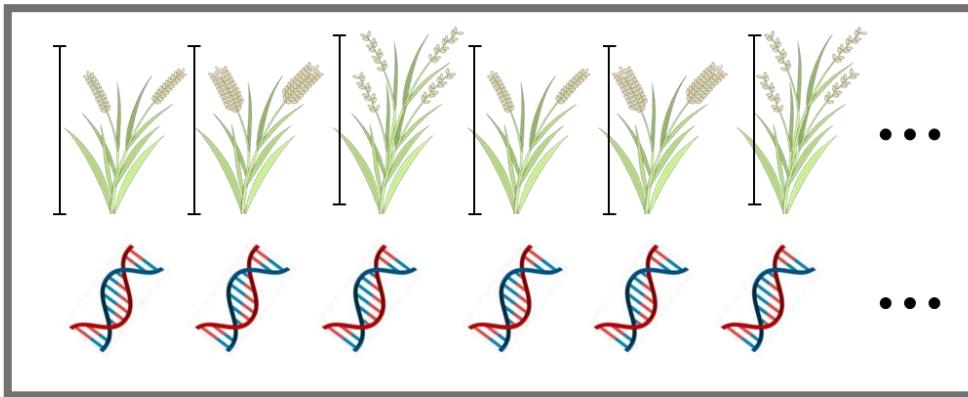
In reality, large dataset is required to make Genomic Prediction model



Recently, many statistical methods can be used

Learning the relationship between genotype & phenotype by statistical approaches

Large Phenotype & Genotype dataset

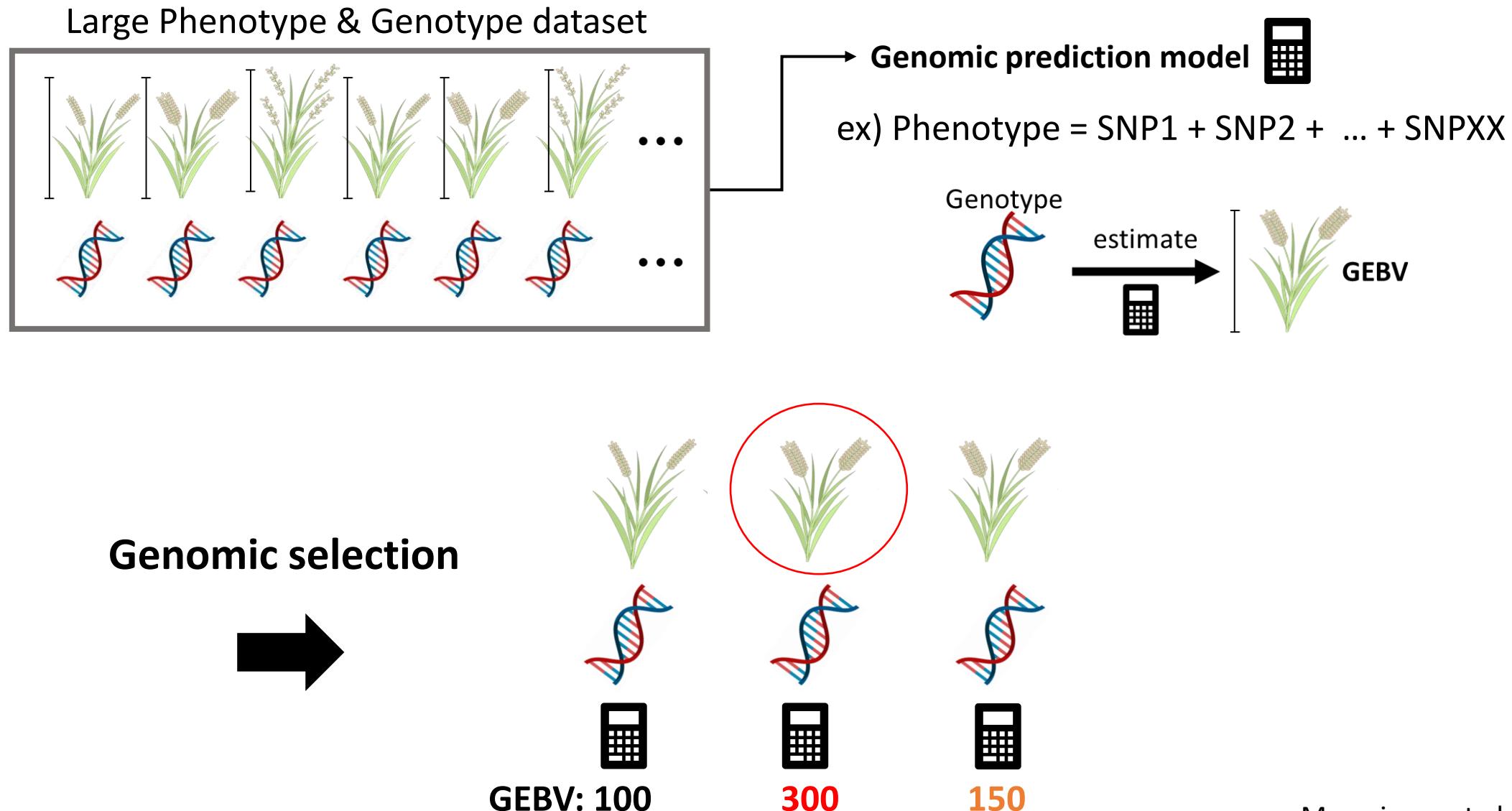


Genomic prediction model



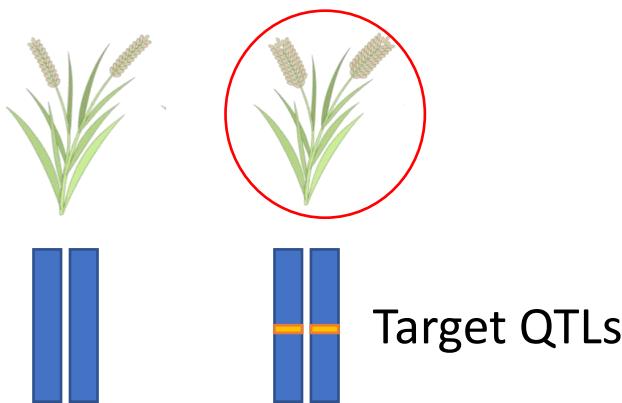
ex) Phenotype = SNP1 effect + SNP2 effect + ... + SNPx effect

We can select good lines based on GEBV estimated by GP model

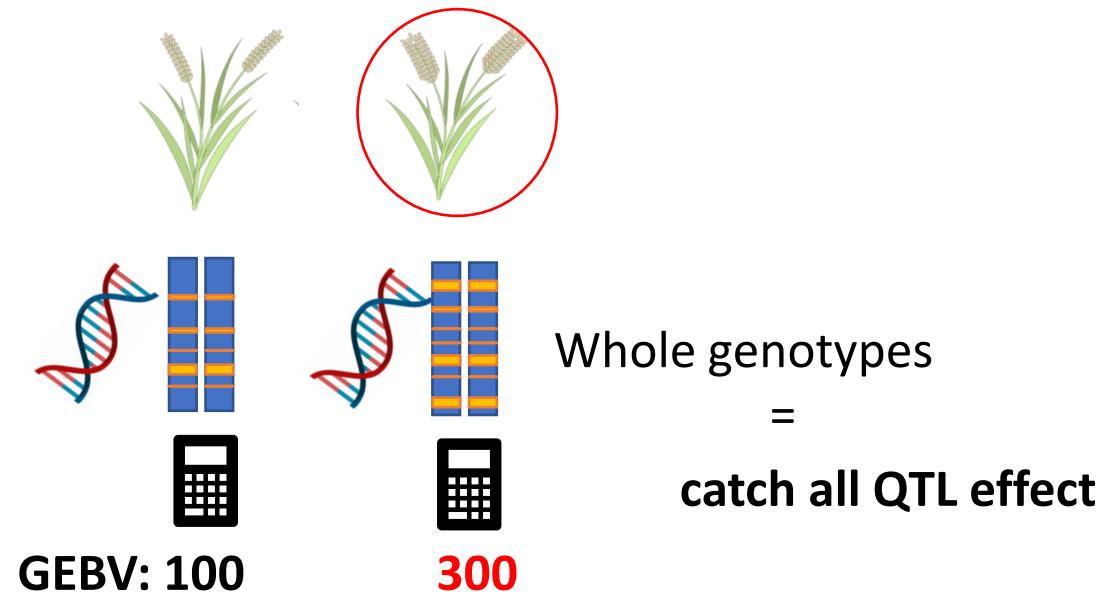


Genomic Selection (GS) can solve the limitations of MAS

Marker-assisted selection

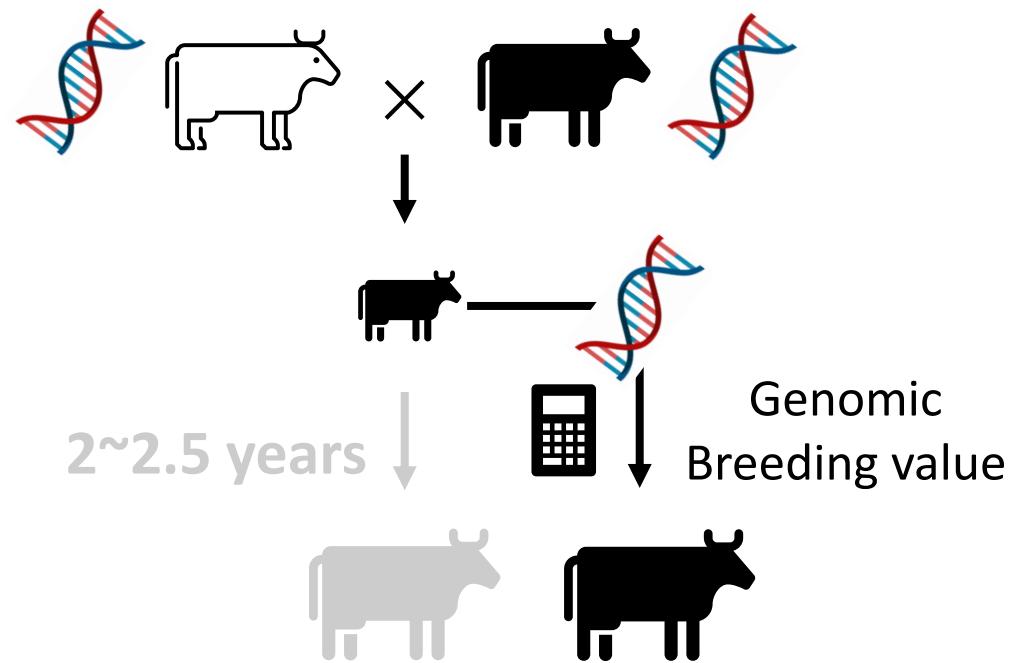


Genomic selection

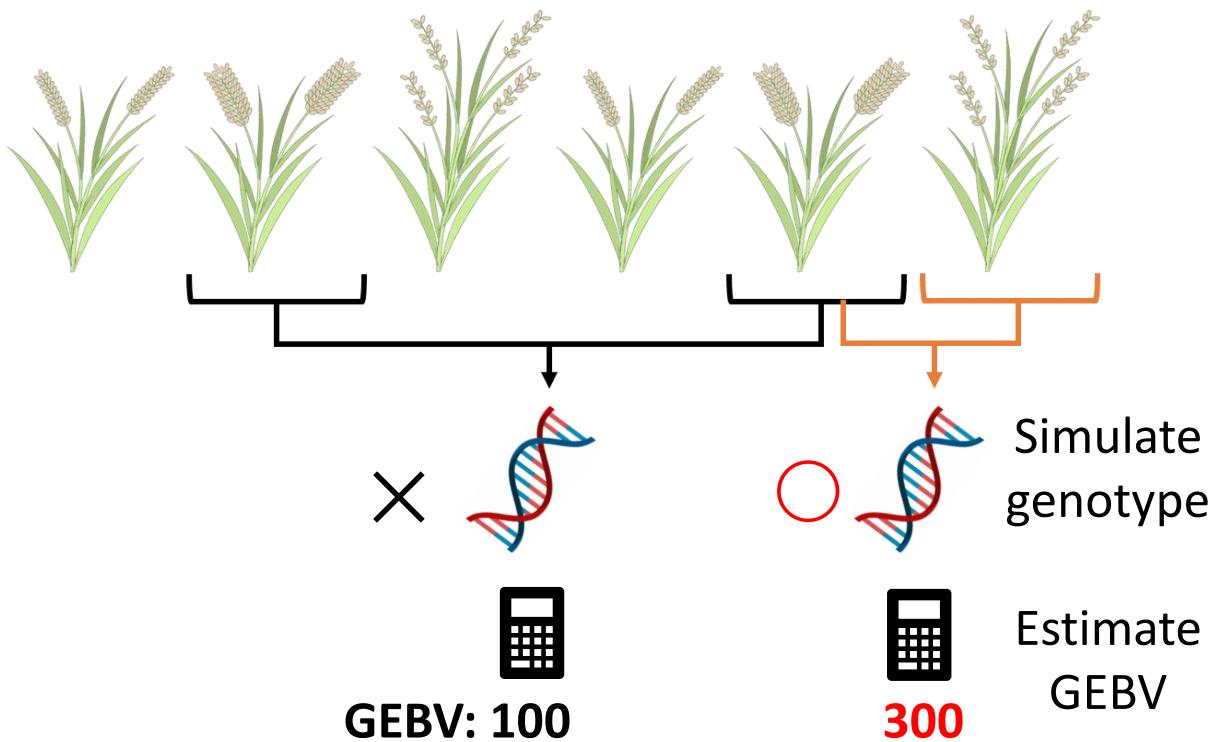


Genomic selection solved limitations of phenotypic selection

No need to wait until phenotyping



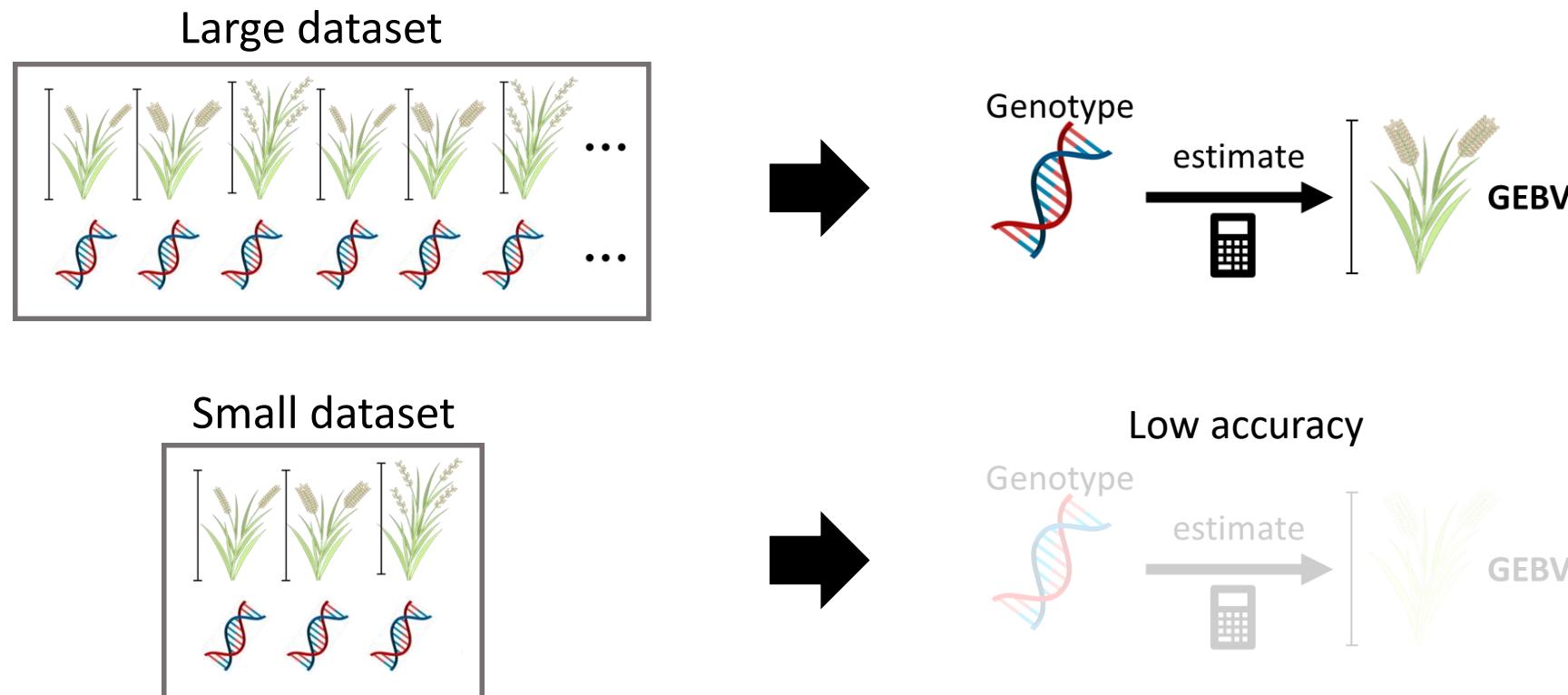
We can estimate phenotype values of next generation
→ Select best combination of parents



Current limitations of Genomic selection / Genomic prediction

- Large dataset is required

Wide variety of lines (Genotypes) is required to catch relationship between phenotype and genotype



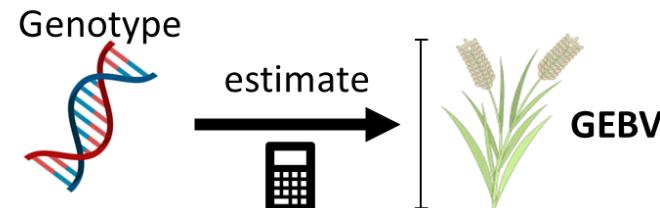
Current limitations of Genomic selection / Genomic prediction

- Most GS research focused on prediction accuracy of the model

We success to generate highly accurate genomic prediction model !! ➡ A lot of research

How to apply Genomic Selection to “Real” breeding program? ➡ few

Genomic prediction model



How to apply?

?

Real breeding



Current limitations of Genomic selection / Genomic prediction

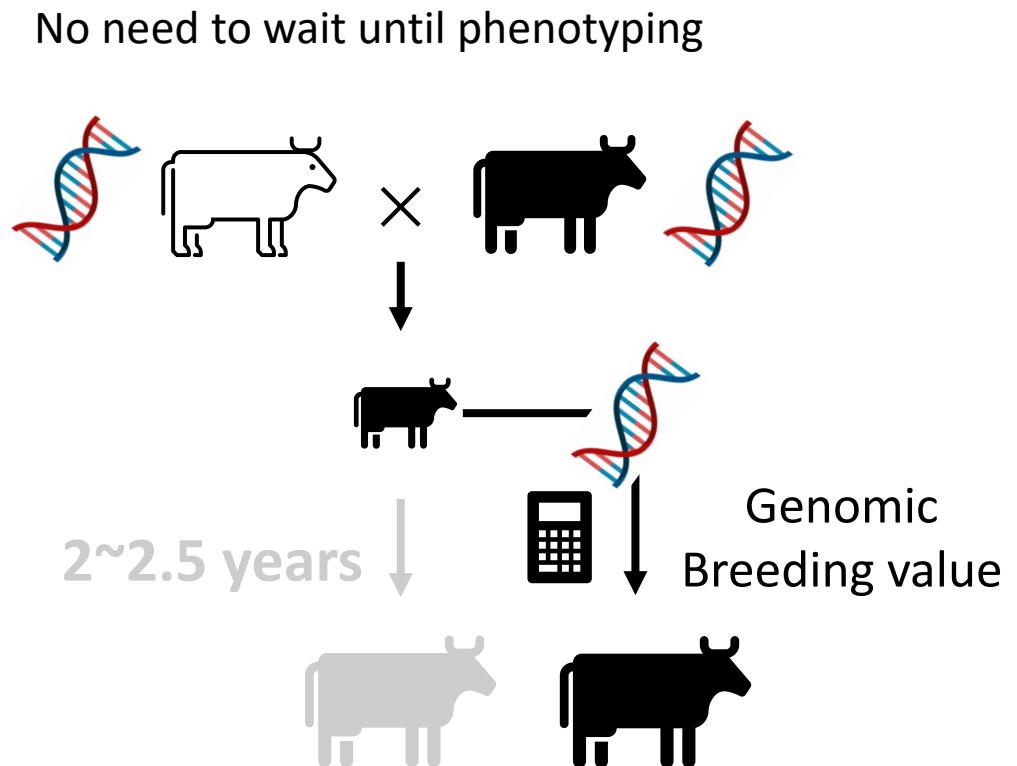
- Genomic Selection is emerged from **Animal breeding**

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

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

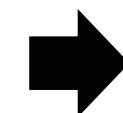
*Research Institute of Animal Science and Health, 8200 AB Lelystad, The Netherlands, [†]Victorian Institute of Animal Science, Attwood 3049, Victoria, Australia and [‡]Institute of Land and Food Resources, University of Melbourne, Parkville 3052, Victoria, Australia

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



c.f.) Genomic selection for plant breeding is different from animal breeding

	Animal	Plant
Cross-breeding	Take time	Easy, speed breeding
Population Size	Small	Large
Environmental effect	Small	Large
Mutation	Difficult to use	Easy to use
Preserve/Share	Difficult	Easy
Benefit of GS	High	Low



In Plant,

- Genomic prediction model with high accuracy
- Apply GP model various way with flexibility

Not only Genomic Selection

→ Plant specific GP strategy may be possible.

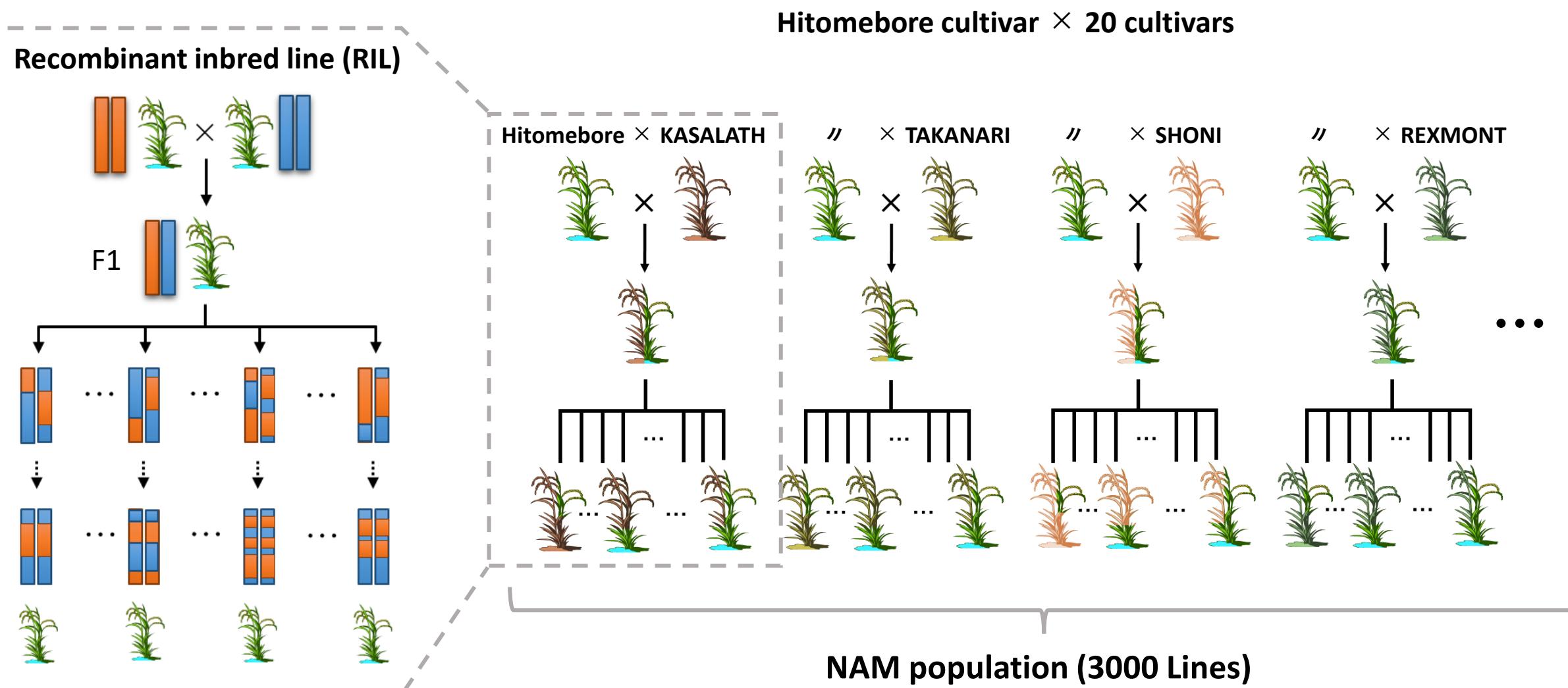


Example of the application of Genomic Prediction using Rice population



In IBRC

Material: NAM population of rice

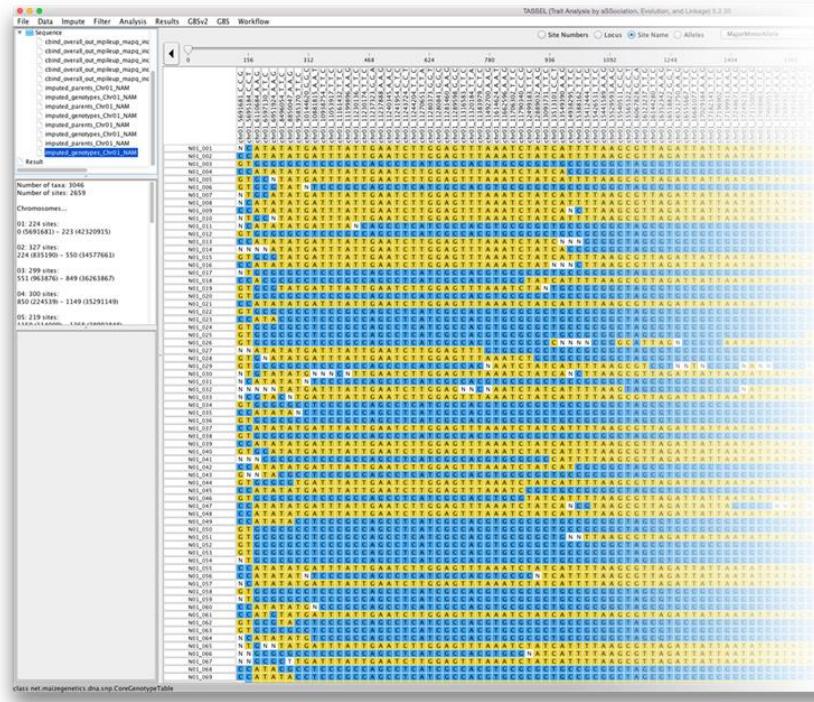


Material: NAM population of rice

- Genome sequencing & phenotyping for all NAM lines.

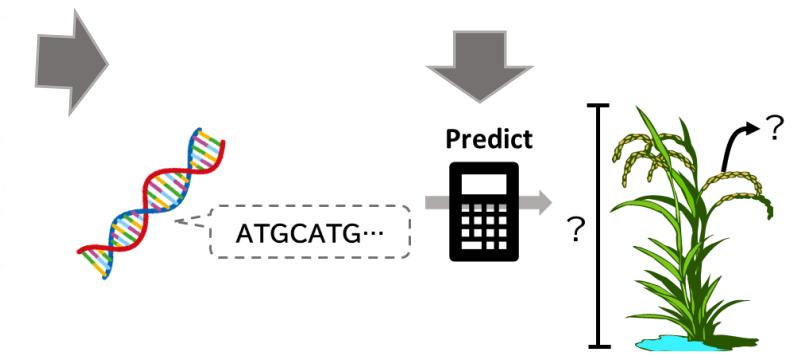
Then, we get...

SNP genotype in whole genome & Phenotype data of 3000 lines

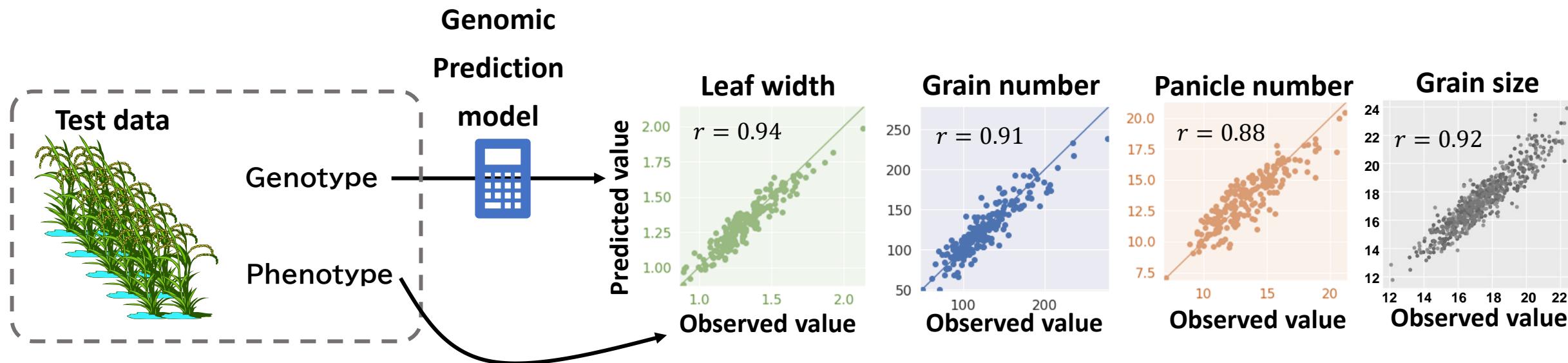


Line	LW_mean	GN_mean	GN_mean_ori	LL_mean	DTH_mean	PN
N01_001	1.275	142.8833334	142.8833334	35.2	113.8	12
N01_002	1.1375	116.0666667	116.0666667	32.975	107.6	13
N01_003	1.3125	118.15	118.15	30.75	113	11
N01_004	1.1125	154.45	154.45	37.475	123.2	
N01_005	1.5	128.8166667	128.8166667	29.725	115.8	
N01_006	1.175	100.3166667	100.3166667	27.075	114.6	14
N01_007	1	79.21666666	79.21666666	25.75	106	
N01_008	1.2125	159.1333334	159.1333334	33.675	105.8	
N01_009	1.15	122.8833333	122.8833333	33.1	108.6	11
N01_010	1.5375	134.4	134.4	34.625	116	13
N01_011	0.875	82.1666667	82.1666667	34.15	108	
N01_012	1.4125	153.7833334	153.7833334	37.175	108	
N01_013	0.85	103.6666667	103.6666667	25	104.2	
N01_014	1.1625	133.6	133.6	30.75	108.2	
N01_015	1.2875	162.6	162.6	24.2	118.2	12
N01_016	1.175	130.7	130.7	36.4	121	
N01_017	1.35	121.8333333	121.8333333	40.225	101.8	10
N01_018	1.0625	158.3	158.3	36.575	124.2	15
N01_019	1.2375	138.6166667	138.6166667	28.525	104	
N01_020	1.3625	108.4	108.4	34.925	109.4	13
N01_021	1.275	133.6333334	133.6333334	33.4	115.2	17
N01_022	1.375	117.15	117.15	31.875	117.2	
N01_023	1.3375	173.8	173.8	30.8	109.8	17
N01_024	1.4125	148.4666666	148.4666666	32.625	111.6	12
N01_025	1.25	118.9833333	118.9833333	31.425	108	12
N01_026	1.575	105.0166667	105.0166667	30.975	100.9	16

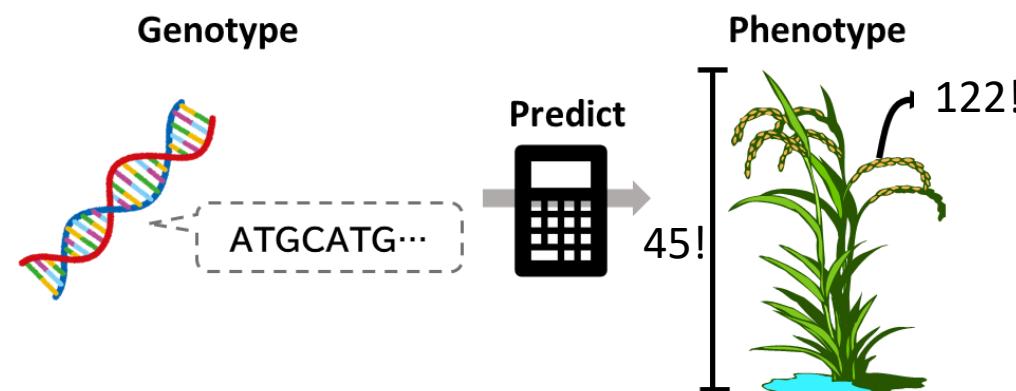
Generate genomic prediction model



We succeed to generate highly accurate genomic prediction model for 4 traits.



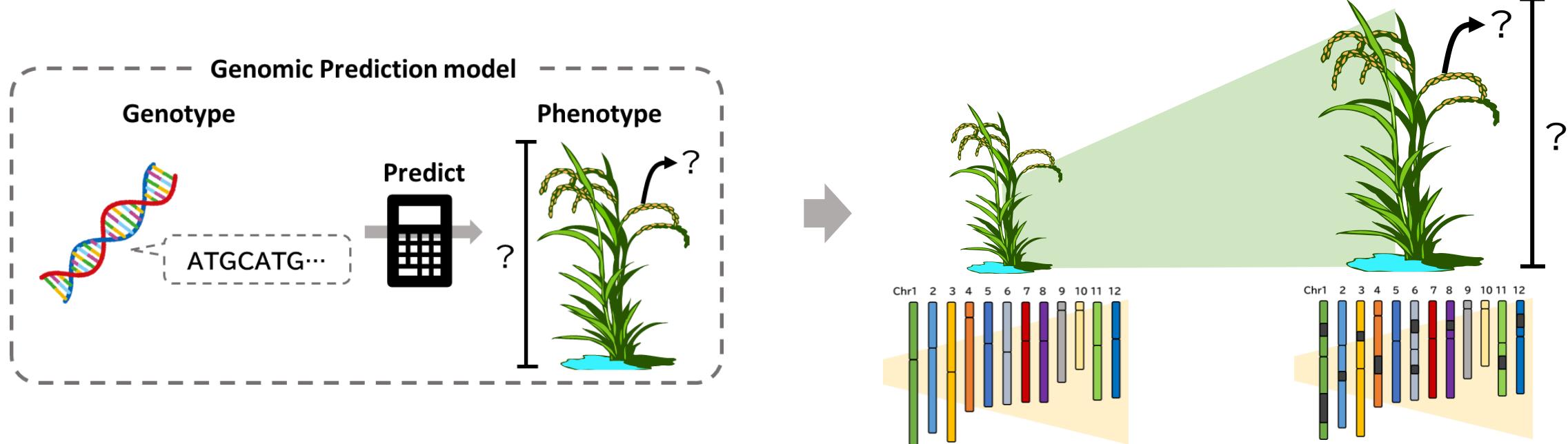
We can predict phenotype(LW, GN, PN, GS) from genotype information.



The example of application of genomic prediction model

Based on genomic prediction model, we can calculate

- How much can we change the traits by genomic breeding?
- What genotype is the ideal for traits?
- ..etc.

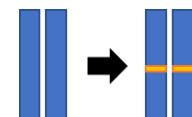


We can consider ideal genotype for trait by Genomic Prediction model

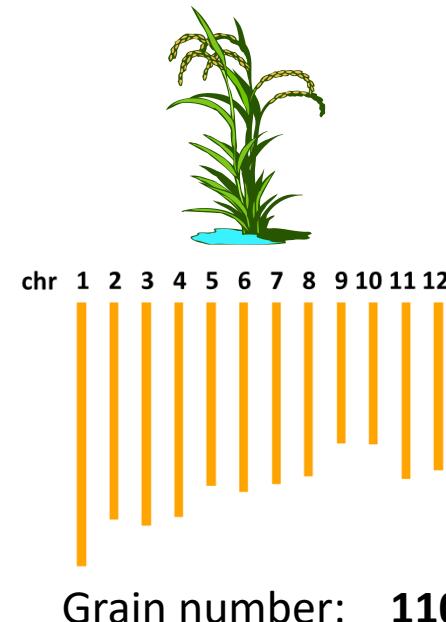
Good prediction model



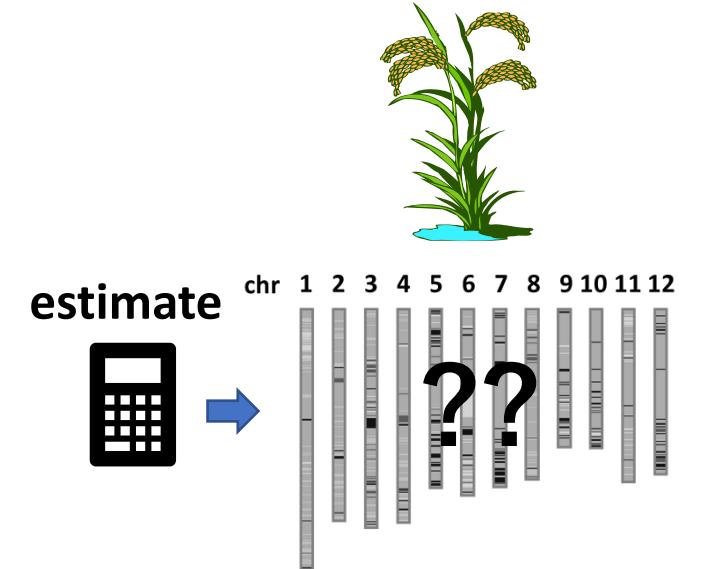
Genomic region	Genotype (from Hitomebore)	Grain number
chr01: 3~4Mbp	to KALUHEENATI	↑
chr03: 12~13Mbp	to KASALATH	↑↑↑
chr06: 8~9Mbp	to C8005	↓↓
...



Hitomebore

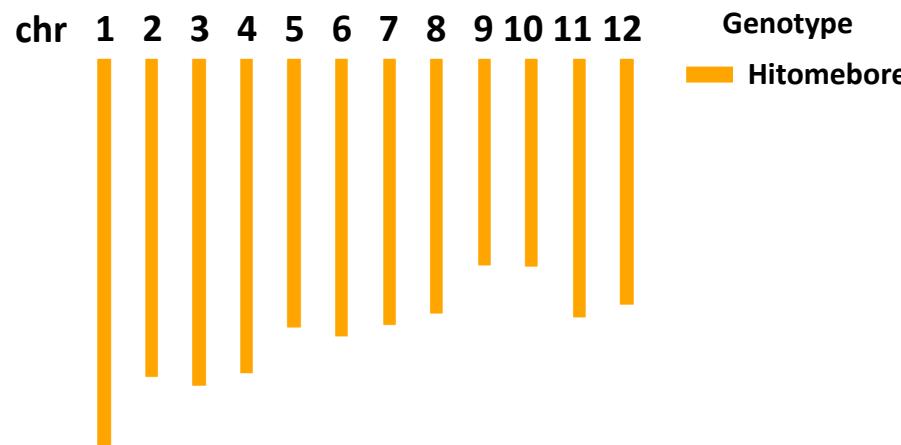


High grain number genotype

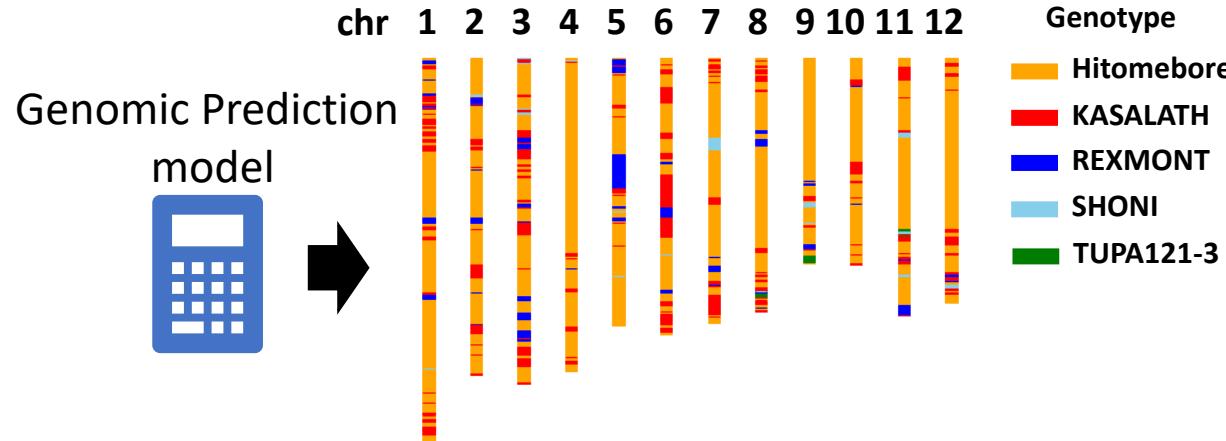


Predict ideal genotype for Grain number by the prediction model

- Hitomebore (reference cultivar)



Grain number per a panicle : 110~



Grain number per a panicle : 591~



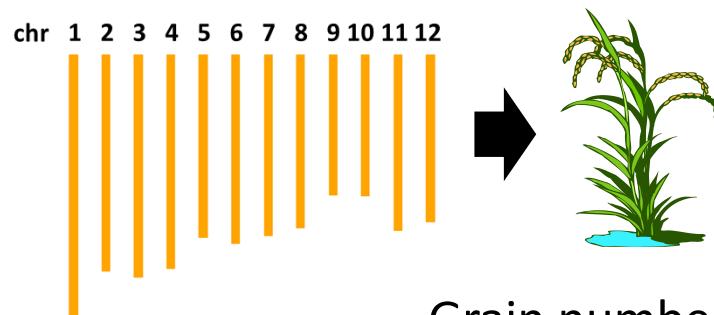
looks Very nice! ☺

Ideal genotype for Grain number has 2 problems.

- It's **Not** ideal genotype for other traits

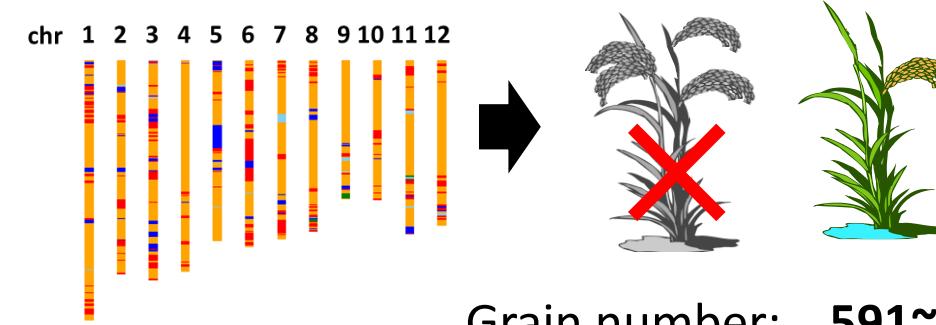
ex) Grain number ←→ Panicle number/Grain size

Hitomebore (reference cultivar)



Grain number: **110~**
Panicle number: **11~12**
Grain size: **16~**

Ideal genotype for Grain number
(**Not ideal for Panicle number/Grain size**)



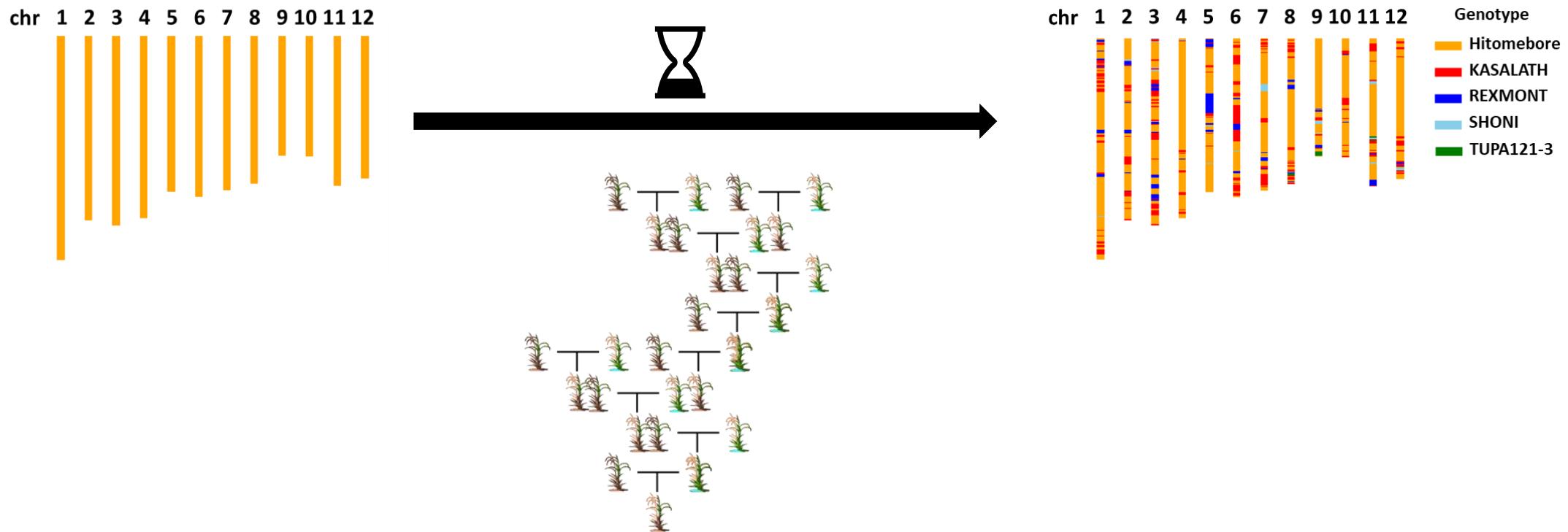
Grain number: **591~**
Panicle number: **1~2...**
Grain size: **10~**



Ideal genotype for Grain number has 2 problems.

- It's **Not** ideal genotype for breeding.

It take so much time to make this genotype... like several hundred years...?



We can also consider feasible genotype based on available materials



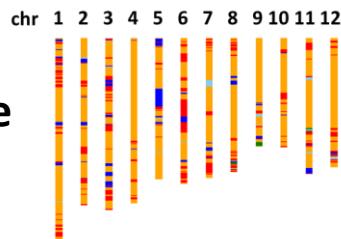
GP model



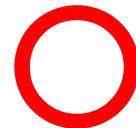
Genetic effect



Impossible genotype



↑↑↑ Grain number
↓↓↓ Panicle number
↓↓↓ Grain size



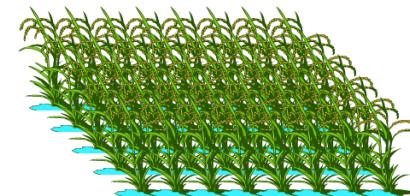
GP model



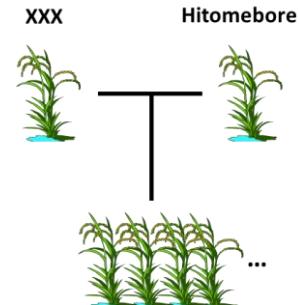
Genetic effect



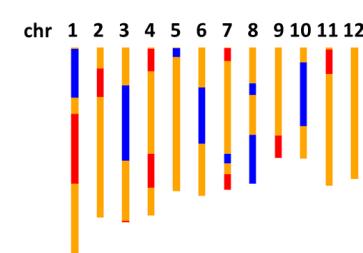
What material can we use?



How many generations are required?

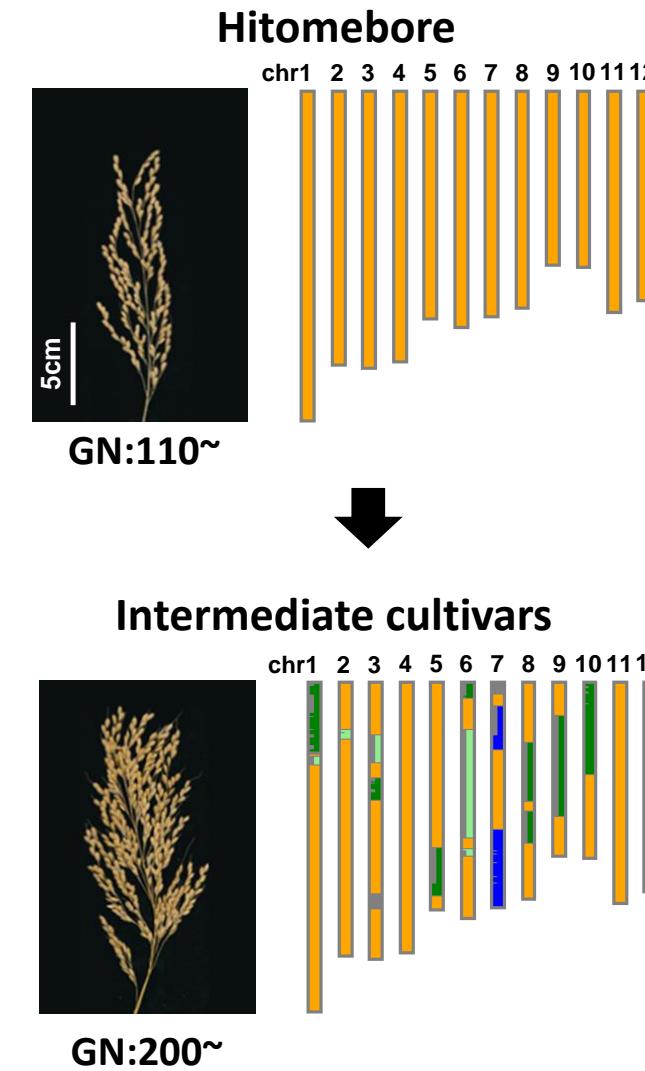
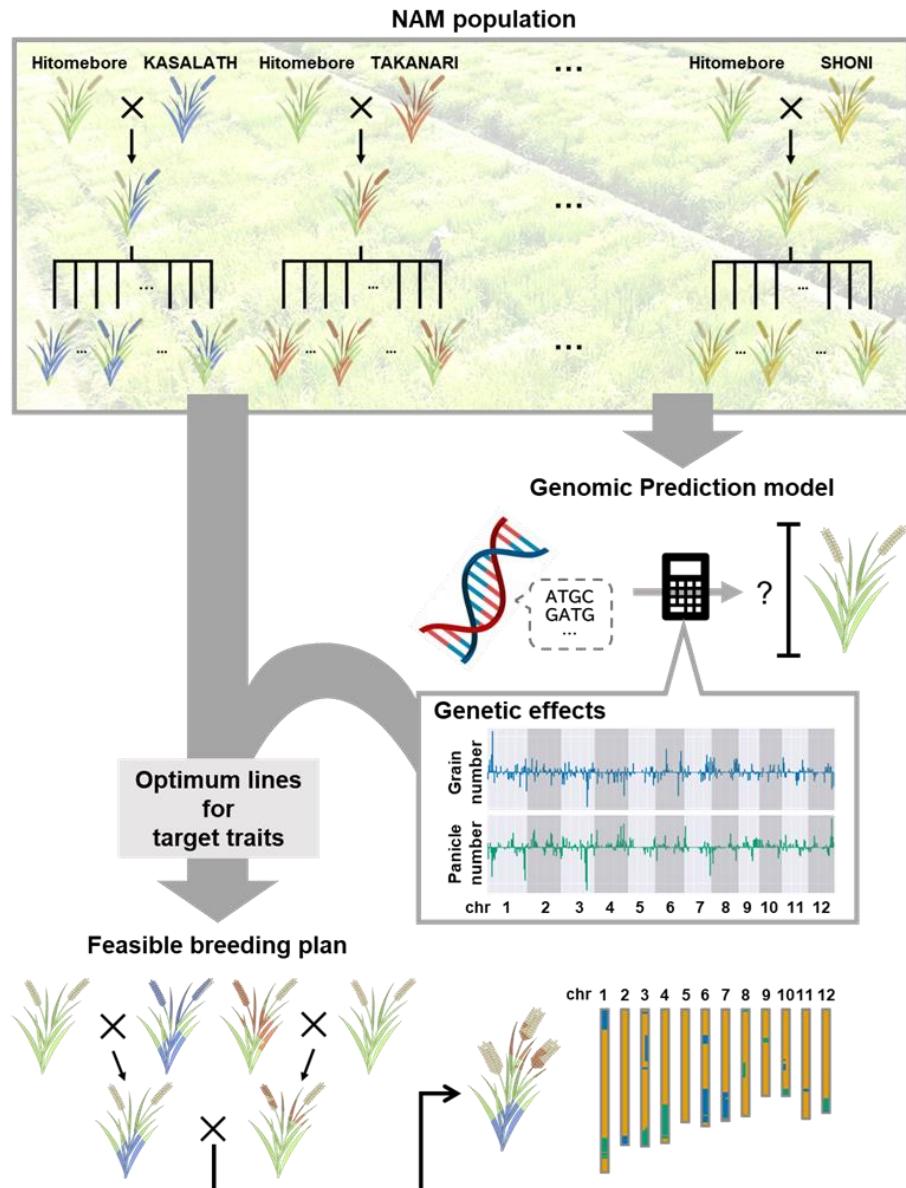


Feasible genotype



↑↑ { Grain number
Panicle number
Grain size

Now, we suggest a new breeding strategy based on Genomic Prediction model

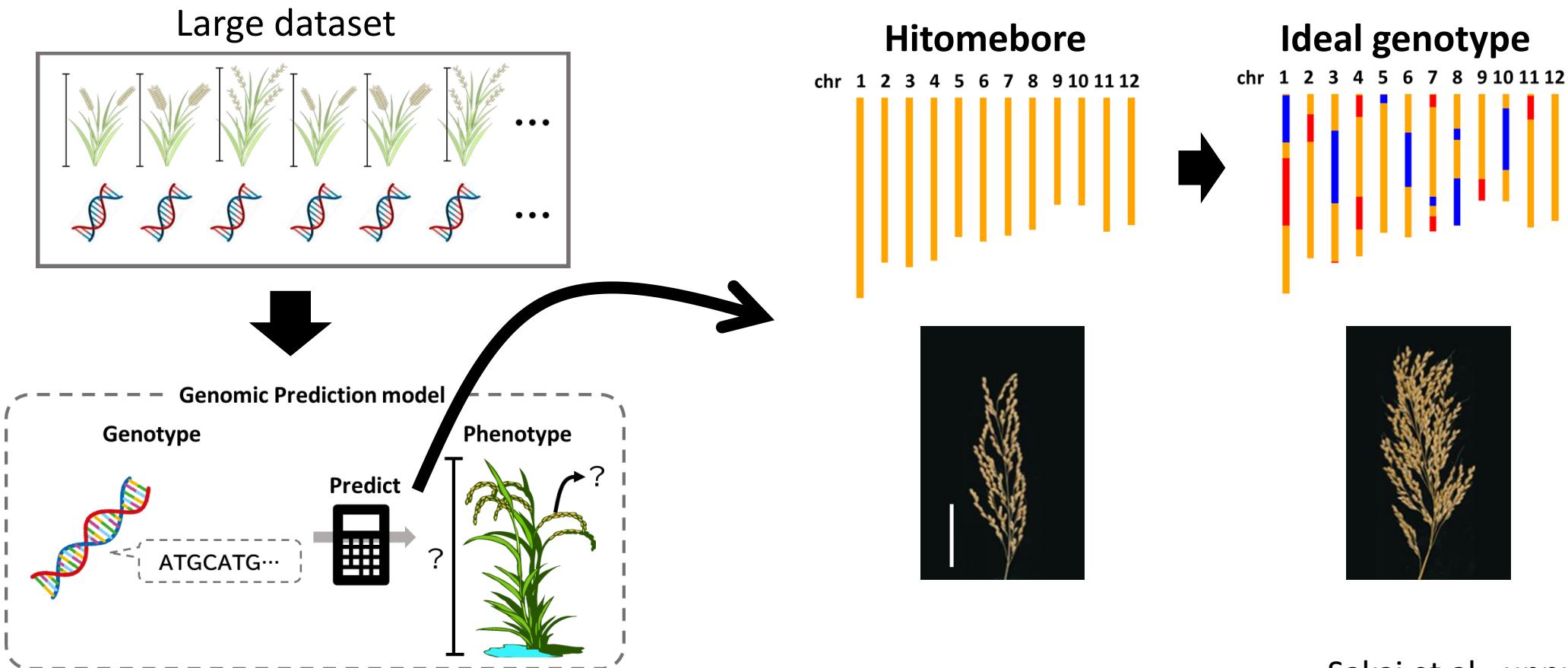


Sakai et al., unpublished

Summary of Genomic Prediction/Selection

Although huge cost & materials are required to make accurate genomic prediction model,

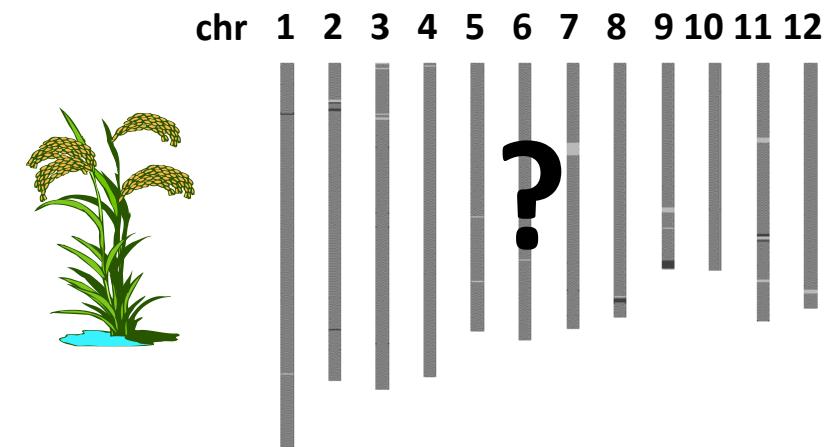
Genomic prediction model can suggest breeding strategy with considering whole genomic regions



Next - Practice part -

We will experience to make & use Genomic Prediction model.

We try to make good genotype for Grain number using our prediction model.
(It' not practice... just play with genomic prediction model that we generated.)

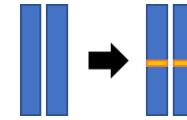


Future direction - include GxE interaction effect -

- In our study, we succeeded to generate good genomic prediction model.
- But we consider only genetic effect in the phenotypic variance.
- There are environmental effect & **gene x environmental (GxE) interaction effect.**

Phenotypic variance = **Genetic variance** + **Environmental variance**

Future direction - include GxE interaction effect -



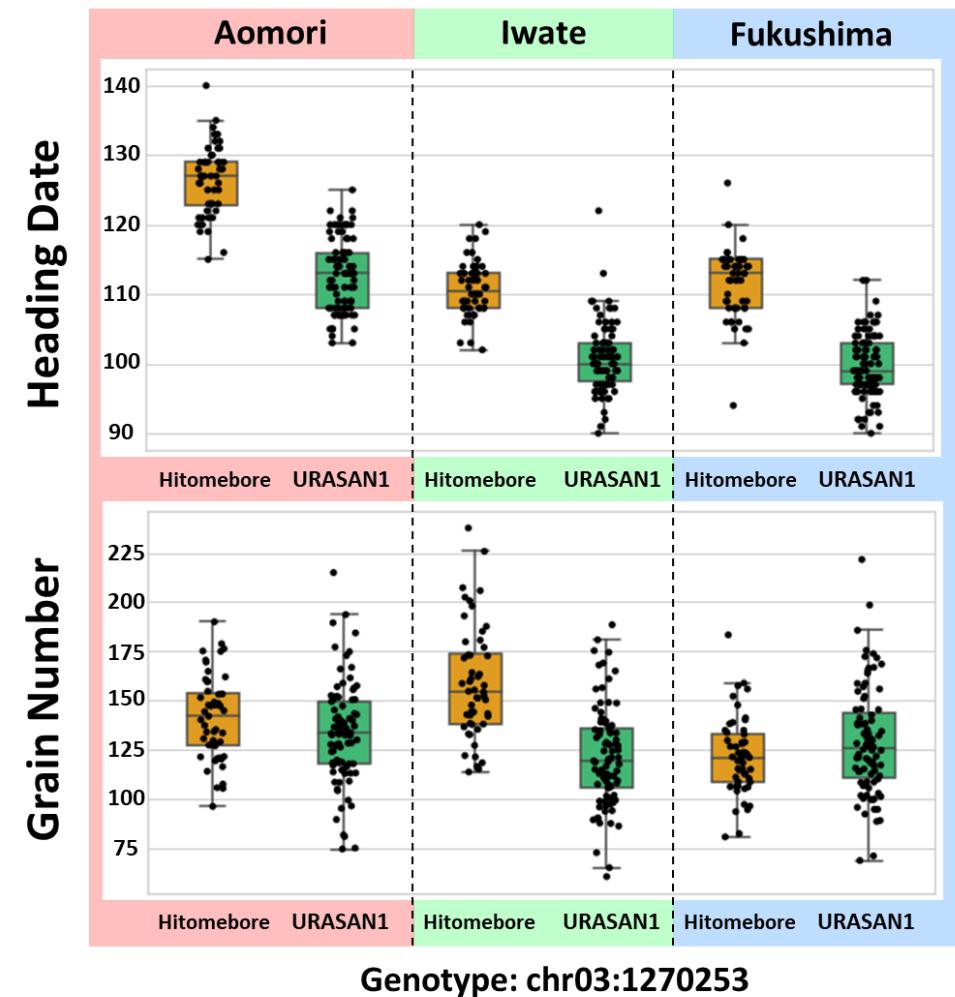
ex) We identified a locus showed GxE interaction effect in rice.

- The QTL affect Heading date in **3 locations**.
But the QTL affect Grain number only in **1 location**.
→ The gene effect on GN is different depending on the location.

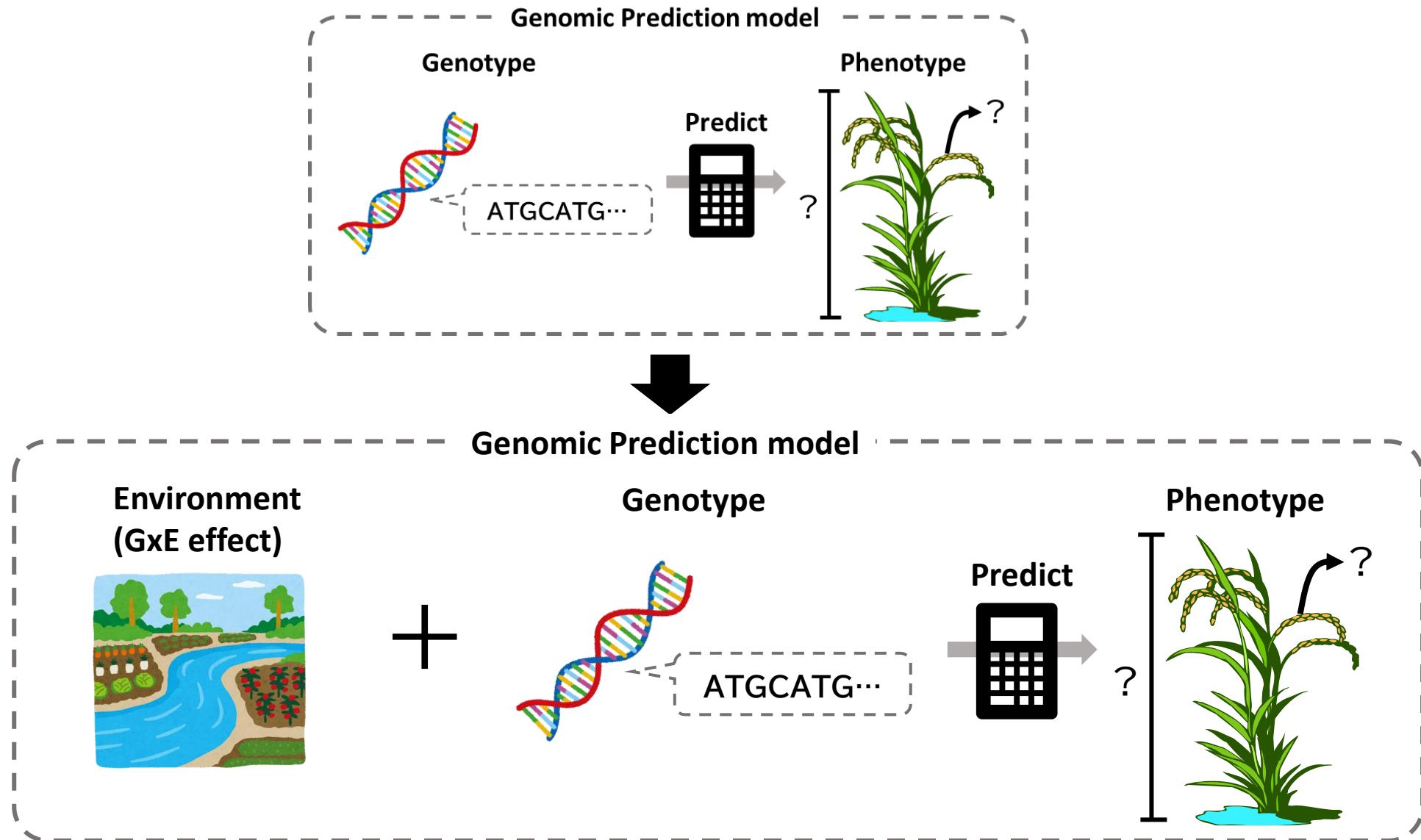


This result suggested

The ideal genotype may be changeable
depending on the environment



Future direction - include GxE interaction effect -



In the future...

Now, it still need some (high) cost to perform whole genome sequence-based methods.

But the sequencing cost is extremely decreasing. And the trend is still continuing.

So, if you have good materials like RILs,
NGS-based methods like MutMap/QTL-seq/GenomicPrediction might be useful in the future.

