# A survey of deleterious variants in highly managed commercial layer lines

Martijn Derks[1], Hendrik-Jan Megens [1], Mirte Bosse[1], Christian Gross[2,3], Marcel Reijnders[3], Dick de Ridder[2], Martien Groenen [1]

[1]Animal Breeding and Genomics Centre, Wageningen University, Wageningen, The Netherlands
[2]Bioinformatics Group, Wageningen University, The Netherlands
[3]The Delft Bioinformatics Lab, Delft University of Technology, The Netherlands

## Background

The level of deleterious genetic variation in highly managed domestic populations is influenced by several factors e.g. effective population size, and artificial selection. **Deleterious variation can impact population fitness substantially.** The majority of deleterious variants are expected to be recessive and to occur at very low allele frequency in the population, making selection against these variants inefficient.

**The current availability, however, of large numbers of whole-genome sequenced (WGS) individuals from the same population opens new possibilities to identify low-frequency deleterious alleles.** The sequences can be used to identify potential phenotype-altering alleles, ranging from embryonic lethal to only mildly deleterious mutations in coding regions. Also, advantageous functional variants can potentially be identified.

**This study presents a catalogue of deleterious variation in three commercial layer lines** that can be added to the current genomic breeding framework to purge or lower the frequencies of these variants in the population.

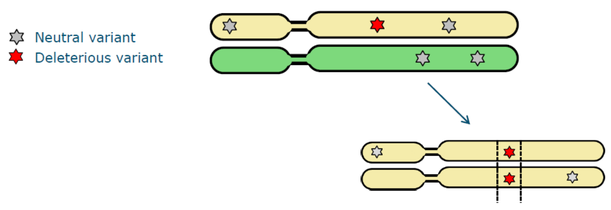


**Figure 1:** Example of deleterious (recessive) and neutral variants.

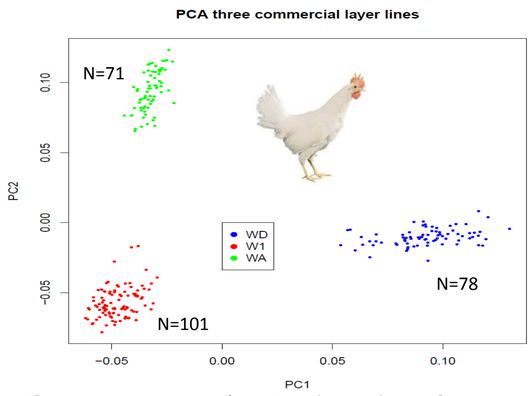


**Figure 2**: PCA analysis and number of sequenced individuals in three commercial layer lines.

## Data & Methods

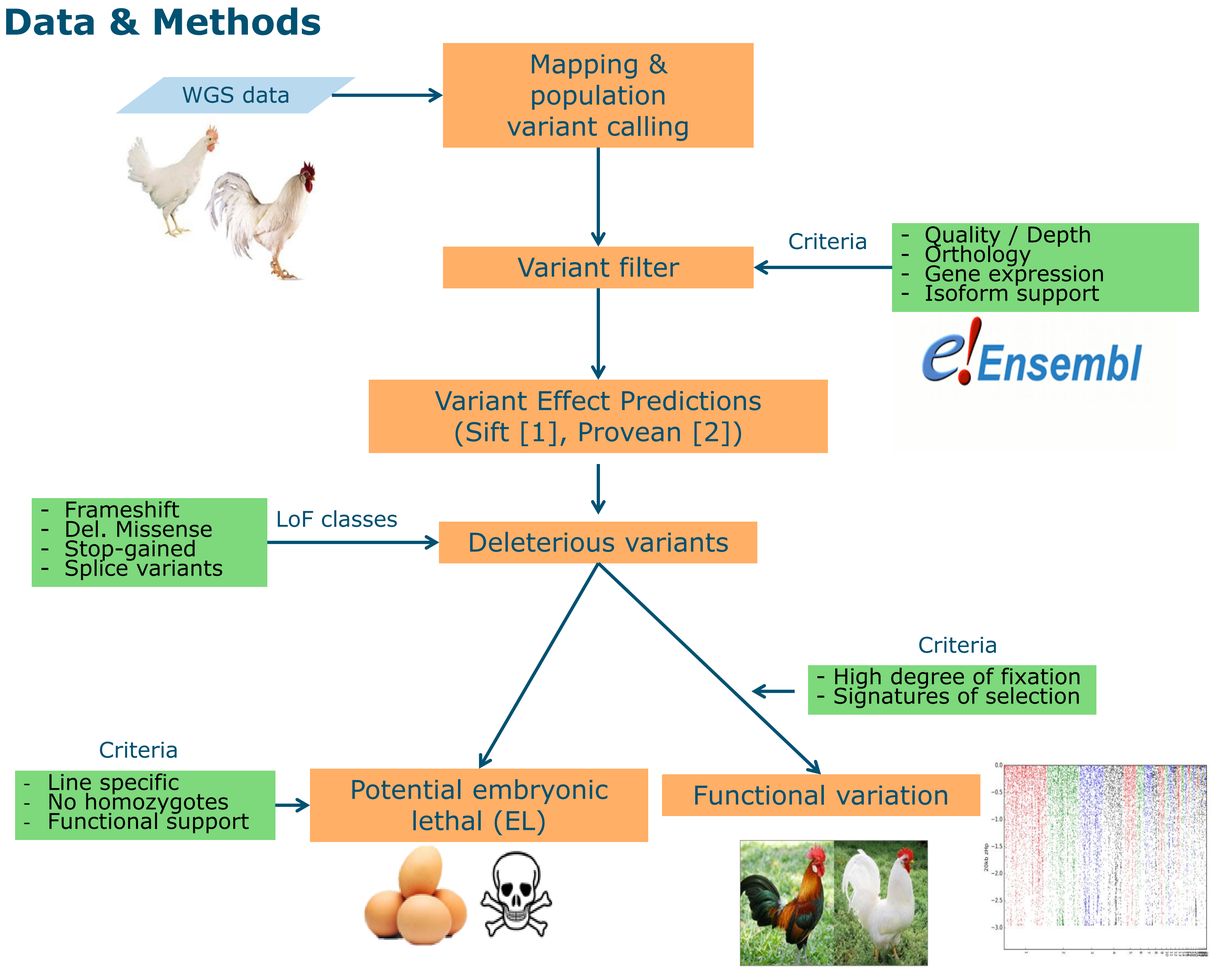


**Figure 3: Pipeline overview to detect deleterious and functional variants using population WGS data.**

## Results

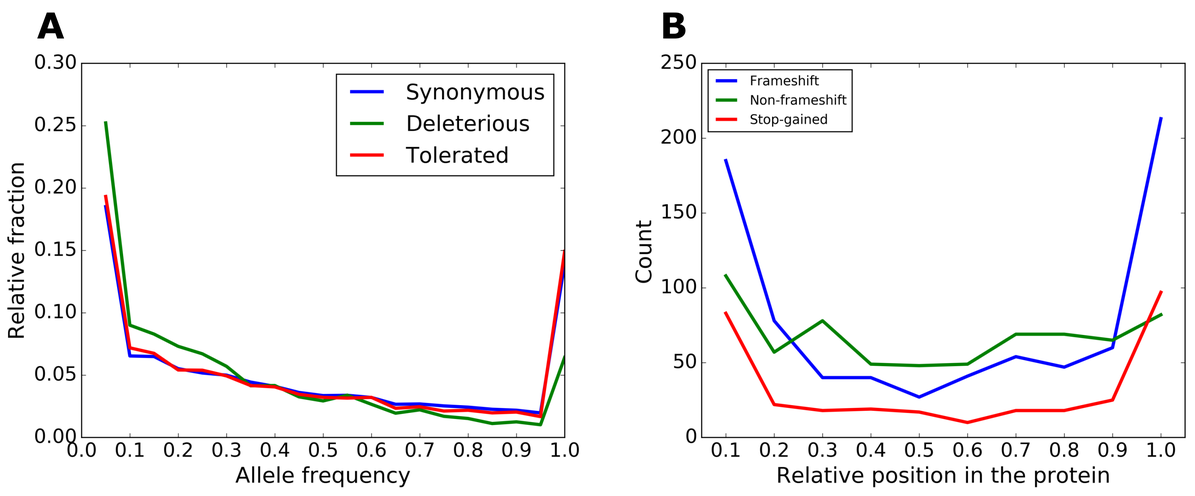


**Figure 4: A. Allele frequency distribution.** Figure shows distribution for synonymous (neutral), non-synonymous tolerated, and non-synonymous deleterious variants. **B. Relative position of frameshift, non-frameshift indels, and stop-gained variants.** Frameshift and stop-gained variants are enriched in N- and C-terminal parts of the protein.

### *Embryonic lethal variants*

**Table 1: Potential EL variants identified per variant class.**

| Functional class | Potential EL |
|---|---|
| Frameshift | 15 |
| Inframe deletion | 8 |
| Inframe insertion | 1 |
| Splice acceptor | 34 |
| Splice donor | 29 |
| Start lost | 2 |
| Stop gained | 13 |
| Del. missense | 383 |

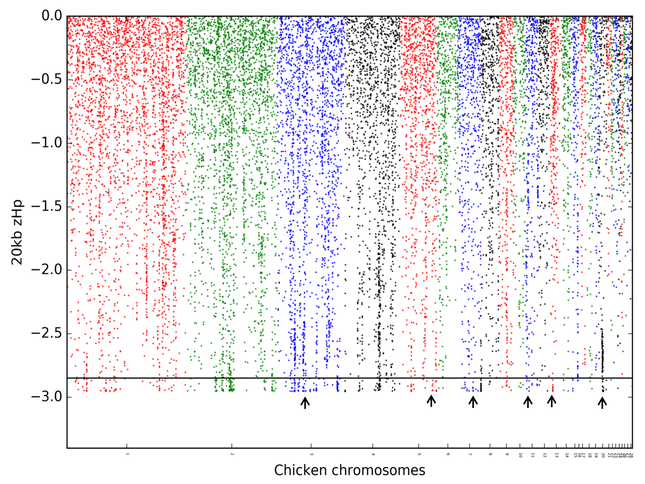


**Figure 5: Genome wide z-scores of heterozygosity (zHp).** Regions with fixed "deleterious" mutations in selective sweeps (zHp<-2.85) are highlighted with an arrow

## Conclusions

- Based on our pipeline we identified a number of deleterious and potential **embryonic lethal mutations in all analysed breeds**.
- The **allele frequency spectrum confirms** that the predicted deleterious variants were **subject to purifying selection.**
- The **enrichment of frameshift and stop-gained variants in N- and C terminal parts** of the protein suggest that these variants are not disruptive as a functional protein can still be generated.
- Breed specific and fixed deleterious variants are identified, these are likely not deleterious and are either the result of **positive selection** or genetic drift effects.
- The set of deleterious variants can be **added to the current genomic breeding framework** to purge or lower the frequencies of these variants in the population.

## References

1. Kumar, P., S. Henikoff, and P. C. Ng, *Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm.* Nat Protoc, 2009. **4**(7): p. 1073-81.
2. Choi, Y., et al., *Predicting the functional effect of amino acid substitutions and indels.* PLoS One, 2012. **7**(10): p. e46688.

Animal Breeding and Genomics Centre
P.O. Box 123, 6708 PB Wageningen
Contact: martijn.derks@wur.nl
T + 31 (0)317 48 30 19, M +31 (0)6 20 26 09 33