Complex Genetic Architecture of a Growth Locus in the Chicken Genome

Jen-Hsiang Ou

Ph.D. Student

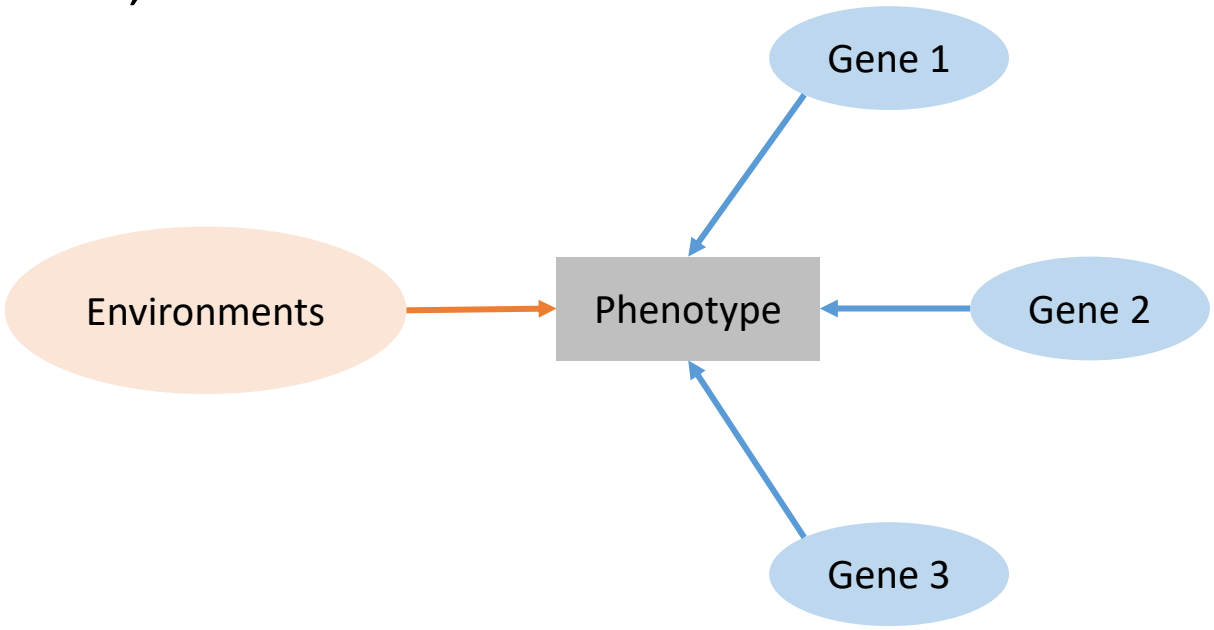
Department of Medical Biochemistry and Microbiology

Uppsala University

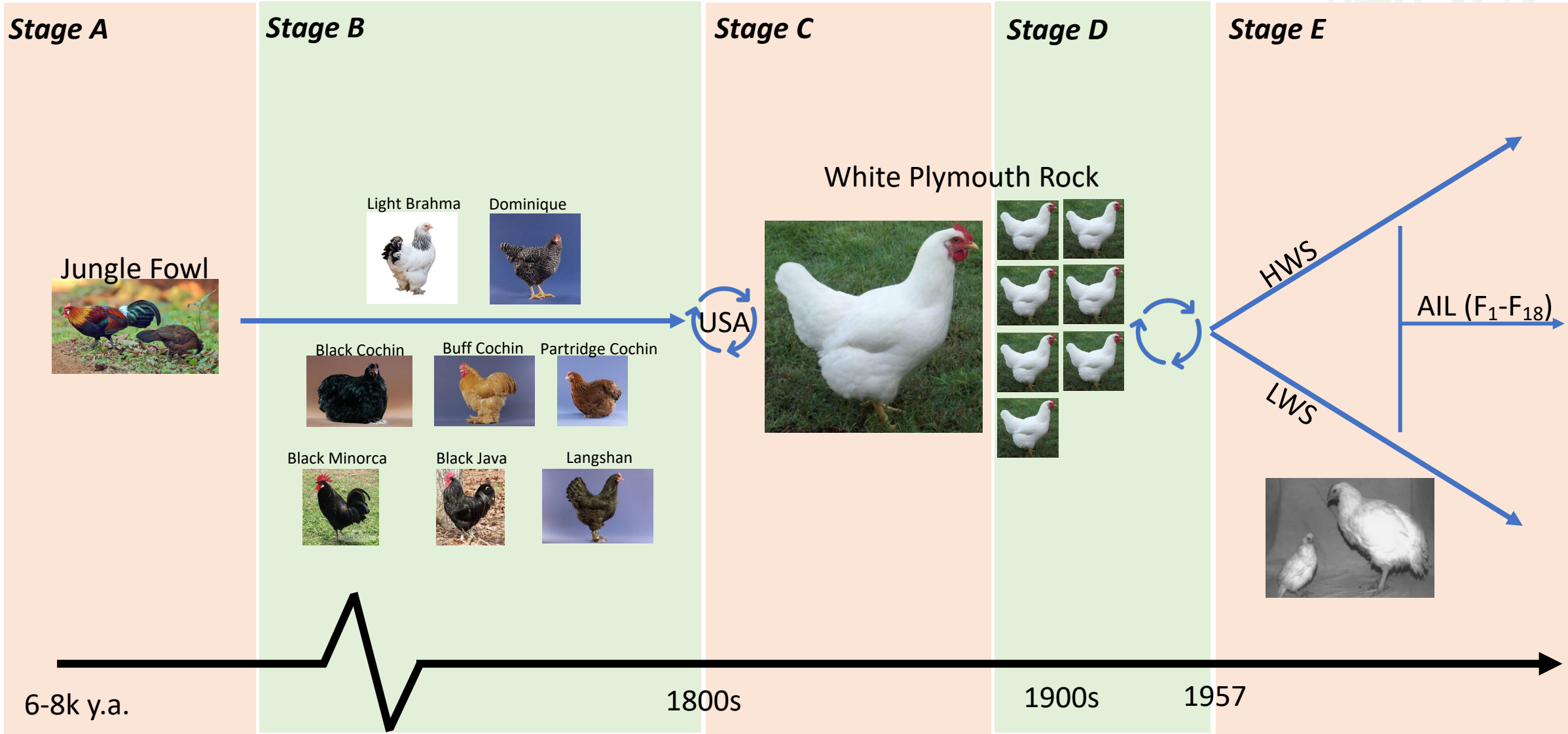
UPPSALA
UNIVERSITET

Complex traits are complicated

- Regulated by multiple factors
- Examples:
  - Cardiac disease, cancer, diabetes, Alzheimer's disease
  - Rice yield
  - Body weight
  - And lots more…

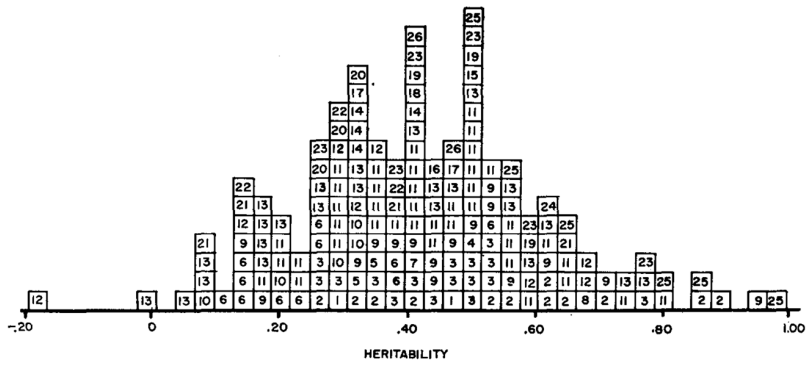


The Virginia Chicken Lines (breeding history)

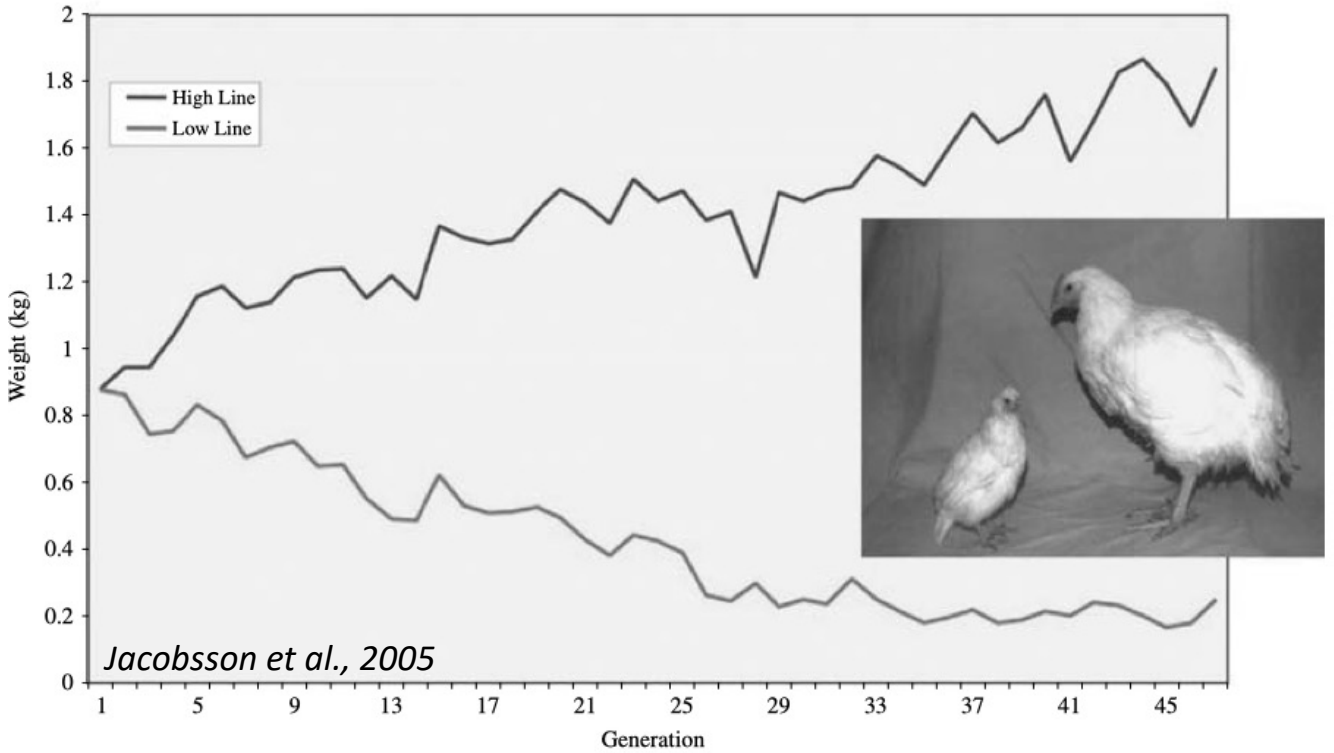


Bidirectional Selection of the Virginia Chicken lines

- Established in 1957 by Paul Siegel and co-workers.
  - "The weight at eight weeks of age is a complex trait influenced by multiple genetic factors." (Siegel, 1962)
- Body weight at eight weeks
  - A moderate to high heritability
  - Selection of a single trait
  - Complex trait



Published heritability of bodyweight w6-w12 (Siegel, 1962)



Advanced Intercross Line (AIL)

- F_2 intercross population → One generation of recombination
  - LD block: wide
  - Detect signals even when marker density is low
  - Don’t have a good resolution
- The advanced intercross line
  - Started from the 41st generation of HWS/LWS (F_0)
  - F_1 to F_{18}
- Sequencing and Genome Alignment
  - F_0 and F_1: high coverage sequencing (~30x in F_0 and ~5x in F_1)
  - F_2 – F_{18}: about 0.4x coverage

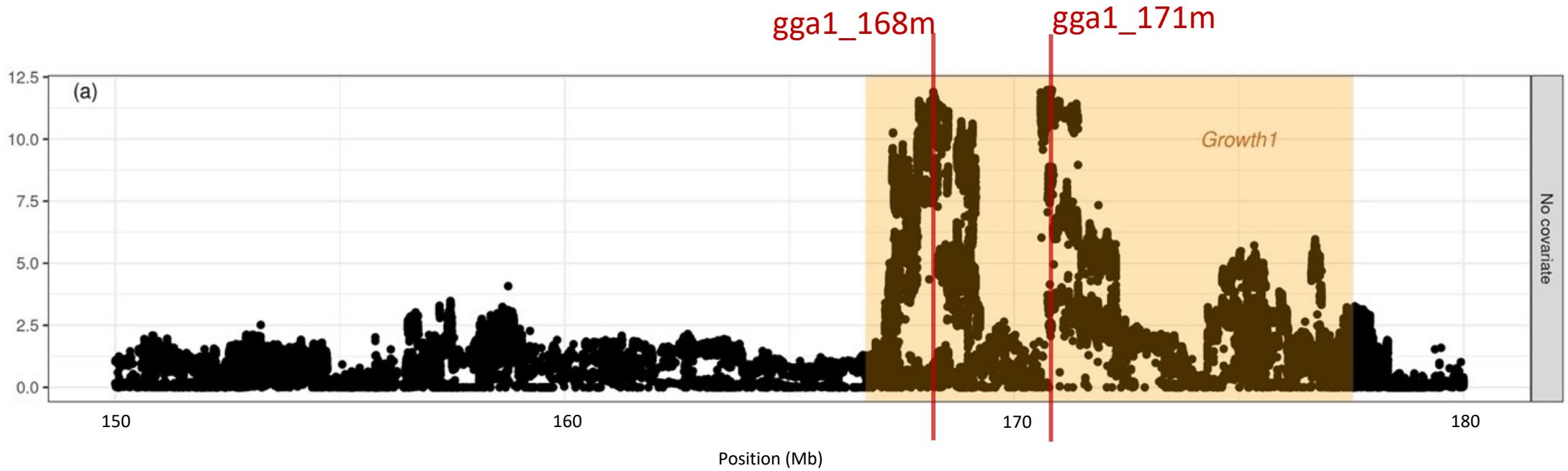
Imputation by AlphaFamImpute

- AlphaFamImpute (Whalen et al., 2020)
- Adding pedigree information to impute genotypes from low-coverage whole-genome sequencing data
- Algorithm of the software
  - Phases and impute parental genotypes
  - The well-classified parental haplotypes are passed to their offspring
  - All segregation states in offspring are projected to parental haplotypes at each locus → to determine the parental haplotype.
  - Offspring data are updated by the new parental genotypes.

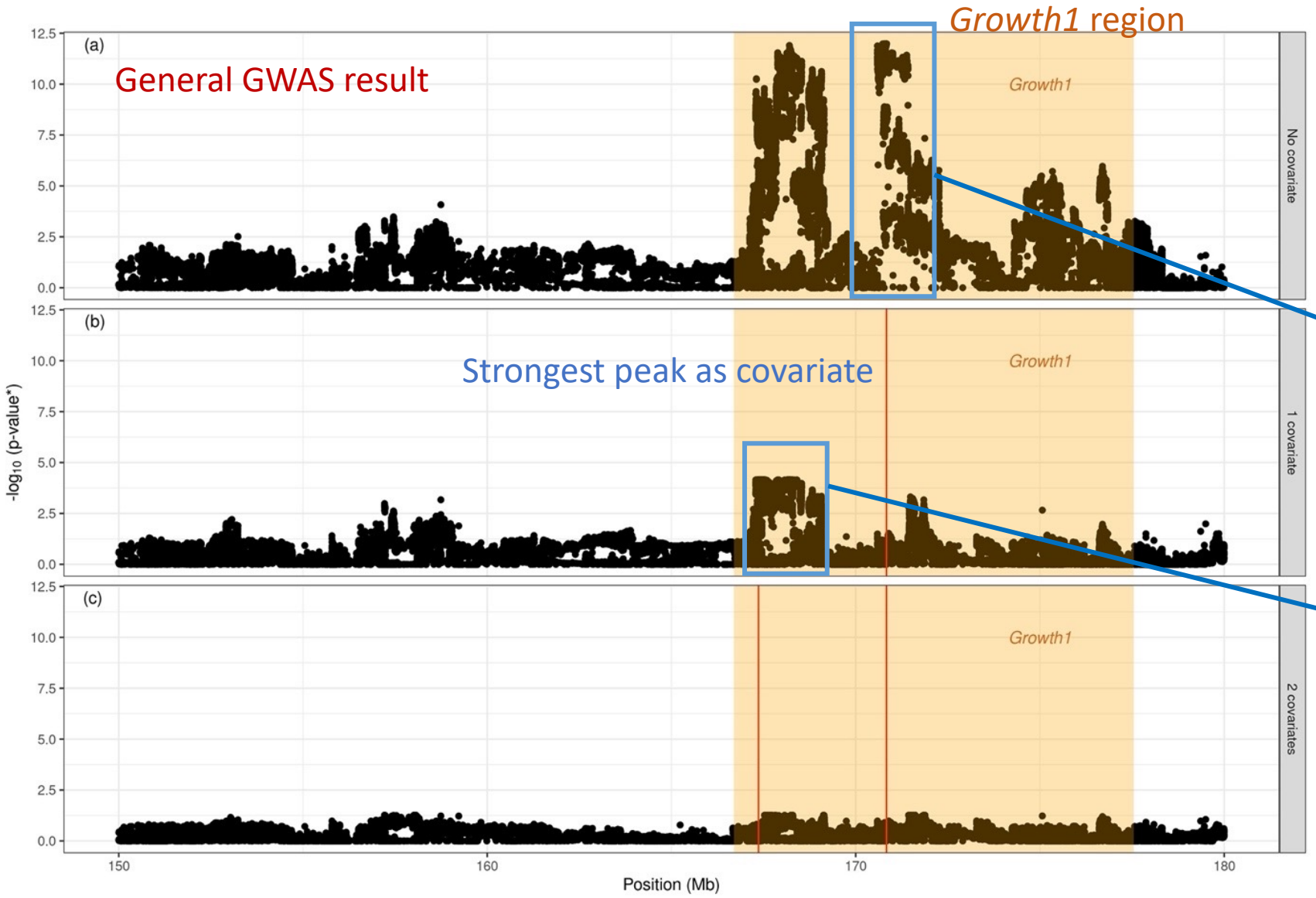
Genome-wide association study

$$y = \mathbf{1}\mu + S\beta_S + g\beta_g + A_j\, a_j + \epsilon$$

- Single-marker association study
- Generation and sex are considered as a fixed effect in the model



Genome-wide association study



$Y = Gen + Sex + \underline{G} + e$

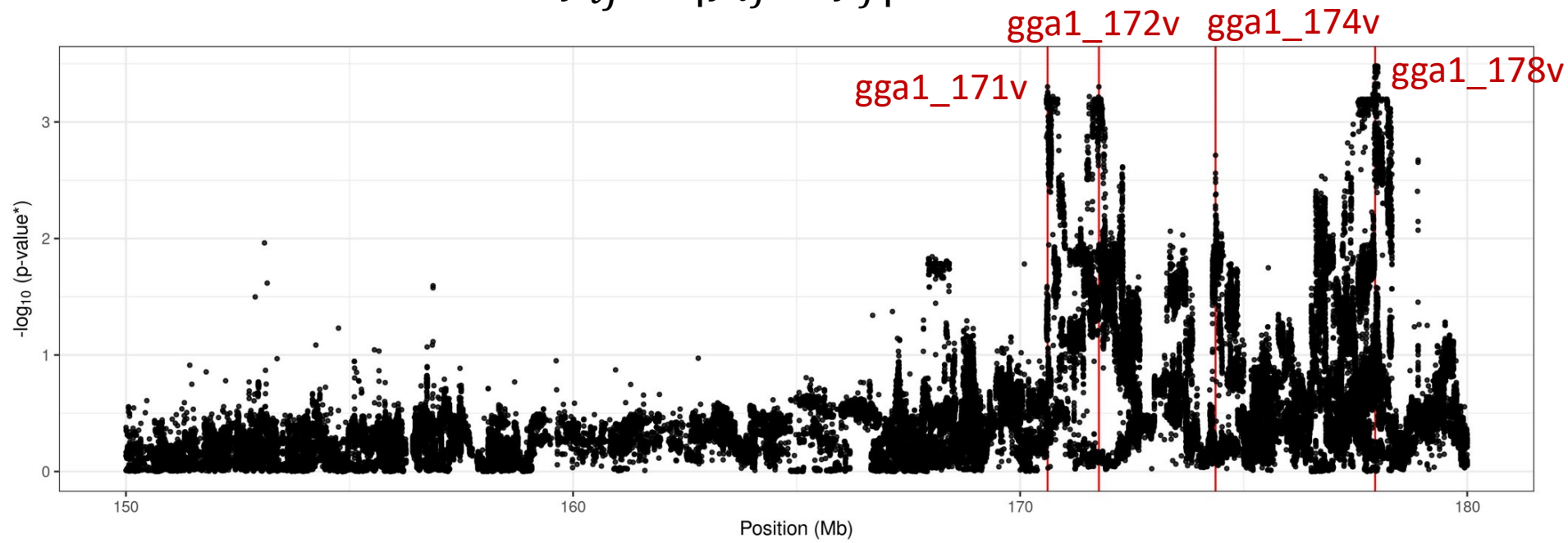
$Y = Gen + Sex + \boxed{S_{170}} + \underline{G} + e$

$Y = Gen + Sex + S_{170} + \boxed{S_{168}} + \underline{G} + e$

Variance-heterogeneity GWAS (vGWAS)

- Interactions and haplotypes (Forsberg et al., 2015)
- Brown-Forsythe test
  - Statistical test for the equality group variance based on performing an ANOVA
  - Transformation responsive variable: to measure the spread in each group

$$y_{ij}^{*} = \left| y_{ij} - \tilde{y}_{j} \right|$$



Hypothesis

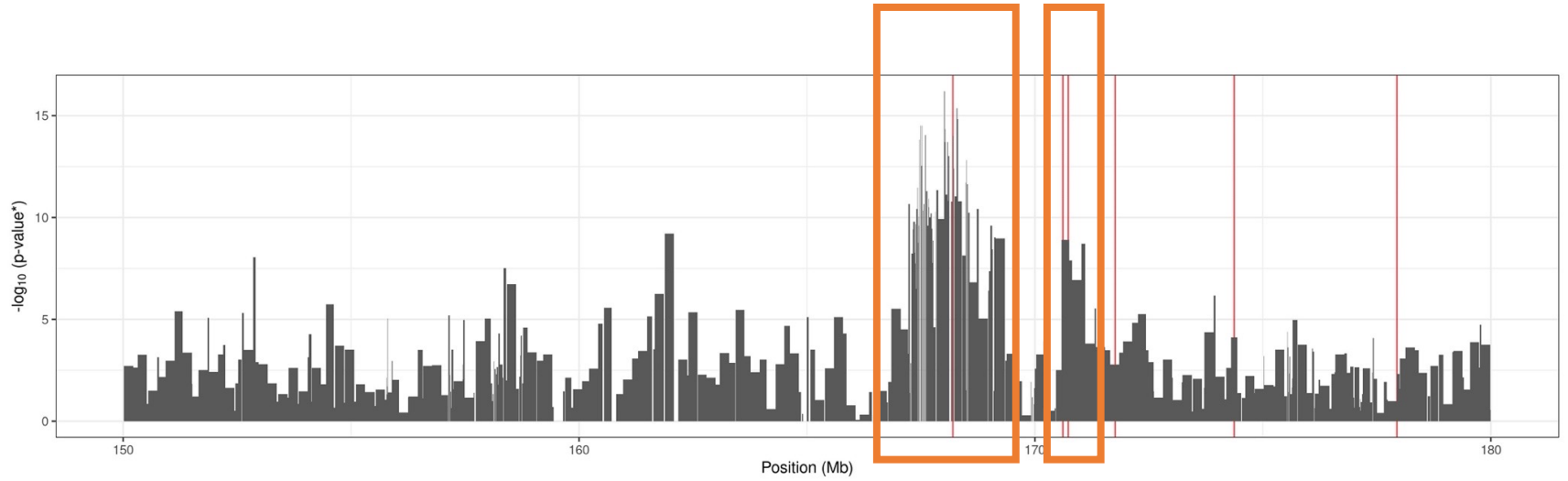
Potential explanations for the complex architecture of this region:

- Haplotype effects
  - ⇒ LD between the functional alleles that are not captured by individual SNP markers
  - ⇒ While we add the right peak as a covariate, the left peak dropped
- Interaction effects
  - ⇒ Interaction between the loci results in nonadditive gene variance
  - ⇒ Not well explained by the additive model

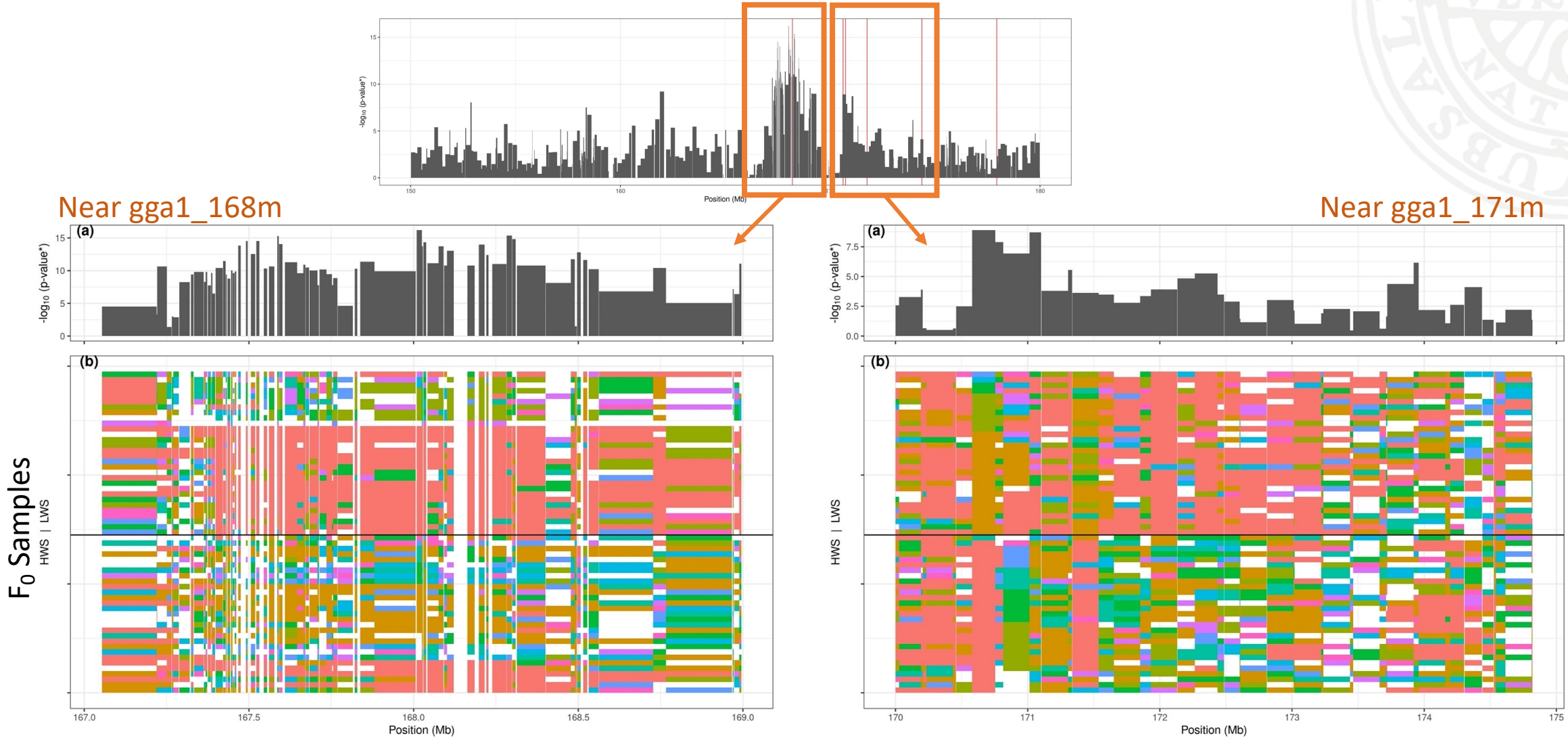
Haplotype-based association study

$$y = \mathbf{1}\mu + S\beta_S + g\beta_g + H_j h_j + \epsilon$$

- Sex and generation added are as fixed effects



Haplotype effect (167-169Mb)



Lillie et al., 2018: “LWS55 samples were fixed for one LWS haplotype while the HWS55 samples carried multiple haplotype which are different from LWS samples.”

The independence marker effects

- A stepwise selection across determined SNP markers
  - Both forward selection and backward elimination gave the same suggestion
  - Sex and generation are fixed effect

	Mean Square	F value	p-value
Sex	1.47×10^7	689.47	< 0.001***
Generation	1.32×10^6	61.72	< 0.001***
gga1_168m	1.51×10^6	70.80	< 0.001***
gga1_171v	6.34×10^5	29.65	< 0.001***
gga1_178v	5.68×10^4	2.66	0.1032
Significant codes: 0 '*' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 " 1**			

The independence marker effects

Target	Genotype	Average	STD	Count	LSD
gga1_168m	RR	0.1194	0.9893	1691	a
	RA	-0.1183	0.0961	1075	b
	AA	-0.4452	0.8843	164	c
gga1_171m	RR	0.1979	1.0384	692	a
	RA	0.0010	1.0069	1552	b
	AA	-0.2019	0.8765	686	c
gga1_171v	RR	-0.1751	0.9158	989	a
	RA	0.0375	1.0125	1538	b
	AA	0.2865	1.0331	403	c
gga1_172v	RR	-0.1306	0.8909	705	a
	RA	0.0004	1.0117	1574	b
	AA	0.1405	1.0418	651	c
gga1_174v	RR	0.0770	1.0208	939	a
	RA	-0.0050	1.0139	1509	b
	AA	-0.1343	0.8617	482	c
gga1_178v	RR	0.0638	1.0719	1071	a
	RA	-0.0053	0.9631	1426	a
	AA	-0.1404	0.8803	433	b

The NOIA model

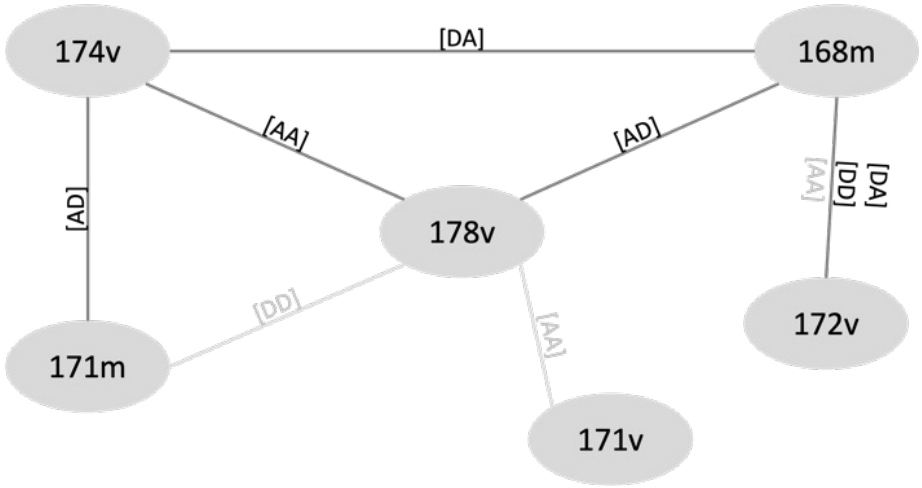
- Natural and Orthogonal InterAction model
  - A statistical framework aiming at unifying, extending, and simplifying existing models of genetic effects
- Practical properties of orthogonality
  - Statistically independent (uncorrelated) genetic effects
  - It leads to a proper decomposition of genetic variance
    - The Sum of the variance component (Var(A), Var(D)...) is exactly equal to the explained genetic variance
  - Make it easier and possible to remove effects without affection others

Interaction network (NOIA)

- All additive effects, dominance effects, and 2nd order interactions
- Sex and generation effects removed

gga1_168m	gga1_171m	gga1_171v	gga1_172v	gga1_174v	gga1_178v	Estimate	Std. Error	p-value
A						-0.245	0.039	0.000
					A	-0.093	0.034	0.006
D					A	0.218	0.082	0.008
D			A			0.313	0.132	0.018
D			D			-0.311	0.141	0.028
					D	0.099	0.045	0.028
	A			D		-0.275	0.132	0.037
A				D		-0.152	0.074	0.040
				D		0.119	0.059	0.046
				A	A	0.116	0.059	0.048
	D				D	-0.217	0.120	0.070
A			A			-0.145	0.080	0.071
		A			A	0.127	0.074	0.088

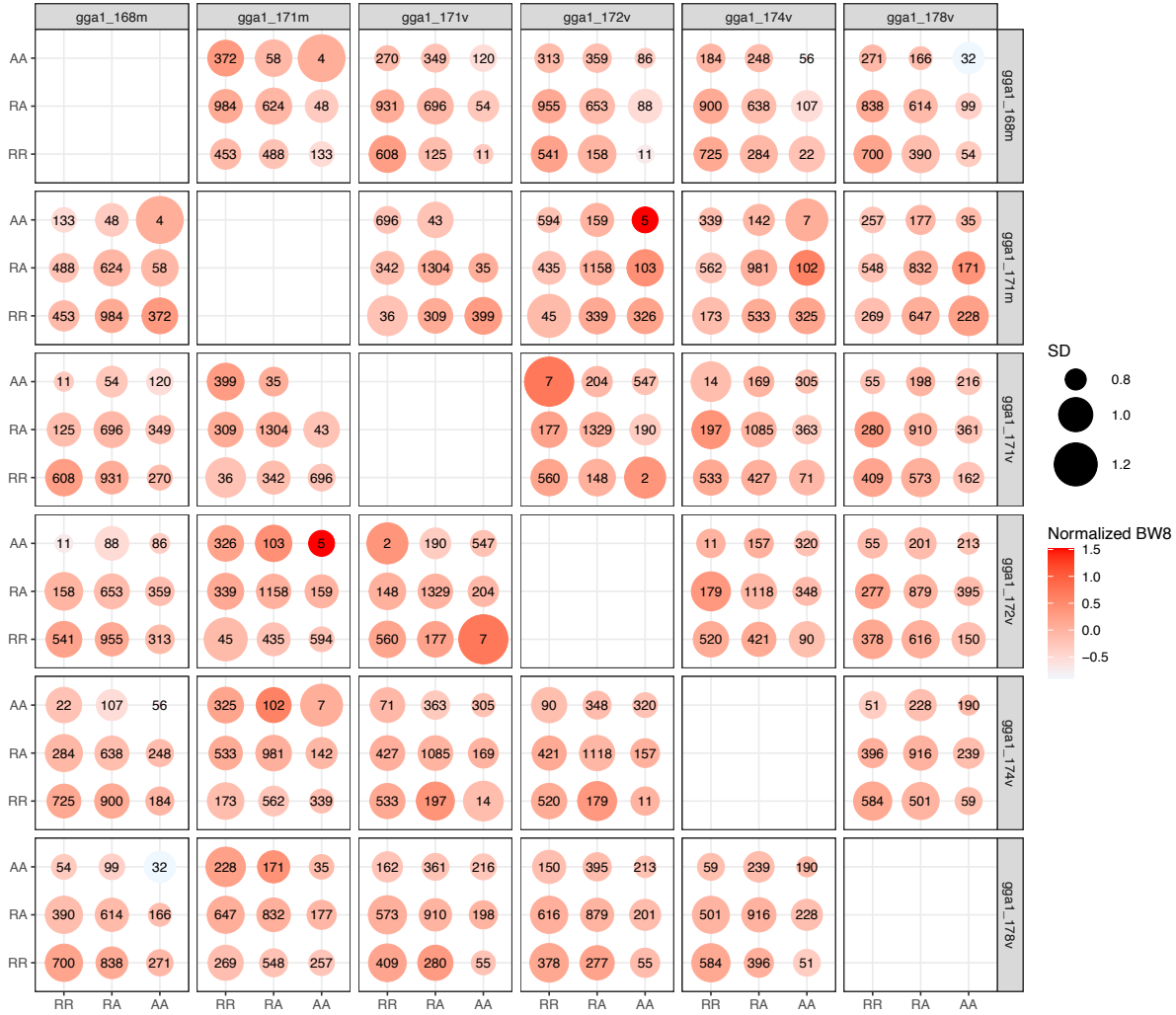
[DA]: Dominance x Additive
[AA]: Additive x Additive



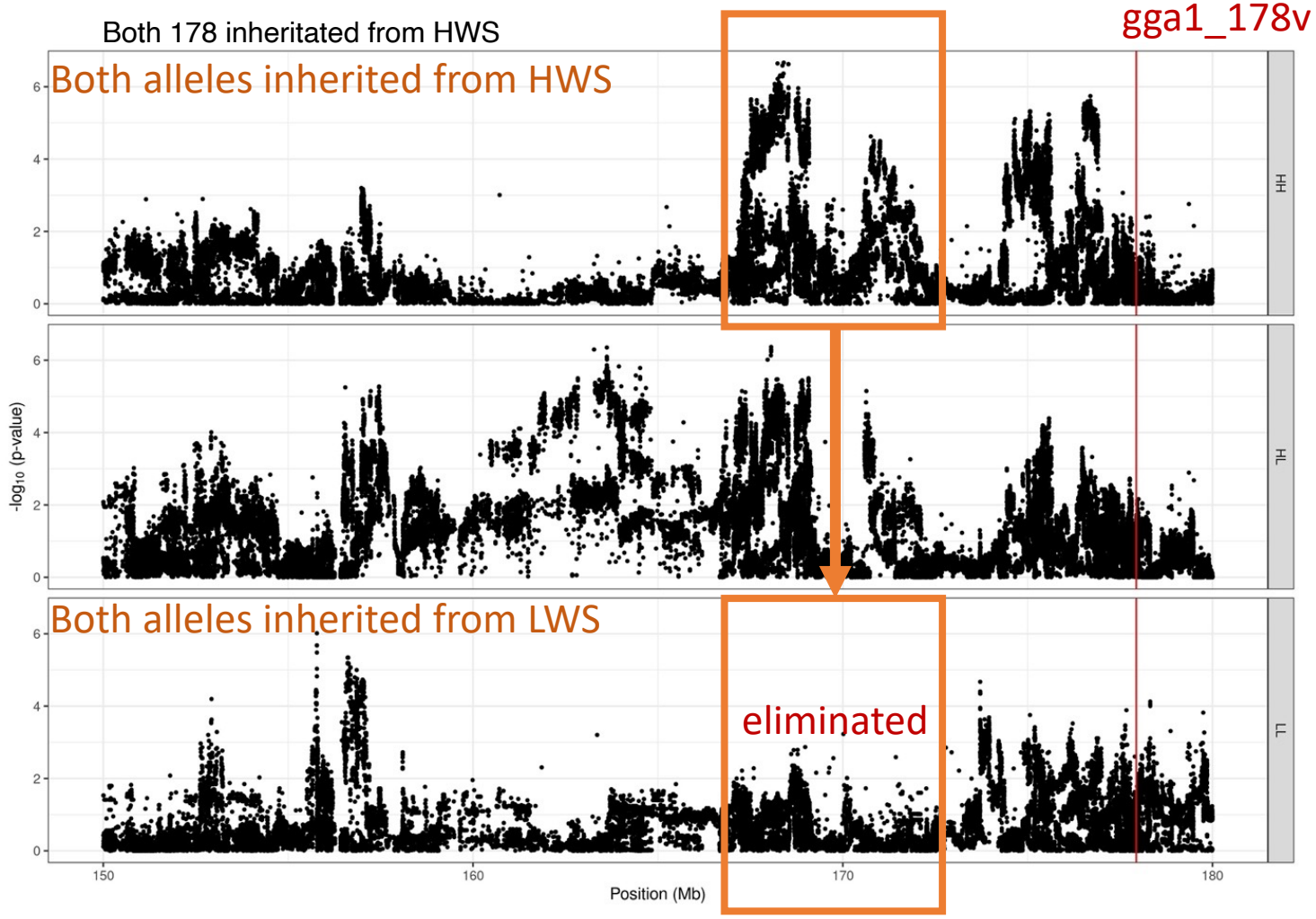
This result support that the regulating body weight by *Growth1* QTL cannot be described simply by the independent effects of the loci.

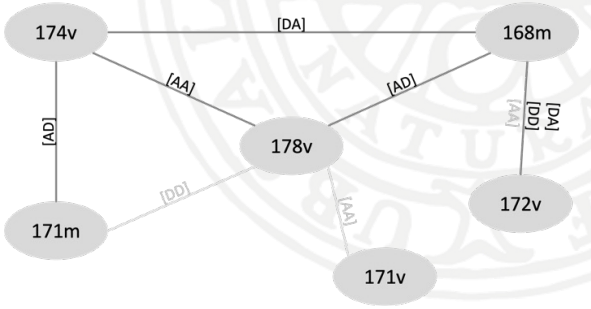
Epistasis effects

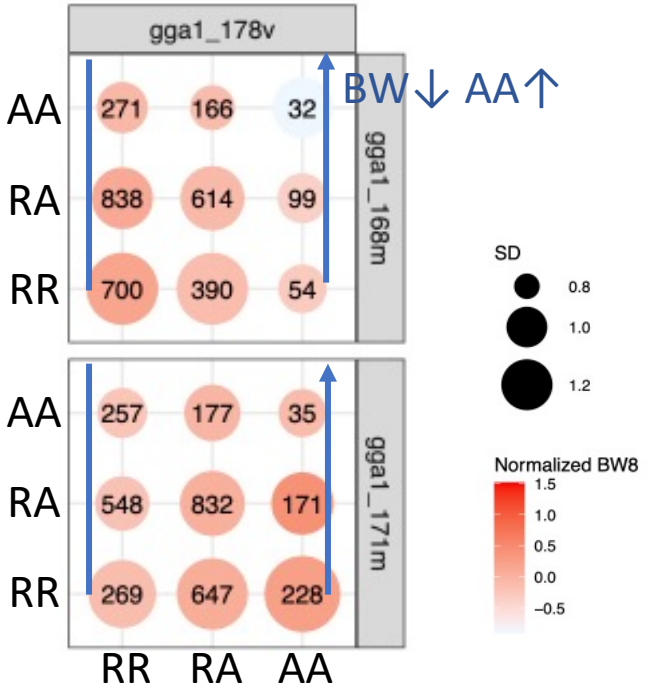
- To confirm statistical epistasis effects determined by the NOIA model
- Results show in two ways
  - Subgroup samples by pairwise markers
    → Average / SD of bodyweight
  - GWAS condition with each selected marker



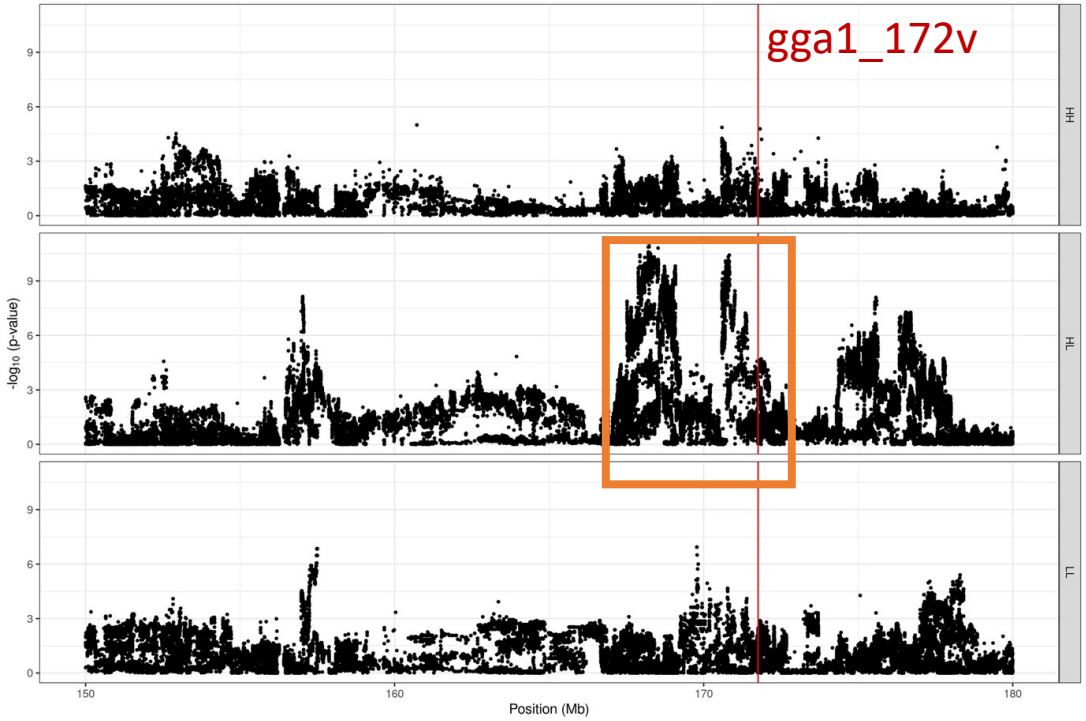
gga1_178v: The central marker of the network

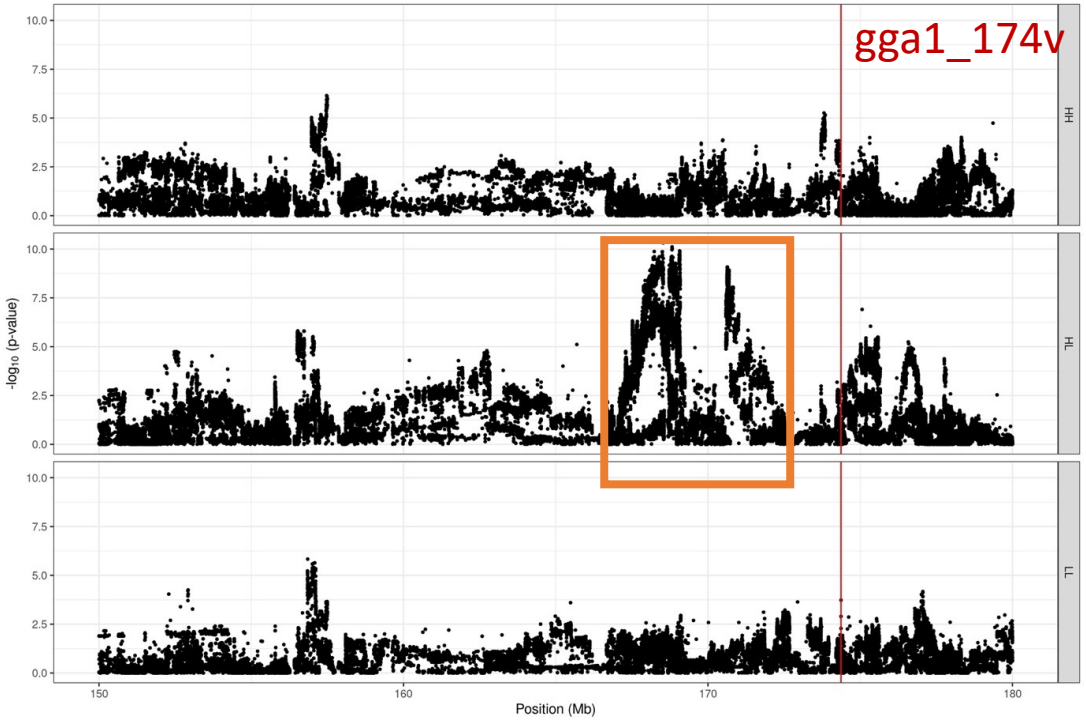


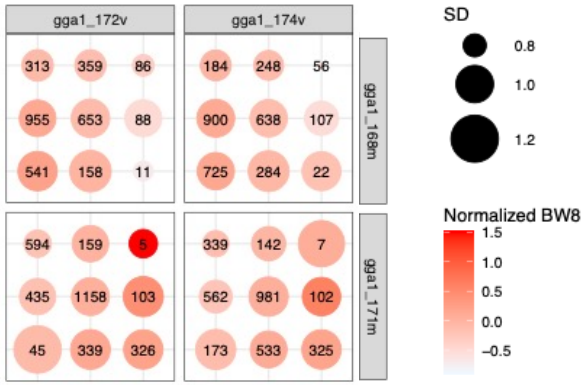


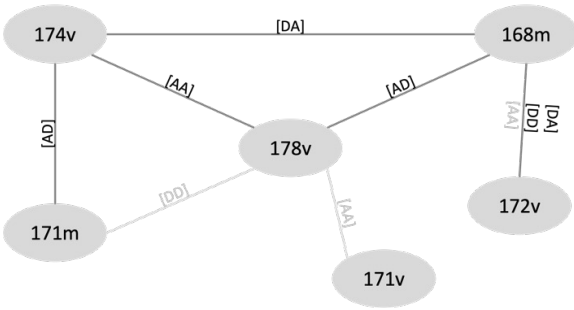


gga1_172v and gga1_174v









Conclusion

- A higher resolution with AIL population extended to generation F_{18}
- Power of association mapping is improved by local epistasis effects for haplotype construction
- A new approach of explaining the complex genetic architecture of chicken *Growth1* region
  - The haplotype-based association can majorly explain the effect while remain unexplained effect
  - NOIA model generate an interaction network of how selected SNP markers interact with each other
  - Conditioned association study confirmed the interaction mechanism

Thank you!

Dr. Carl-Johan Rubin
Dr. Örjan Carlborg
Dr. Guo Ying
Tilman Rönneburg

UPPSALA
UNIVERSITET