# Genome-wide association studies for body weight in broilers using sequencing and SNP chip data

*Eirini Tarsani, Valentina Riggio, Andreas Kranis*

# Aim of this presentation

- To compare the SNP signals between the GWAS using the SNP chip and whole-genome sequencing (WGS) data.

# Material and methods (1)

**SNP chip data:**

**60k** animals and 45,834 autosomal SNPs (chr: 1-28)

After quality control (QC) :

- **56635 samples (29865 females and 26770 males)** remained with call rate >0.90,
- **44,959 autosomal SNPs** remained with call rate >0.95, minor allele frequency >0.05 and HWE test p-value> 0.000001

*Software: PLINK*

**Sequence data:**

- We used the same 60k animals as for chip data.
- These SNP genotypes were imputed to sequence level for a total of 12,806,034 variants with high levels of accuracy (>0.95).
- Genotype imputation was conducted with Beagle software.

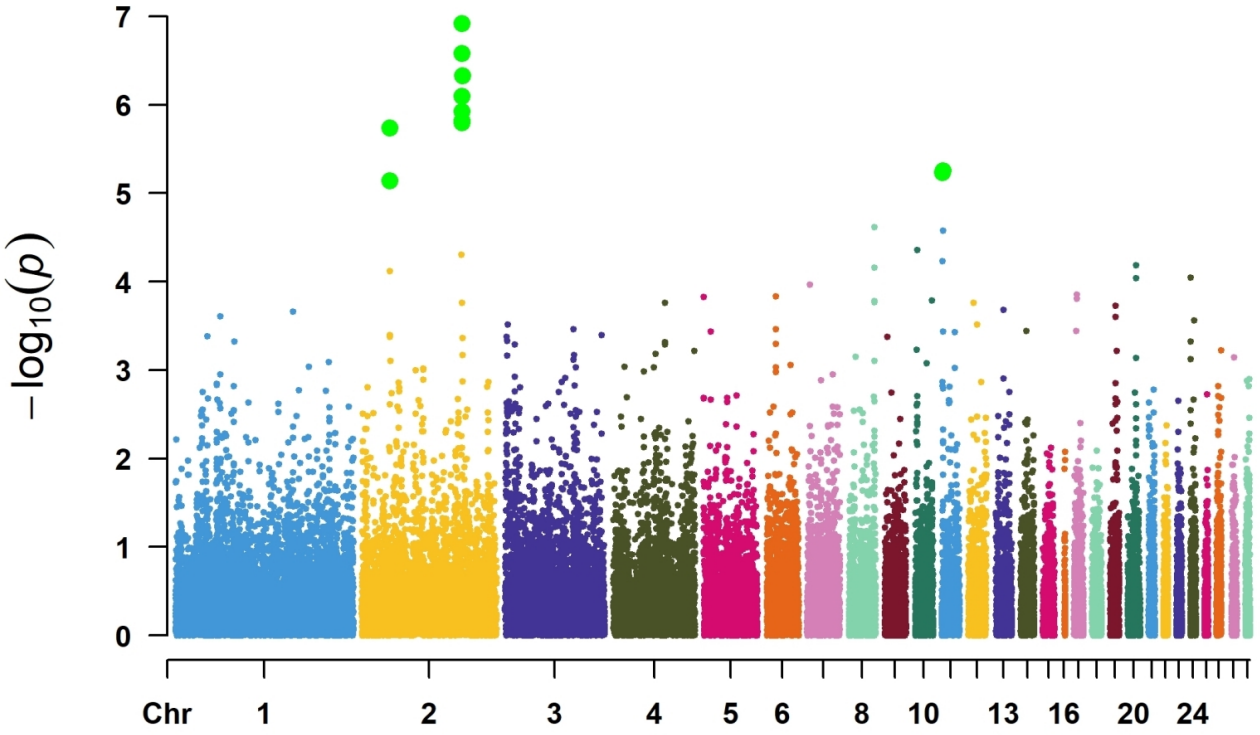
# Material and Methods (2)

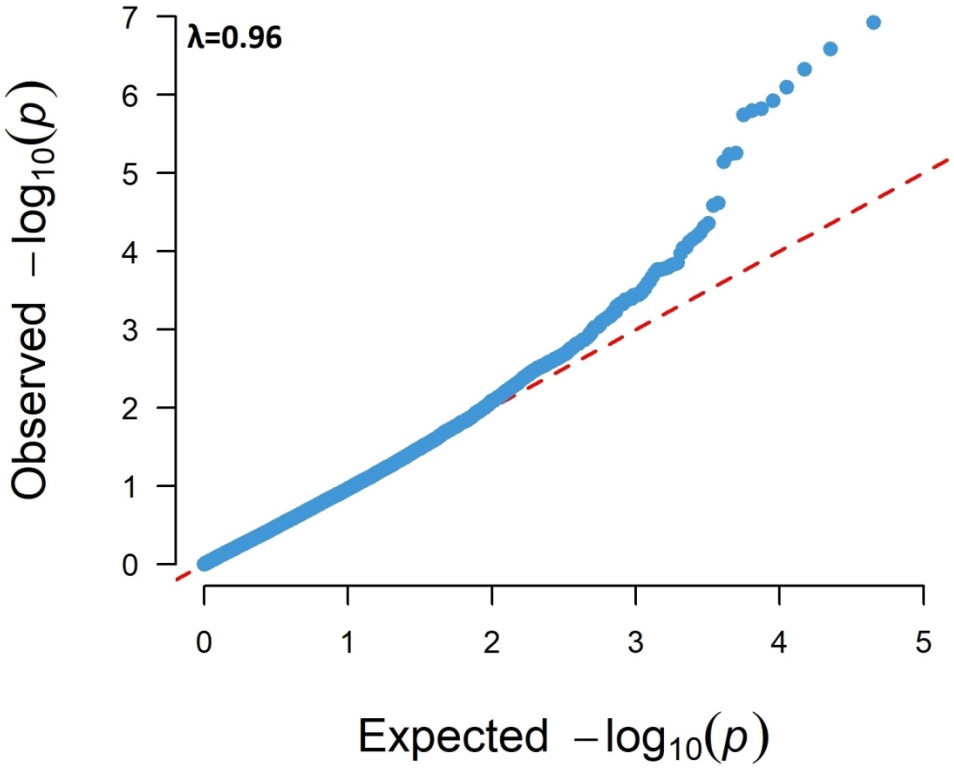
**GWAS using chip and sequence data, respectively**

- Software: GCTA
- Same Linear mixed model: Each time, a single locus mixed (additive) model GWAS:
- SNPs (fitted as fixed effect covariates) and adjusted BW for the fixed effects (no covariates)
- Animals' relationships were included in the mixed model as random effects via the genomic relationship matrix (GRM).
- Genome-wide significant SNP: FDR p-value < 0.05
- Significant markers were annotated using Ensembl Variant Effect Predictor (VEP) tool against the gene annotation database and genome reference sequence file from Ensembl for *Gallus gallus*.

# Results by GWAS using chip data

**In total: 11 genome-wide significant SNPs**



# Results by VEP for the significant markers using SNP data

| Position of the marker (bp) | Chromosome | P-value (FDR p-value) | Consequence type of variant | Positional candidate gene(s) including significant marker |
|---|---|---|---|---|
| **31128475** | **2** | **1.83379e-06 (0.009)** | **intron** | ***GPNMB*** |
| 31294098 | 2 | 7.25205e-06 (0.027) | intron | *TRA2A* |
| 111130624 | 2 | 8.01306e-07 (0.007) | Upstream gene variant | *LOC112531857* |
| 111156392 | 2 | 1.60245e-06 (0.009) | intron | *LYN* |
| 111182748 | 2 | 1.51017e-06 (0.009) | intron | *KIF20AL* |
| 111287788 | 2 | 1.20372e-07 (0.002) | intron | *CHCHD7* |
| 111314287 | 2 | 1.18994e-06 (0.008) | intron | *CHCHD7,SDR16C5* |
| 111452734 | 2 | 2.61663e-07 (0.003) | intron | *Non-coding RNA gene* |
| 111597499 | 2 | 4.70903e-07 (0.005) | intron | *IMPAD1* |
| **1236869** | **11** | **5.79365e-06 (0.023)** | **intron** | ***Non-coding RNA gene*** |
| **1288067** | **11** | **5.599e-06 (0.023)** | **intron** | ***PLEKHG4*** |

# Results by VEP for the significant markers using SNP data

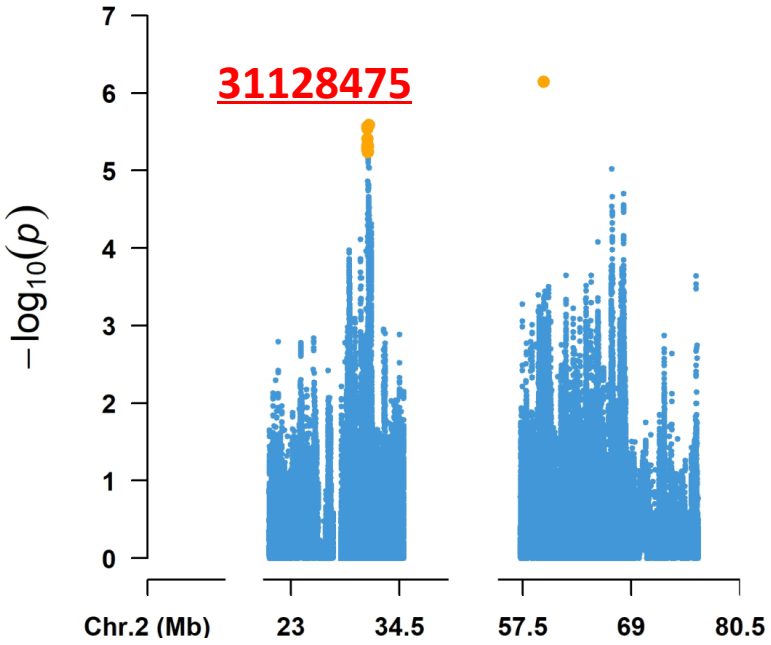
| Position of the marker (bp) | Chromosome | P-value (FDR p-value) | Consequence type of variant | Position including |
|---|---|---|---|---|
| **31128475** | **2** | **1.83379e-06 (0.009)** | **intron** | ***GPNMB*** |
| 31294098 | 2 | 7.25205e-06 (0.027) | intron | |
| 111130624 | 2 | 8.01306e-07 (0.007) | Upstream gene variant | |
| 111156392 | 2 | 1.60245e-06 (0.009) | intron | |
| 111182748 | 2 | 1.51017e-06 (0.009) | intron | |
| 111287788 | 2 | 1.20372e-07 (0.002) | intron | *CHCHD7* |
| 111314287 | 2 | 1.18994e-06 (0.008) | intron | *CHCHD7,SDR16C5* |
| 111452734 | 2 | 2.61663e-07 (0.003) | intron | *none* |
| 111597499 | 2 | 4.70903e-07 (0.005) | intron | *IMPAD1* |
| **1236869** | **11** | **5.79365e-06 (0.023)** | **intron** | ***none*** |
| **1288067** | **11** | **5.599e-06 (0.023)** | **intron** | ***PLEKHG4*** |

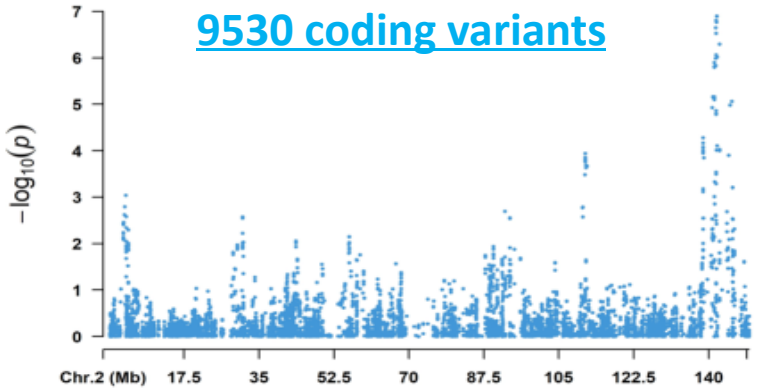
***270 sequence variants were lied within GPNMB gene***

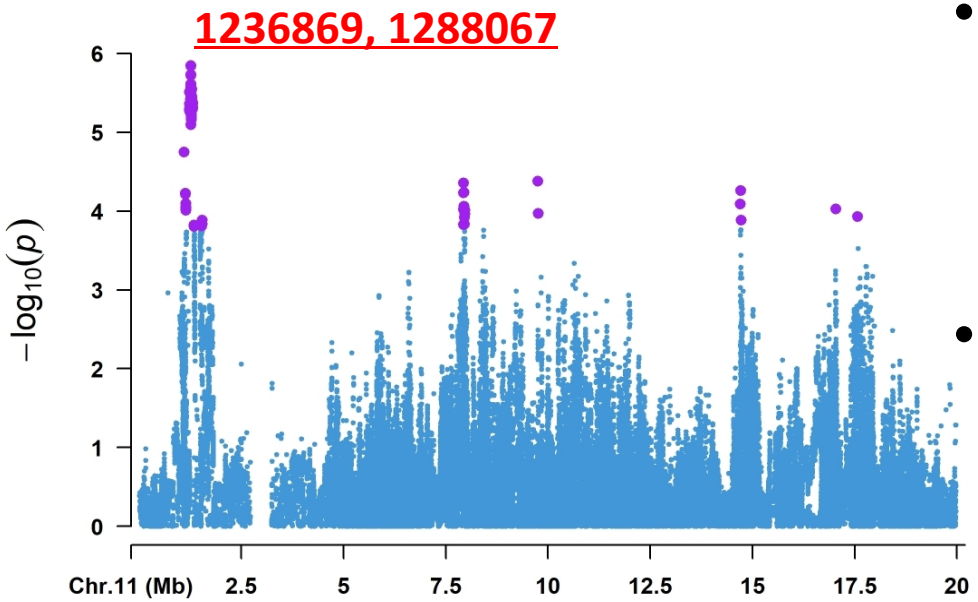
***GPNMB*** *(glycoprotein nmb)* gene regulates abdominal fat deposition in chickens (Wang et al., 2022)

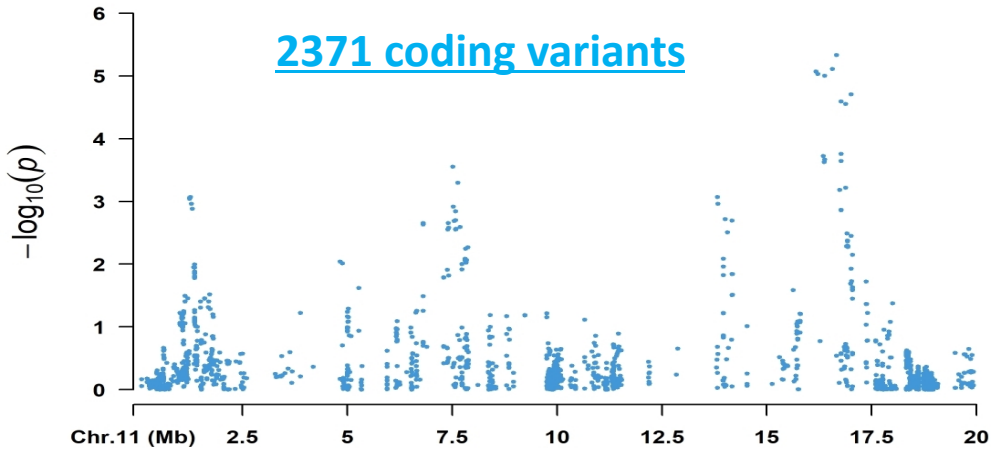
Poult Sci: 10.1016/j.psj.2022.102216.

# Common markers between GWAS using SNP vs. sequence data



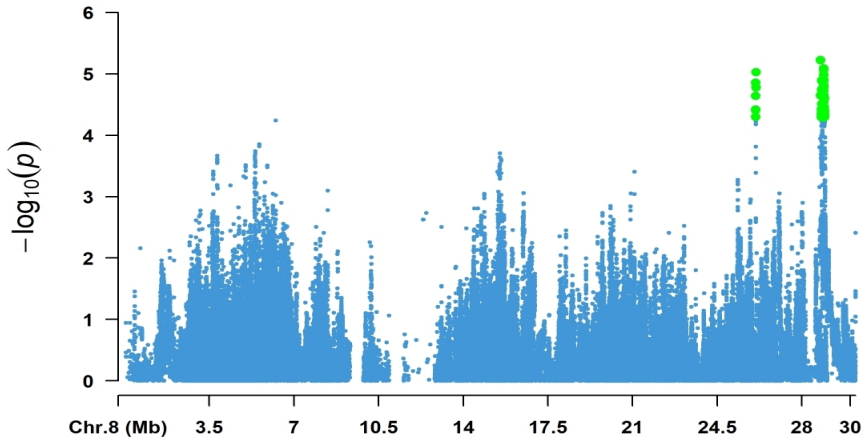



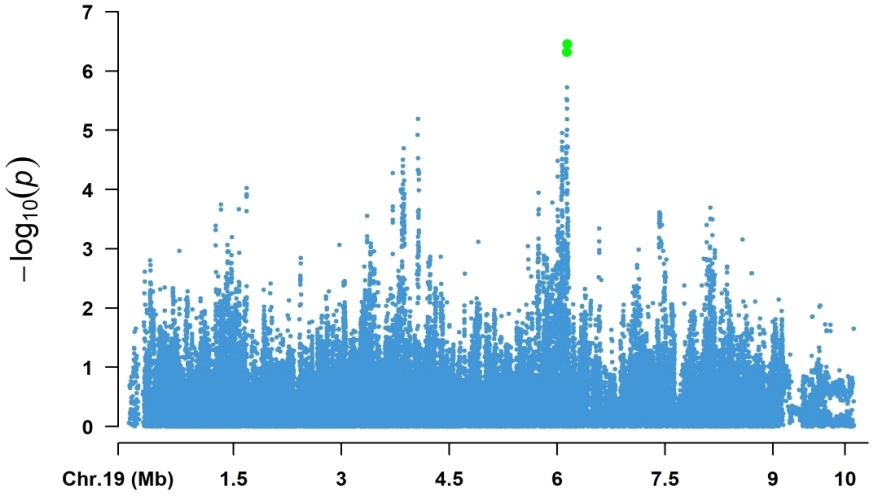



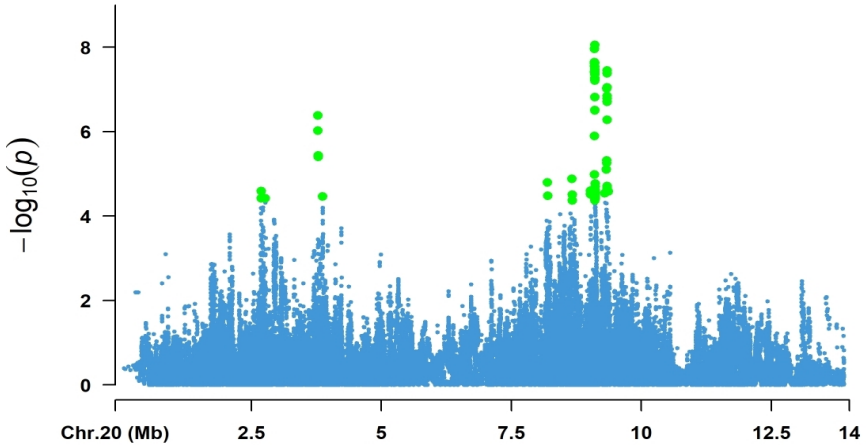



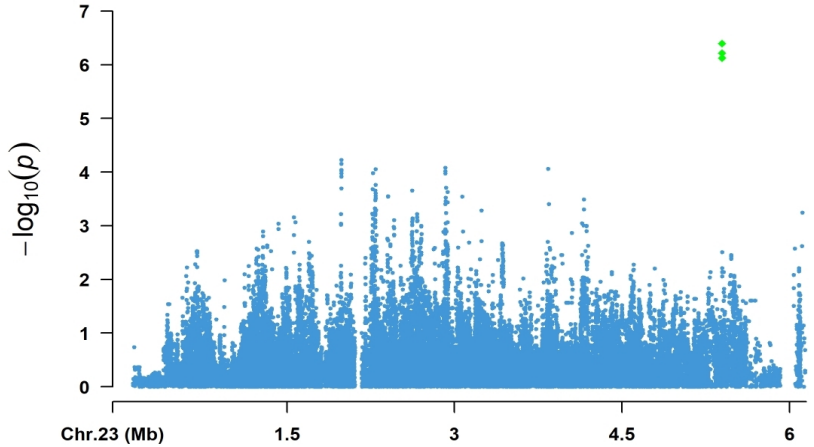

- 3 markers on GGA2 and GGA11 were common between the GWAS
- In total: 44 significant SNPs were identified by GWAS using sequence data for GGA2 while 531 SNPs were found for GGA11.

# GWAS using sequence data highlighted more significant variants e.g.




GGA8 (243 SNPs),
GGA19 (2 SNPs),

GGA20 (122 SNPs),

GGA23 (3 SNPs)

# Conclusions

- With the use of sequencing data, results were similar to those obtained with SNP chip data. However, additional GWAS signals were found.
- Our results suggest an improvement in power and precision when using WGS data. However, the use of sequence data is still time-consuming.
- In spite of the identified markers for BW, the high proportion of genetic variance attributed to regions harbouring non-significant SNPs supports the hypothesis that the genetic architecture of BW is polygenic.
- Under the assumption of polygenicity, a “large” number of genes with “small” effects is expected to control BW.
- Next step: Genomic prediction

# Thank you for attention!

Aviagen is acknowledged for data provision