



Dissecting the genetic architecture of complex traits and its impact on genetic improvement programs - lessons learnt from the Virginia chicken lines

Mats E. Pettersson¹, Örjan Carlborg¹

¹ Department of Animal Breeding and Genetics, Swedish University of Agricultural Sciences, P.O. Box 7023, SE-75007 Uppsala, Sweden.

ABSTRACT - Long-term selection experiments provide a valuable resource for understanding the genetic changes that take place in animal improvement programs. The Virginia chicken lines have for more than 50 years been subjected to single-trait, bi-directional, divergent selection for high and low juvenile body weight. It is one of the most well studied experimental populations in agricultural animals and has played an instrumental role in the understanding of the genetic basis of complex traits. In this paper, we summarize the findings from the more recent publications focusing on the efforts to identify the loci contributing to selection response as well as discuss the currently ongoing research. We conclude by describing some future research prospects that promise to bring new, interesting insights into the biology of complex traits in the near future and their implications for future animal improvement programs.

Keywords: body weight, chicken, epistasis, metabolic, QTL, selection

Dissecando a arquitetura genética de características complexas e seu impacto sobre os programas de melhoramento genético - lições aprendidas com o frango Virginia

RESUMO - Experimentos de seleção a longo prazo constituem um recurso valioso para a compreensão das alterações genéticas em programas de melhoramento animal. As linhagens de frango de Virgínia há mais de 50 anos foram submetidas à seleção unicaracterística, bi-direcional e divergente para alto e baixo peso corporal em animais jovens. Trata-se de uma das populações experimentais em agropecuária que mais estudadas, desempenhando um papel fundamental na compreensão da base genética de características complexas. Neste artigo, resumimos os resultados das publicações mais recentes que focaram em identificar locos que contribuem para a resposta de seleção, bem como discutir as pesquisas atuais. Concluímos descrevendo algumas perspectivas de pesquisas que prometem trazer descobertas interessantes sobre a biologia dos multicares em um futuro próximo e suas implicações para programas de melhoramento animal.

Palavras-chave: epistase, metabólico, QTL, peso corporal, poedeiras, seleção

Introduction

Breeding of plants and animals has been a key feature in human society for thousands of years. Traditionally, breeding has been based on phenotypic selection, where the individuals with properties most appealing to the breeder have been selected as parents for the next generation. The results of such breeding schemes are often spectacular, leading to huge phenotypic changes in what is, in evolutionary terms, the blink of an eye. However, these changes are simply a reflection of underlying changes to the genome of that organism – the genotype. The effect of selection on the genomic level is less well studied, and the changes induced by directional breeding are still largely unknown.

Among the many systems and model organisms used to try and comprehend the connection between complex phenotypes and the corresponding genetic changes, a significant amount of work has been dedicated to the Virginia chicken lines. These lines are a chicken resource population designed to study the genetic, genomic and phenotypic effects of long-term, single trait, divergent artificial selection¹. In 1957, founders for one high- and one low- body weight line were selected from a seven way cross between partially inbred White Plymouth Rock chickens. Once a year, the birds with the highest and lowest eight-week body weight, respectively, are selected as parents for the next generation. Of particular importance is that the experiment involves bi-directional selection and that the population history,

including population sizes and selection intensities, is well known.

The Virginia lines are a useful resource for genetic dissection of complex traits. The strength of the system comes from the fact that the genomes of the two lines share a very recent ancestry, but have been subjected to about 50 generations of strong selection for high and low body weight, respectively. The effect on the phenotype is dramatic, at generation 40 the high line birds were nine times heavier than the low line birds. A study based on genotyping birds from both lines on a 60k Illumina SNP (Single Nucleotide Polymorphism) chip (Johansson et al., in revision) shows that the effects on the genomic level is equally dramatic, with more than a hundred regions, distributed all across the genome, being fixed for alternate alleles.

The Virginia lines have been used for several QTL studies aiming to find genes involved in regulation of body weight and other traits that have shown correlated responses in the selection program (Calborg et al., 2006; Wahlberg et al., 2009). The aforementioned study using the 60k SNP chip also provides an initial look at the genome wide effects of selection in the lines. In addition, the lines were used in a massive chicken genome re-sequencing study published in Nature (Rubin et al., 2010), and further analysis of the re-sequencing data is currently in progress (Marklund et al., in preparation).

QTL analysis

The largest body of genomics work using the Virginia lines is focused on detection of QTL (Quantitative Trait Loci) for body-weight and other correlated traits. A large number of QTL have been detected in these studies (Jacobsson et al., 2005; Park et al., 2006; Wahlberg et al. 2009). The first study was performed by Jacobsson et al. (2005) on an intercross between the high and low line with approximately 800 F₂ individuals, genotyped for 145 markers distributed across 26 linkage groups covering approximately 80% of the chicken genome. This study detected 13 QTL associated with growth, designated *Growth1-13*. Park et al. (2006) extended the investigation of the same dataset by looking for association with other, but related, traits including abdominal fat deposition and breast muscle weight. This study reported several additional QTL and also showed a high degree of colocalization between QTL for the different traits as well as a partial, but not complete, overlap with the previously reported *Growth1-13*. We summarize the findings from these studies in Table 1. Despite the large population size for an experimental intercross, relatively few genome-wide QTL were, however, reported.

Table 1 - QTL reported in the Virginia lines. A summary of the single QTL reported for growth and related traits in the Virginia lines. Compiled from Jacobsson et al. (2005), Park et al. (2006) and Wahlberg et al. (2009)

Chr	Publication	Start position	End position	Designation/ Trait
Chr.1	4	1000000	31400000	Breast muscle
Chr.1	4	1000000	31400000	Insulin
Chr.1	3	158400000	191800000	Growth1
Chr.1	4	169100000	191800000	Abdominal fat
Chr.1	4	169100000	191800000	Breast muscle
Chr.1	4	169100000	191800000	Shank
Chr.1	4	171700000	191800000	IGF-1
Chr.1	5	171900000	176000000	Growth1
Chr.2	4	15300000	37900000	Insulin
Chr.2	3	37900000	59400000	Growth2
Chr.2	4	96100000	144900000	Triglycerides
Chr.2	3	101600000	131900000	Growth3
Chr.3	3	24200000	79800000	Growth4
Chr.3	4	24200000	51600000	Abdominal fat
Chr.3	4	24200000	45200000	Breast muscle
Chr.3	4	24200000	45200000	Cholesterol
Chr.3	4	45200000	79800000	Lung
Chr.3	3	79800000	103800000	Growth5
Chr.4	3	3500000	19200000	Growth6
Chr.4	5	5100000	35000000	Growth6_7
Chr.4	3	44700000	62100000	Growth7
Chr.4	4	62100000	81500000	Breast muscle
Chr.5	3	21600000	44200000	Growth8
Chr.7	3	11900000	38000000	Growth9
Chr.7	5	21800000	33300000	Growth9
Chr.7	4	25300000	38000000	Abdominal fat
Chr.9	4	22600000	24800000	Cholesterol
Chr.10	4	700000	6700000	Spleen
Chr.11	4	5700000	20600000	Spleen
Chr.13	3	9400000	17200000	Growth10
Chr.17	4	700000	5700000	Bursa
Chr.20	3	400000	3100000	Growth11
Chr.20	4	400000	9700000	Glucose
Chr.20	3	3100000	11200000	Growth12
Chr.20	4	3100000	12000000	Cholesterol
Chr.20	5	6000000	12000000	Growth12
Chr.26	4	1900000	4200000	Bursa
Chr.26	4	2400000	4700000	Shank
Chr.27	4	2879303	7388499	Shank
Chr.27	4	2879303	7388499	Glucose
Chr.28	3	2500000	2500000	Growth13

The confidence intervals of the initially detected QTL regions were rather broad, partially due to relatively sparse microsatellite map used. Additional SNP markers were added by Wahlberg et al. (2009), 384 highly informative SNP:s selected from a 13k SNP chip, and a total map of > 500 markers were used for a re-analysis of the same set of F₂ individuals. In this study, 6 of the original 13 QTL were confirmed as significant for at least one body-weight trait, and the overall profile of the QTL signal was very consistent with that found by Jacobsson et al. (2005).

The Virginia lines have, thus, proven to be highly useful for QTL detection and have been key to the mapping

of genetic regions involved in growth regulation in chickens.

Detection of epistatic interactions

Beyond the single QTL analysis, the Virginia lines have also been used in several groundbreaking studies on the role of genetic interactions, often referred to as epistasis, in complex trait determination. The prevalence of epistasis in the system was originally shown by Carlborg et al. (2006). Here, an epistatic network including six QTL (*Growth2*, *Growth4*, *Growth6*, *Growth9*, *Growth12* and *Growth14*) was identified. Four of the loci (*Growth4*, *Growth6*, *Growth9* and *Growth12*) was shown to explain approximately 45% of the difference between the two lines in the F_2 cross used by Jacobsson et al. (2005) and Park et al. (2006). The results indicated that the genotype at *Growth9* had a significant influence on the effects of the other loci: the peripheral loci (*Growth4*, *Growth6* and *Growth12*) had no or very small effects in individuals homozygous for the allele inherited from the low-weight line at *Growth9*. They did, however, have strong effects in other genetic backgrounds, especially when *Growth9* was homozygous for the allele inherited from the high weight line. In the same way, homozygosity for the high-weight alleles at *Growth4*, *Growth6* and *Growth12* releases the genetic effects on growth by *Growth9*, whereas the low weight line alleles hides these effects (Figure 1). This mechanism closely resembles what has been described as Selection Induced Genetic Variation (SIGV).

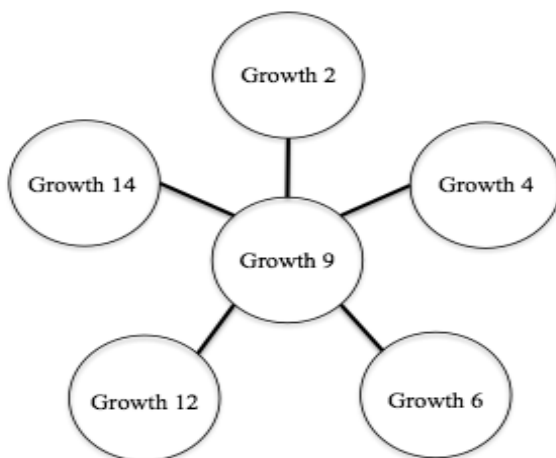


Figure 1 - Graphical representation of the epistatic network regulating growth in the Virginia lines F_2 cross. Each line represents a significant epistatic interaction. Adapted from Carlborg et al. (2006).

Fine mapping and replication of QTL effects using an Advanced Intercross Line

The founders from the Virginia line F_2 intercross (Carlborg et al., 2006; Wahlberg et al., 2009) has also been used to breed an advanced intercross line (AIL). This AIL, which is maintained with approximately 100 individuals in each generation, aims to break up the linkage disequilibrium created by crossing the divergently selected Virginia lines and through this reach a population where the stretches of genetic material inherited from each founder line (haplotypes) are much shorter than in the F_2 population. This is valuable for fine-mapping efforts, since shorter haplotypes allows better resolution when determining QTL locations. When this AIL reached the F_8 generation, approximately 400 individuals were hatched and phenotyped more comprehensively to generate a data set to replicate and fine-map the previously reported QTL (Jacobsson et al., 2005; Wahlberg et al., 2009). This data set is currently being investigated (Besnier et al., in revision; Besnier et al., in preparation; Pettersson et al., unpublished). Preliminary results indicate that the most of the previously reported loci, as well as the epistatic interactions, were replicated in the AIL. The estimates of the genetic effects is generally lower, which is to be expected given the Beavis effect, and for example the epistatic interaction between the four loci that previously explained about 40% of the difference between the lines now explain roughly 25% (Besnier et al., in revision).

Large-scale genome wide changes

The Virginia lines have also been utilized in studies aiming to determine the genome wide effects of long-term directional selection. Recently, they were included in a massive chicken genome re-sequencing project by Rubin et al. (2010). Here, a large number of chicken lines were re-sequenced to approximately 5X genome coverage with the aim to explain the major genetic differences separating layers and broilers - the two major types of commercially grown chickens. Several regions were found that were likely to have undergone selective sweeps during the differential selection period, as well as a several mutations that are possibly functionally contributing to some of the differences between layers and broilers. The well-documented history of the Virginia lines was highly useful in this study as it allowed a particular sweep, in the gene *TBC1D1*, to be dated with good precision since it is present in both the high and low body weight lines.

In addition to this large-scale study, there is also an ongoing investigation aiming at discovering the genomic changes that have taken place during the Virginia line

selection experiment. For this, a number of birds from both lines and from different time points in the selection process, have been genotyped using a 60k Illumina SNP chip. The results clearly show that a large number (~100) of genomic regions encompassing roughly 1500 SNPs have been fixed for different haplotypes in the two lines. This is a dramatic amount given the short time since the lines were separated. Naturally, some of these changes are due to genetic drift; although the breeding scheme has been set up to minimize drift, the effective population size is by necessity not large (~40). However, even taking drift into account, the results strongly support the hypothesis that the differences between the two lines are caused by a large number of loci rather than a few major ones (Johansson et al., in revision).

Future prospects

In addition to the future prospects described earlier, an effort is in progress to use genomics to identify the functional mutations underlying the identified QTL and selective sweeps. For this, the Virginia lines will be sequenced with high coverage to identify virtually all SNPs differentiating the lines. A bioinformatics pipeline is being developed that will utilize comparative genomics data to identify the most plausible causative mutations in the selected regions, and thus reveal potential metabolic pathways and other biological systems that have been affected by the differential selection process (Marklund et al., unpublished).

One of the most interesting new developments in animal improvement programs is genomic selection, where data on the genome-wide genotype of individuals is utilized in the evaluation of breeders. Today, we know very little of what to expect in terms of how many loci, which genetic mechanisms, including epistasis, that contribute to phenotypic response to selection in populations, both in a short- and a long-term perspective. The studies on the Virginia lines have already, and will continue to, provide important insights to the basic genetic mechanisms involved in long-term response to selection. Several of the insights gained have particularly important implications for animal improvement programs. We provided the first experimental evidence for an important contribution of epistatic interactions to long-term selection response (Le Rouzic et al., 2008). The study showed how epistatic interactions can regulate the release of standing genetic variation in response to selection and highlights the importance of future attempts to identify and decipher epistasis in other selected populations as well to explore how common these mechanisms are. Knowing more about these mechanisms and their existence in particular

populations will help us to better understand how to predict the limits of long-term selection. The better insights gained into the number of loci contributing to short- and long-term selection is also useful input to the efforts of understanding how to predict the genetic value of individuals using genome-wide marker sets. Our study indicate that current response to selection is primarily the result of stronger selection acting on 10s of genes, whereas long-term selection acts on considerably more genes. There might thus be a conflict between the most rapid short-term response and potential of long-term response that needs to be balanced when designing genomic selection strategies. It will be of great interest to see if this pattern is replicated in other populations under selection. In the future, we hope that the insights gained from identifying the mutations in genes and pathways that have contributed to the selection response in the Virginia lines will help us gain a better understanding about the molecular mechanisms underlying the particular traits studied in these lines as well as more general questions including the biological explanation for e.g. correlated responses to selection.

Final Considerations

In summary, we have here reviewed some of the more recent insights gained into the biology of long-term selection response from the studies of the Virginia chicken lines. Much work remains, but the results highlight the usefulness of using genomics tools in quantitative genetic studies. In the future, we hope to see more studies where this is done and as genomic characterization of individuals and populations becomes more affordable with time the future looks bright.

References

- CARLBORG, Ö.; JACOBSSON, L.; ÅHGREN, P. et al. Epistasis and the release of genetic variation during long-term selection. *Nature Genetics*, v.38, p.418-419, 2006.
- DUNNINGTON, E.A.; SIEGEL, P.B. Long-term divergent selection for eight-week body weight in White Plymouth Rock chickens. *Poultry Science*, v.75, p.1168-1179, 1996.
- EITAN, Y.; SOLLER, M. Selection induced genetic variation: A new model to explain direct and indirect effects of sixty years of commercial selection for juvenile growth rate in broiler chickens, with implications for episodes of rapid evolutionary change. In: WASSER, S.P. (Ed.) **Evolutionary theory and processes: modern horizons**. Papers in Honour of Eviatar Nevo. Netherlands: Kluwer Academic, 2004. p.153-176.
- JACOBSSON, L.; PARK, H.; WAHLBERG, P. et al. Many QTLs with minor additive effects are associated with a large difference in growth between two selection lines in chickens. *Genetic Research*, v.86, p.115-125, 2005.

- LE ROUZIC, A.; ALVAREZ-CASTRO, J.M.; CARLBORG, O. Dissection of the genetic architecture of body weight in chicken reveals the impact of epistasis on domestication traits. **Genetics**, v.179, p.1591-1599, 2008.
- PARK, H.B.; JACOBSSON, L.; WAHLBERG, P. et al. QTL analysis of body composition and metabolic traits in an intercross between chicken lines divergently selected for growth. **Physiology Genomics**, v.25, p.216-223, 2006.
- RUBIN, C.J.; ZODY, M.C.; ERIKSSON, J. et al. Whole-genome resequencing reveals loci under selection during chicken domestication. **Nature**, v.464, p.587-591, 2010.
- WAHLBERG, P.; CARLBORG, O.; FOGGIO, M. et al. Genetic analysis of an F-2 intercross between two chicken lines divergently selected for body-weight. **BMC Genomics**, v.10, p.248-260, 2009.