

Dissection of the Genetic Architecture of Body Weight in Chicken Reveals the Impact of Epistasis on Domestication Traits

Arnaud Le Rouzic,¹ José M. Álvarez-Castro² and Örjan Carlborg²

Linnaeus Centre for Bioinformatics, Uppsala University, 751 24 Uppsala, Sweden

Manuscript received March 14, 2008

Accepted for publication May 6, 2008

ABSTRACT

In this contribution, we study the genetic mechanisms leading to differences in the observed growth patterns of domesticated White Leghorn chickens and their wild ancestor the red jungle fowl. An epistatic QTL analysis for several body-weight measures from hatch to adulthood confirms earlier findings that polymorphisms at >15 loci contribute to body-weight determination in an F₂ intercross between these populations and that many loci are involved in complex genetic interactions. Here, we use a new genetic model to decompose the genetic effects of this multilocus epistatic genetic network. The results show how the functional modeling of genetic effects provides new insights into how genetic interactions in a large set of loci jointly contribute to phenotypic expression. By exploring the functional effects of QTL alleles, we show that some alleles can display temporal shifts in the expression of genetic effects due to their dependencies on the genetic background. Our results demonstrate that the effects of many genes are dependent on genetic interactions with other loci and how their involvement in the domestication process relies on these interactions.

UNDERSTANDING the impact of epistasis on the evolution of multifactorial traits remains a major challenge in complex-trait genetics. Epistasis is more complicated to model, detect, and interpret than marginal (*i.e.*, additive and dominance) genetic effects since the effect of specific alleles at a locus depends on allelic frequencies at other loci. In a population under natural or artificial selection, allele frequencies will change over time and, as a result of this, so will the genetic effects. Explorations of the impact of genetic interactions on phenotypic evolution thus rely on the study of populations in which both genetic and phenotypic information is available. This requires models that are able to decouple the effect of genetic interactions on the displayed genetic variance and to estimate the effect of allele substitutions in different genetic backgrounds.

Domestication of animals and plants provides outstanding examples of rapid evolution. The genetic architecture (*i.e.*, the number of genes and alleles, as well as the nature of interactions among them) that underlies a trait of agricultural interest determines how fast and how far a domesticated species is able to respond to long-term directional selection (LE ROUZIC *et al.* 2007; LE ROUZIC and CARLBORG 2008). Dissecting the genetic differences between domesticated strains and the cor-

responding wild populations is a particularly relevant approach to unravel mechanisms involved in the domestication process. Wild and domestic populations normally display large phenotypic differences for a wide range of traits and as domestication has been a rapid process in an evolutionary perspective, a reasonably low number of major genetic factors are expected to contribute to these differences. As wild and domestic populations for agricultural traits produce viable offspring, quantitative trait loci (QTL) detection is a particularly efficient methodology to dissect the genetic architecture involved in domestication (see, *e.g.*, DOEBLEY *et al.* 1995 and TANKSLEY *et al.* 1996 for plants or ANDERSSON *et al.* 1994 for animals).

The increase in the body weight in farm animals is a good example for which a quantitative trait has been drastically modified during domestication, leading to, *e.g.*, a twofold increase in body size in adult layer-type chickens compared to their wild ancestor. The growth of an animal is a complex process involving the basic genetics of metabolism and health in addition to the general adaptation to a particular environment. Some recent studies aiming to dissect the molecular basis of chicken growth using data from crosses between artificially selected lines or between wild and domesticated strains have found that (i) the genetic architecture of body weight is a polygenic trait (up to 20 loci involved) (CARLBORG *et al.* 2003; KERJE *et al.* 2003; JACOBSSON *et al.* 2005) and (ii) a significant part of the genetic variation in body weight is due to epistatic effects (CARLBORG *et al.* 2003, 2006).

¹Corresponding author: Centre for Ecological and Evolutionary Synthesis, Department of Biology, P.O. Box 1066, Blindern, 0316 Oslo, Norway.
E-mail: a.p.s.lerouzic@bio.uio.no

²Present address: Department of Animal Breeding and Genetics, Swedish University of Agricultural Sciences, 750 07 Uppsala, Sweden.

The history of population-based models of genetic effects for quantitative traits stems from the foundation of quantitative genetics (FISHER 1918) and profited from landmark contributions to incorporate epistasis half a century ago (COCKERHAM 1954; KEMPTHORNE 1954). More recently, the need for a “physiological” (CHEVERUD and ROUTMAN 1995) or a “functional” (HANSEN and WAGNER 2001) approach to properly investigate the importance of epistasis in the evolutionary processes has been pointed out, and the theory to integrate this new conceptual scaffold together with the previous statistical framework has been accomplished (ÁLVAREZ-CASTRO and CARLBORG 2007). These new developments enable researchers to use the statistical models to properly obtain orthogonal estimates of genetic effects from real data and then to translate them to have a functional meaning—effects of allele substitutions performed in individual genotypes—instead.

In this contribution, we aim to dissect the architecture of the genetic differences in body weight between domesticated and wild chicken. An epistatic QTL analysis was performed in an F₂ population obtained from an intercross between a single male red jungle fowl (“wild”) and several females from a domesticated layer-type chicken population (“White Leghorn”). A subset of eight major loci was selected for further study, and we analyzed the functional effects (*i.e.*, genetic effects estimates that are independent from genotypic frequencies) of Leghorn and jungle fowl QTL alleles in both wild and domesticated multilocus genetic backgrounds.

MATERIALS AND METHODS

Biological material: We performed genome scans for individual as well as interacting QTL in a red jungle fowl × White Leghorn F₂ intercross. One red jungle fowl male was mated to three White Leghorn females, producing 1046 F₂ offspring in total, 827 of which had measured genotypes. This population has previously been used in QTL mapping for behavior traits (SCHÜTZ *et al.* 2002), egg production (WRIGHT *et al.* 2006), and morphology (CARLBORG *et al.* 2003; KERJE *et al.* 2003; RUBIN *et al.* 2007) and a full description of the mapping population can be found in those references.

Previous QTL analyses on this population were based on a linkage map including 94 markers covering 2552 cM on 25 autosomes (KERJE *et al.* 2003). Here, we use an updated marker map with 439 markers covering 3214 cM on 32 linkage groups (L. ANDERSSON, personal communication).

Phenotypes: The body weight (BW) of the F₂ chickens was measured five times, at 1, 8, 46, 112, and 200 days of age. As described in CARLBORG *et al.* (2003), four additional phenotypes, corresponding to the difference between consecutive measurements, were used to estimate the growth rates (GR) at different points in life. Among the 827 genotyped chickens, 57 had incomplete growth data (missing points in the time series, most of them being due to death before 200 days) and 5 had a pathological growth (*e.g.*, decrease of the weight between two consecutive points). These individuals were removed from the data set and our analyses were based on the remaining 765 individuals.

To get an analytical description of the shape and the dynamics of the growth over time, a Gompertz growth function

was fit to the body-weight time series. The Gompertz function is a particular case of the Richards growth function (RICHARDS 1959) and has been shown to be adequate for chicken growth modeling (ROGERS *et al.* 1987). The function used was

$$B = A \cdot \exp(b_2 \cdot b_3^t),$$

where B is the expected body weight, t is the age of the chicken in days, and A , b_2 , and b_3 are the three parameters of the Gompertz function. The parameter A (denoted $Asym$ in the rest of this article) has a direct biological meaning; it represents the expected maximum (asymptotic) body weight. Another biologically meaningful parameter, $x_{mid} = -\log(b_2)/\log(b_3)$, is the estimate of the age at which the growth rate is maximum (inflection point of the growth curve). The non-linear regressions were performed with the module “SSgomperz” in the R software (R DEVELOPMENT CORE TEAM 2007). The distribution of the regression parameters is provided in supplemental Figure 1.

A principal-components analysis (PCA) over these 13 variables (*i.e.*, five body weights, four growth rates, and four parameters from the Gompertz regression) was performed to extract a smaller number of mutually independent variables for use in the mapping procedure. Our aim was to explore how well the principal components (PCs) capture the variance of QTL affecting growth. Our analysis thus uses the principles of both “functional” approach (MA *et al.* 2002; YANG *et al.* 2006) and dimensional reduction often used in multitrait approaches (see, *e.g.*, KOROL *et al.* 2001).

QTL mapping: QTL mapping was performed using a three-step strategy. First, QTL were mapped on the basis of their marginal effect (additive and dominance), using a standard least-squares-based method (HALEY *et al.* 1994), with statistical significance assessed by randomization testing (CHURCHILL and DOERGE 1994). Significant QTL were successively added to the model as described before (CARLBORG and ANDERSSON 2002) to obtain a total model:

$$y = \mu + \sum_j (a_j + d_j) + e.$$

Second, a two-locus interaction model, including all epistatic interactions (additive by additive, additive by dominance, dominance by additive, and dominance by dominance), was used to screen for epistatic QTL as described in CARLBORG *et al.* (2003), using the model

$$y = \mu + \sum_j (a_j + d_j) + aa_{12} + ad_{12} + da_{12} + dd_{12} + e.$$

Significance for QTL pairs was assessed using empirical significance thresholds (CARLBORG and ANDERSSON 2002). Finally, a randomization test was performed to determine which model provides the best fit for all significant QTL detected. For a detailed description of the procedure used, we refer the reader to CARLBORG *et al.* (2003).

Estimation of multilocus genetic effects: A subset of eight highly significant, unlinked loci was selected for further analysis: loci 1A [chromosome 1, 105 cM (1.105)], 1C (1.481), 3B (3.174), 6A (6.60), 8A (8.65), 11B (11.53), 12A (12.35), and 27A (27.23) (see Table 1), among which all pairwise interactions were considered. The genetic effects of each selected QTL, as well as all pairwise interactions between the eight loci, were computed for all traits, including the principal components, using the regression

$$\mathbf{Y} = \mathbf{Z} \cdot \mathbf{S}_g \cdot \mathbf{E}_g + \boldsymbol{\varepsilon},$$

where \mathbf{Y} is the vector of observed phenotypes, \mathbf{Z} is the matrix linking phenotypes to corresponding genotypes, \mathbf{S}_g is the

genetic-effects design matrix of the statistical formulation of the “NOIA” model (ÁLVAREZ-CASTRO and CARLBORG 2007), and $\boldsymbol{\varepsilon}$ is the vector of random (environmental) effects. The \mathbf{Z} matrix is computed from the genotypic probabilities provided by a Haley–Knott regression (HALEY *et al.* 1994), as detailed in ÁLVAREZ-CASTRO *et al.* (2008). This regression estimates \mathbf{E}_S , the vector of statistical genetic effects. To reduce the model to account for only pairwise epistasis, we made the columns of the \mathbf{S} matrix corresponding to higher-order epistasis into columns of zeros. The estimates of genetic effects obtained are in this way average effects of allele substitutions in the sample of individuals of the QTL study. From these estimates, “functional” genetic effects \mathbf{E}_B , *i.e.*, genetic effects corresponding to allelic substitutions in a given genetic background “ B ,” can be obtained using the transformation tool

$$\mathbf{E}_B = \mathbf{S}_B^{-1} \cdot \mathbf{S}_S \cdot \mathbf{E}_S,$$

where \mathbf{S}_B is the genetic-effects designed matrix fitting the desired meaning of the new estimates. Estimates of genotypic values, *i.e.*, the genotype-to-phenotype map \mathbf{G} , can be obtained by

$$\mathbf{G} = \mathbf{S}_S \cdot \mathbf{E}_S.$$

These operations (regression and transformation tool) are described in more detail in ÁLVAREZ-CASTRO and CARLBORG (2007) and implemented in an R package (LE ROUZIC and ÁLVAREZ-CASTRO 2008). In this way, we have obtained estimates of functional effects in two particular genotypes: the homozygous jungle fowl wild genetic background (“1” alleles at the other seven loci) and the homozygous Leghorn “domestic” genetic background (“2” alleles at the other loci).

RESULTS

Chicken growth: There was a high level of correlation among the phenotypic traits (body weights at 1, 8, 46, 112, and 200 days of age, as well as the four growth rates between successive ages) (supplemental Table 1). The correlation between successive body weight measurements for the same individual over time can be modeled using a growth model, in our case the Gompertz growth function. Fitting this model on the data provides estimates for three parameters (Asym, b_2 , and b_3) from the five points measured in each series. However, even though the b_2 and b_3 (as well as the age at the inflection point x_{mid}) parameters bring a longitudinal dimension to the analysis, they are not orthogonal characters since two pairs of parameters ($A - b_2$ and $b_2 - b_3$) are substantially correlated (0.48 and -0.53 , respectively). Orthogonal scales for the phenotypes were obtained through a principal-components analysis including all 13 previously described phenotypic traits. The four first PCs explain 42.5, 25.0, 15.3, and 8.5% the total phenotypic variance, respectively (91.3% altogether).

Figure 1 and supplemental Figure 2 illustrate the decomposition of chicken growth according to the three first principal components. PC1 is clearly a total weight variable. It affects the body weight at all measured points in life of the chicken equally. PC2 and PC3

describe “switches” in the growth process, *i.e.*, early fast and late slow growth *vs.* early slow and late fast growth. This change in growth pattern is late in the case of PC2 and early in the case of PC3. Consequently, PC1 describes the general body-weight trend for an individual, while PC2 and PC3 determine the shape of the growth curve.

A network of interacting loci: Marginal-effect QTL detection was performed for the five raw body-weight measurements, the four growth rates between successive ages, and the four parameters describing the growth curve, as well as for the four first PCs. It revealed 18 genomic regions (QTL) that are significantly associated with at least one of these traits (Table 1, supplemental Data 1). Among these 18 detected loci, 11 are significant (by their marginal and/or interaction effects) at a 5% genomewide significance threshold. Almost all QTL affect several traits, which in some cases can be due to correlations between traits (for instance, BW200, Asym, and PC1 are strongly correlated, supplemental Table 1), but not always. For instance, the QTL located on chromosome 27 affects both PC1 and PC3, which are by definition not correlated. Phenotypes related to the final body weight (BW200, Asym, PC1, and to a lesser extent BW112 and the late growth rates) seem to have a broader genetic basis than the early growth traits, both according to the number of QTL and according to how much of the phenotypic difference between the founder populations they explain. The PCA clearly makes it possible to detect all QTL using fewer genome scans. PC1 alone is the trait for which the largest number of underlying loci was detected.

A two-dimensional genome scan for pairs of interacting QTL using a genetic model accounting for genetic interactions (epistasis) was also performed. Considering all traits together, a total of 41 pairs of interacting loci were significant using a 5% genomewide significance threshold for each trait (supplemental Data 2). When comparing the detected interacting pairs of loci across traits, the loci overlap to a large degree with those of the one-dimensional scan. In particular, the two loci that have the most pronounced effects in the one-dimensional scan (1A and 1C) also show the most epistatic interactions with the other loci (Table 1).

The genetic architecture underlying phenotypic change in body weight during domestication: Here we focus on studying a network containing the eight loci that have the most pronounced effects in chicken body-weight determination (underlined loci in Table 1). The decomposition of genetic variance is provided in Table 2. By applying the transformation and translation tools of the NOIA model (ÁLVAREZ-CASTRO and CARLBORG 2007) to the marginal and epistatic effects estimated for this reduced network, we can predict phenotypic values for all possible multilocus genotypes. These tools are used to compute the effects of allele substitutions performed in domestic and wild reference individual

TABLE 1
Summary of QTL detected for the 17 analyzed traits

Trait	<u>1A</u>	1B	<u>1C</u>	2A	2B	3A	<u>3B</u>	4A	4B	<u>6A</u>	7A	<u>8A</u>	11A	<u>11B</u>	<u>12A</u>	16A	20A	<u>27A</u>
BW1	1.107 (3)		<u>(3)</u>			(2)					(2)	<u>8.65 (3)</u>	(1)	<u>11.53 (2)</u>				
BW8	<u>1.105 (2)</u>		<u>1.462 (1)</u>	<u>2.39 (1)</u>		(1)	(1)			(1)					<u>(3)</u>	16.0		<u>27.26 (3)</u>
BW46	<u>1.104 (1)</u>		<u>1.480 (2)</u>			(1)	<u>3.161</u>	4.118	<u>6.61 (2)</u>						<u>(4)</u>			<u>27.26 (3)</u>
BW112	<u>1.103 (2)</u>	<u>1.139 (1)</u>	<u>1.484 (3)</u>		2.112 (1)		3.181 (2)		<u>6.61 (2)</u>					(1)	(1)			<u>27.26 (3)</u>
BW200	1.91 (1)	<u>1.138 (3)</u>	(5)			3.91 (2)					7.5 (1)	<u>8.66 (3)</u>						
GR18	<u>1.105 (2)</u>		<u>1.462 (1)</u>	<u>2.39</u>		(1)		4.117	6.54 (1)					<u>11.44 (2)</u>	<u>(2)</u>	16.0		<u>27.26 (4)</u>
GR846	<u>1.104 (1)</u>	<u>1.139 (1)</u>	<u>1.480 (4)</u>						6.63				11.4 (2)		<u>12.23 (4)</u>			
GR112200	1.73	<u>1.139 (1)</u>	<u>1.499 (4)</u>															
Asym	<u>1.103 (3)</u>	<u>1.139 (1)</u>	<u>1.485 (1)</u>		2.112					(1)				(1)				<u>27.26 (2)</u>
b_2	<u>1.105 (3)</u>	<u>(1)</u>	<u>1.452 (3)</u>															<u>(1)</u>
b_3																		
x_{mid}		<u>1.137 (1)</u>											11.0 (1)		12.63		20.35 (1)	
PC1	<u>1.104 (2)</u>	<u>1.139</u>	<u>1.485 (2)</u>				3.160 (2)		6.57 (2)					<u>11.58 (1)</u>	<u>(2)</u>	<u>16.0</u>	(1)	<u>27.26 (4)</u>
PC2		<u>(1)</u>																
PC3	<u>1.107 (1)</u>		(1)											(1)				27.21 (1)
PC4								4.8										27.20

BW: body weights at 1, 8, 46, 112, and 200 days. GR: growth rates (increase in body weight between consecutive ages). Asym, b_2 , b_3 , and x_{mid} : parameters resulting from the nonlinear regression (Gompertz function). PC1–PC4: the first four principal components. The chromosomal locations refer to the location of the significant QTL peaks in the one-dimensional scan (if any). The numbers in parentheses indicate the number of interactions involving each locus, the genome-wide significance level being indicated as follows: underlined, <5%; regular type, <10%; italics, <20%. Underlined column headings are the eight loci selected for the second part of the analysis.

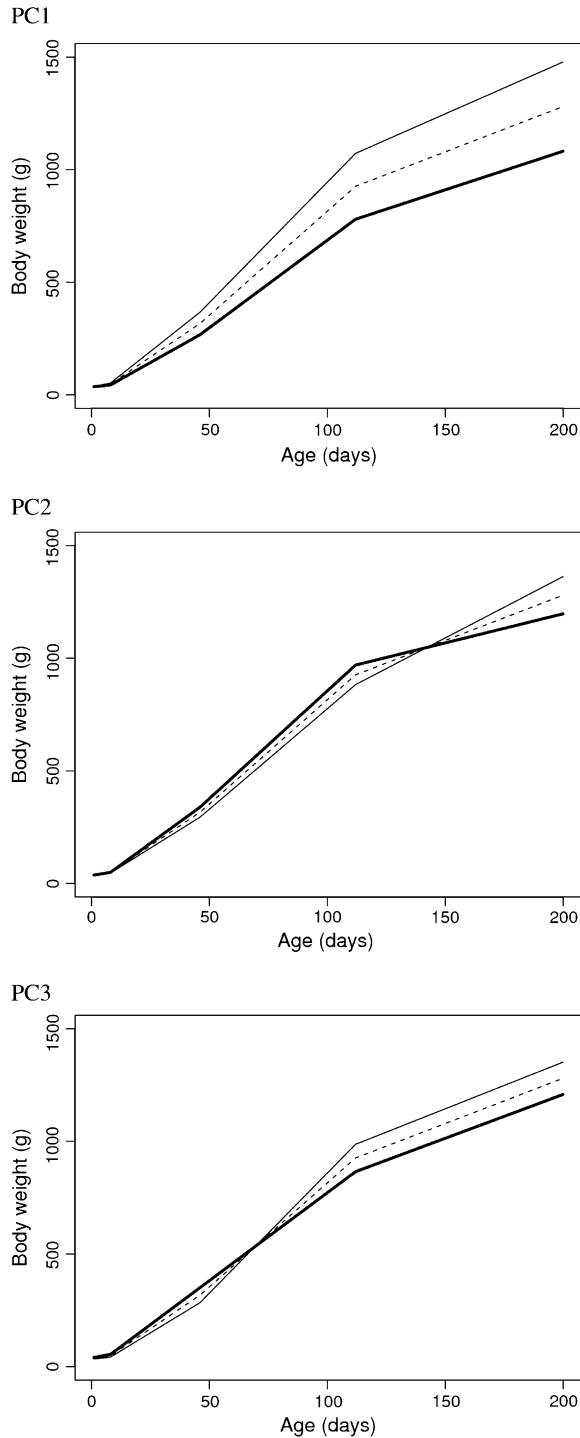


FIGURE 1.—Graphic illustration of growth pattern features described by the principal components. The mean phenotype of the F_2 population (dotted curve) is taken as a reference to describe the effects of the three first principal components. All PCs are set to 0, except one (top, PC1; middle, PC2; and bottom, PC3) that is varied from -3 (solid thin line) to 3 (solid thick line). The values -3 and 3 roughly represent the maximum and minimum values for the PCs in the population (see supplemental Figure 2).

genotypes. In other words, we inspect the functional properties of the genetic network.

Figure 2 describes the predicted effects of allelic substitutions for the eight major loci on the overall growth pattern in two genetic backgrounds: the wild background (alleles 1 at all other loci) and the domestic background (alleles 2 at all other loci). Most loci display effects that depend on the genotype at other loci, *i.e.*, epistasis, and the epistatic patterns are different for different loci in the network. Locus 1A has a major individual effect on increasing the body weight. The effect is similar in direction in both backgrounds, but appears to be earlier and larger in the wild background. Most loci (3B, 6A, 11B, 12A, and 27A) show background-specific patterns, where the effect on the final weight is often displayed only in either the wild or the domestic background, but not in both at the same time. The domestic allele in 3B and 11B increases growth early in life in a domestic and late in a wild genetic background. Two of the loci (6A and 11B) affect body weight at 200 days only in a wild background, whereas locus 12A changes body weight only in a domestic background (Figure 2).

DISCUSSION

The genetic architecture of chicken growth: The growth of a complex organism is a quantitative trait, and it is therefore expected that multiple genes are involved in determining the large phenotypic differences between the domesticated White Leghorn chicken and its ancestor the red jungle fowl. The results from the original QTL analysis of this data set (CARLBORG *et al.* 2003) support this as many QTL with both individual effects and loci with large epistatic effects were detected.

Our analysis, which accounts for individual as well as all pairwise epistatic effects in a subset of eight prominent loci, shows a strong trend toward a temporal shift of the allelic effects due to multilocus interactions with the genetic background in which the effects are measured. Alleles that tend to increase late body weight in a wild genetic background display their effects much earlier in life in a domestic background. This trend is only due to complex genetic interactions among these eight loci as no other factors have been included in the analysis. Moreover, the dynamics and robustness of the expression of allelic effects have been evaluated in reduced models, *i.e.*, where each of the eight loci are successively excluded, the effect being then measured in seven-locus networks. The analysis of these seven-locus networks (results not shown) did not indicate that the temporal phenotypic expression was due to any specific interaction with an individual locus; for instance, loci 3B and 11B keep their specific expression pattern (effect on early growth in the domestic background and on late growth in the wild background)

TABLE 2
Decomposition of genetic variance for the 17 analyzed traits

	Var(A)	Var(D)	Var(AA)	Var(AD)	Var(DD)	Var(G)	Var(P)	h^2	H^2
BW1	0.19	0.24	0.70	1.78	1.05	3.97	13.8	0.01	0.28
BW8	3.50	0.82	1.88	4.13	3.03	13.36	36.64	0.10	0.36
BW46	677.76	32.94	159.79	423.00	165.40	1,458.9	2,973.9	0.23	0.49
BW112	6,848.1	747.8	1,851.5	3,812.4	2,111.7	15,371.6	35,546	0.19	0.43
BW200	13,099	1,490.4	2,835.5	7,574.1	3,655.9	26,655.2	71,090	0.18	0.37
GR18	2.67	0.32	1.80	2.18	1.82	8.78	24.20	0.11	0.36
GR846	602.72	25.77	143.75	412.86	157.43	1,342.54	2,781.6	0.22	0.48
GR46112	3,376.4	569.3	1,300.8	2,222.2	1,409.4	8,878.1	22,977	0.15	0.39
GR112200	1,085.8	255.6	421.3	1,709.9	639.5	4,112.1	13,396	0.08	0.31
Asym	14,954	1,770	3,066	9,401	4247	33,438	87,176	0.17	0.38
b_2	0.0077	0.0047	0.0147	0.0185	0.0102	0.0559	0.1868	0.04	0.30
b_3 ($\times 10^{-6}$)	0.078	0.135	0.498	1.11	0.509	2.33	8.74	0.01	0.27
x_{mid}	1.35	0.89	2.38	8.80	3.50	16.91	65.61	0.02	0.26
PC1	2.66	0.07	0.25	0.50	0.26	3.74	5.53	0.48	0.68
PC2	0.07	0.04	0.18	0.49	0.24	1.00	3.25	0.02	0.31
PC3	0.21	0.02	0.16	0.20	0.09	0.68	1.99	0.11	0.34
PC4	0.01	0.01	0.09	0.15	0.08	0.35	1.11	0.01	0.32

Var(G) is the total of all genetic variances. H^2 stands for the broad-sense heritability ($\text{Var}(G)/\text{Var}(P)$), and h^2 is the narrow-sense heritability ($\text{Var}(A)/\text{Var}(P)$).

even when any of the other loci are excluded from the regression. The epistatic interactions that lead to the temporal shifts are thus due to multilocus interactions in a complex network rather than to a few specific and strong pairwise effects.

Chicken domestication: The domestication of chickens occurred ~ 6000 BC from the red jungle fowl *Gallus gallus* (FUMIHITO *et al.* 1994). The relationships between the different chicken breeds (including egg-, meat-, and fighting-type breeds) are somewhat complicated (MOISEYEVA *et al.* 2003), in particular because they might have resulted from multiple independent domestication events and because late introgressions from the wild species are likely (LIU *et al.* 2006). Although the White Leghorn is an egg-layer breed, it is likely that during the long domestication process, its ancestors have been subjected to direct or indirect selection for the total weight, as Leghorn chickens are now around twice as large as the wild *G. gallus*.

One of the main implications of epistatic patterns detected in our analysis is that the effects of the domestic Leghorn alleles (*i.e.*, the alleles that differ between the domesticated egg-layer chickens and the jungle fowl) depend on the genetic state of the population in which they arose by mutation or were introduced by other means. For instance, the domestic alleles in loci 6A and 11B do not increase the adult body weight in the Leghorn background: if these alleles were fixed through artificial selection for larger chickens, they must have been fixed in a background that closely resembles that of the original wild jungle fowl population. In contrast, the domestic alleles in loci 3B and 27A decrease the body weight in a genetic background similar to the jungle fowl. They are thus not expected to

be fixed by artificial selection for increased body weight early in domestication. Our results thus strongly suggest that the contribution of the loci detected in this wild \times domestic intercross to phenotypic evolution will have changed considerably during the domestication process. It is therefore not expected that the increase in allelic frequency for the loci will have been simultaneous as, *e.g.*, the domestic allele at loci 6A and 11B is more or less neutral in the domesticated chickens, indicating either that the selection on these loci took place early in domestication or that they have a major effect on other selected traits. Locus 27A, on the other hand, has a very low effect in a wild background and is thus expected to have been selected late in the domestication process. The domestic alleles at some loci, *e.g.*, loci 1A and 1C, increase body weight in all genetic backgrounds and these alleles could thus have spread in the population at any time. The domestic alleles at other loci, such as 8A or 12A, appear to have even slightly negative effects on body weight. The fixation of these alleles might be unrelated to artificial selection and due to, *e.g.*, genetic drift or genetic linkage (Hill–Robertson effect). It may also be due to pleiotropic effects on another selected trait (fertility, egg production, muscle–fat ratio, etc.). As the Leghorn breed has not been directly selected in its recent history for increased body weight but rather for increased egg production, pleiotropy appears to be a plausible explanation.

Epistasis, pleiotropy, and the genetic analysis of complex traits: The potential impact of epistasis on the genetic architecture of quantitative traits has been intensively addressed by theory (*e.g.*, GOODNIGHT 1995; RICE 2000; HANSEN and WAGNER 2001; BARTON and TURELLI

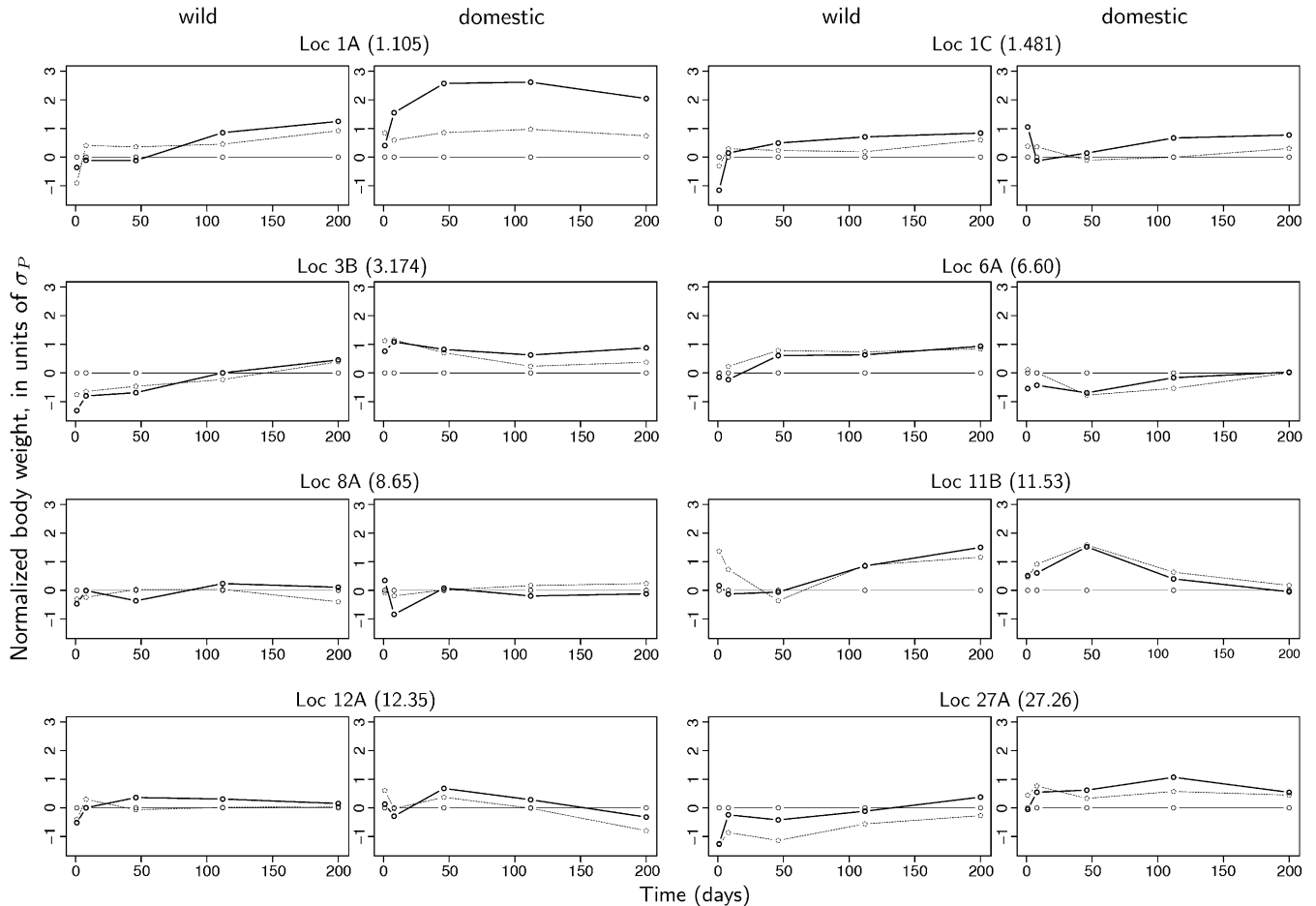


FIGURE 2.—Genetic effects of individual loci in wild and domestic genetic backgrounds. The estimated phenotypic effect of “2” (“Leghorn” or “domestic”) alleles at each of the 8 loci included in the regression analysis in the “wild” (left plots) and “domestic” genetic backgrounds (right plots). The reference is always the “11” (homozygote wild) genotype (solid thin line) in both backgrounds, and plots are scaled according to the residual variance at each age. Additive and dominance effects can be extracted in each plot from the dotted and thick solid lines representing the heterozygote (“12”) and the homozygote domestic (“22”) genotypes, respectively. Epistasis is evident in the comparison of the genetic effects in the alternative genetic backgrounds (if there was no epistasis, right and left plots should be identical). For instance, being 22 instead of 11 at locus 1A increases the body weight at 200 days by $\sim 1.2\sigma_P$ in the “jungle fowl” background (11 genotype at all other loci) and by about $2\sigma_P$ in the “Leghorn” background (22 genotype at all other loci).

2004; CARTER *et al.* 2005; HANSEN *et al.* 2006; TURELLI and BARTON 2006), and due to important progress in methodological, statistical, and computational issues it has been recently confirmed and generalized from empirical data (*e.g.*, CARLBORG and HALEY 2004; MALMBERG and MAURICIO 2005; ZENG *et al.* 2005). However, despite improvements in the quality and the quantity of tools for detection of epistatic interactions, our ability to interpret the output of these QTL analyses in term of biologically relevant genetic effects is still limited. In particular, the statistical models used for QTL detection are based on the average effects of allelic substitutions (and the corresponding variance) in a population. They are therefore suitable for detection of loci, but as the estimates they provide depend on the genotypic frequencies in the particular population, this so-called “statistical epistasis” (CHEVERUD and ROUTMAN 1995)

is of little or no interest for the traditional geneticist. For interpretation, physiological (CHEVERUD and ROUTMAN 1995) or functional (HANSEN and WAGNER 2001) genetic effects are desirable, *i.e.*, the effects of allelic substitutions in a specific genetic background or genotype to draft a genotype–phenotype map. However, these modeling paradigms have often been intermixed or misunderstood in the literature and only recently a suitable mathematical tool has been developed to transform statistical effects (the output of QTL analysis) into functional effects (ÁLVAREZ-CASTRO and CARLBORG 2007).

In a time-series study, the levels and interpretation of epistasis depend on the way the trait is analyzed. A phenotypic measurement, such as the body weight at 46 days of age, considered as an independent trait, has an apparently solid genetic basis with some marginal-effect

loci, as well as several significant interaction effects (see also CARLBORG *et al.* 2003). However, the biological significance of the measured genetic effects for the QTL affecting this trait is despite this unclear, as jungle fowl and White Leghorn chickens, as well as their F₂ progeny, are likely to be at different physiological stages 46 days after hatching. Our analyses clearly show that much epistasis detected for individual traits is due to temporal shifts in the genetic effects of loci from interactions with the genetic background. A potential explanation for this could be that individuals with Leghorn alleles at all eight major loci in the genetic network will be at a later physiological stage than individuals with jungle fowl alleles at all loci. This is a very interesting result that indicates the possibility that a common reason for detected statistical epistasis might actually be due to physiological rather than molecular-level interactions. In future studies it would therefore be highly interesting to measure body weights at, *e.g.*, a physiological stage instead of at a particular age to explore this further. This would most likely decrease the differences in the temporal effects of the loci and hence decrease the general levels of statistical epistasis for the studied traits.

The growth of a complex animal, such as a bird, probably involves hundreds or thousands of genes being active at different stages of development, and at one level or another these are likely to interact. Only a subset of these genes (*i.e.*, the polymorphic ones having a strong allele-substitution effect) can be detected as QTL. Here, we analyzed a large number of alternative descriptors of growth (body weights, growth rates, and parameters linked to the shape of the growth curve as well as orthogonal composite descriptors of all the other traits). In spite of the decomposition of this parameter set into principal components, which are by definition not correlated, it has not been possible to distinguish several groups of genes, *e.g.*, being involved specifically in the final body weight or in determining the shape of the growth curve. Virtually all significant QTL affect the first principal component, related to the overall body weight, and there is a surprising lack of genetic support for individual loci affecting the subsequent PCs, despite a nonnegligible part of the total phenotypic variance being associated with these: the QTL analysis was unable to detect any solid genetic factors for PC2, while PC3 has a few weak but significant underlying genetic factors identical to PC1. Consequently, our results do not support the idea that the growth process could be decomposed into genetically independent parts or modules. The detailed analysis of individual gene effects evidences that the same locus can affect different stages of growth depending on the genetic background, generating strong epistatic interactions. However, experimental evidence is lacking to extend these results to other organisms and to assess whether it might be related to the artificially driven evolutionary history of domesticated species.

We are grateful to Carl Nettelblad for useful discussion on the data analysis, Per Jensen for sharing published phenotypic data, and Leif Andersson for sharing unpublished genetic data and providing useful comments on this work. This work was partially funded by the Swedish Foundation for Strategic Research. Ö.C. acknowledges financial support from the Knut and Alice Wallenberg foundation.

LITERATURE CITED

- ÁLVAREZ-CASTRO, J.-M., and Ö. CARLBORG, 2007 A unified model for functional and statistical epistasis and its application in quantitative trait loci analysis. *Genetics* **176**: 1151–1167.
- ÁLVAREZ-CASTRO, J.-M., A. LE ROUZIC and Ö. CARLBORG, 2008 How to perform meaningful estimates of genetic effects. *PLoS Genet.* **4**: e1000062.
- ANDERSSON, L., C. S. HALEY, H. ELLEGREN, S. A. KNOTT, M. JOHANSSON *et al.*, 1994 Genetic mapping of quantitative trait loci for growth and fatness in pigs. *Science* **263**: 1771–1774.
- BARTON, N. H., and M. TURELLI, 2004 Effects of genetic drift on variance components under a general model of epistasis. *Evolution* **58**: 2111–2132.
- CARLBORG, Ö., and L. ANDERSSON, 2002 The use of randomization testing for detection of multiple epistatic QTL. *Genet. Res.* **79**: 175–184.
- CARLBORG, Ö., and C. HALEY, 2004 Epistasis: Too often neglected in complex trait studies? *Nat. Rev. Genet.* **5**: 618–625.
- CARLBORG, Ö., S. KERJE, K. SCHÜTZ, L. JACOBSSON, P. JENSEN *et al.*, 2003 A global search reveals epistatic interaction between QTL for early growth in the chicken. *Genome Res.* **13**: 413–421.
- CARLBORG, Ö., L. JACOBSSON, P. AHGREN, P. SIEGEL and L. ANDERSSON, 2006 Epistasis and the release of cryptic variation during long-term selection. *Nat. Genet.* **38**: 418–420.
- CARTER, A. J. R., J. HERMISSE and T. F. HANSEN, 2005 The role of epistatic gene interactions in the response to selection and the evolution of evolvability. *Theor. Popul. Biol.* **68**: 179–196.
- CHEVERUD, J. M., and E. J. ROUTMAN, 1995 Epistasis and its contribution to genetic variance components. *Genetics* **139**: 1455–1461.
- CHURCHILL, G. A., and R. W. DOERGE, 1994 Empirical threshold values for quantitative trait mapping. *Genetics* **138**: 963–971.
- COCKERHAM, C. C., 1954 An extension of the concept of partitioning hereditary variance for analysis of covariances among relatives when epistasis is present. *Genetics* **39**: 859–882.
- DOEBLEY, J., A. STEC and C. GUSTUS, 1995 Teosinte branched1 and the origin of maize—evidence for epistasis and the evolution of dominance. *Genetics* **141**: 333–346.
- FISHER, R. A., 1918 The correlation between relatives on the supposition of Mendelian inheritance. *Trans. R. Soc. Edinb.* **52**: 339–433.
- FUMIHITO, A., T. MIYAKE, S. SUMI, M. TAKADA, S. OHNO *et al.*, 1994 One subspecies of the red junglefowl (*Gallus gallus gallus*) suffices as the matriarchic ancestor of all domestic breeds. *Proc. Natl. Acad. Sci. USA* **91**: 12505–12509.
- GOODNIGHT, C. J., 1995 Epistasis and the increase in additive genetic variance: implications for phase 1 of Wright's shifting-balance theory. *Evolution* **49**: 502–511.
- HALEY, C. S., S. A. KNOTT and J. M. ELSE, 1994 Mapping quantitative trait loci in crosses between outbred lines using least squares. *Genetics* **136**: 1195–1207.
- HANSEN, T. F., and G. P. WAGNER, 2001 Modelling genetic architecture: a multilinear theory of gene interaction. *Theor. Popul. Biol.* **59**: 61–86.
- HANSEN, T. F., J.-M. ÁLVAREZ-CASTRO, A. J. R. CARTER, J. HERMISSE and G. P. WAGNER, 2006 Evolution of genetic architecture under directional selection. *Evolution* **60**: 1523–1536.
- JACOBSSON, L., H.-B. PARK, P. WAHLBERG, R. FREDRIKSSON, M. PEREZ-ENCISO *et al.*, 2005 Many QTLs with minor additive effects are associated with a large difference in growth between two selection lines in chickens. *Genet. Res.* **86**: 115–125.
- KOROL, A. B., Y. I. RONIN, A. M. ITS KOVICH, J. H. PENG and E. NEVO, 2001 Enhanced efficiency of quantitative trait loci mapping analysis based on multivariate complexes of quantitative traits. *Genetics* **157**: 1789–1803.

- KEMPTHORNE, O., 1954 The correlation between relatives in a random mating population. *Proc. R. Soc. Lond. Ser. B* **143**: 102–113.
- KERJE, S., Ö. CARLBORG, L. JACOBSSON, K. SCHÜTZ, C. HARTMANN *et al.*, 2003 The twofold difference in adult size between the red junglefowl and White Leghorn chickens is largely explained by a limited number of QTLs. *Anim. Genet.* **34**: 264–274.
- LE ROUZIC, A., and J. M. ÁLVAREZ-CASTRO, 2008 Estimation of genetic effects and genotype-phenotype maps. *Evol. Bioinform.* (in press).
- LE ROUZIC, A., and Ö. CARLBORG, 2008 Evolutionary potential of hidden genetic variation. *Trends Ecol. Evol.* **23**: 33–37.
- LE ROUZIC, A., P. B. SIEGEL and Ö. CARLBORG, 2007 Phenotypic evolution from genetic polymorphisms in a radial network architecture. *BMC Biol.* **5**: 50.
- LIU, Y.-P., G.-S. WU, Y.-G. YAO, Y.-W. MIAO, G. LUIKART *et al.*, 2006 Multiple maternal origins of chickens: out of the Asian jungles. *Mol. Phylogenet. Evol.* **38**: 12–19.
- MA, C.-X., G. CASELLA and R. WU, 2002 Functional mapping of quantitative trait loci underlying the character process: a theoretical framework. *Genetics* **161**: 1751–1762.
- MALMBERG, R. L., and R. MAURICIO, 2005 QTL-based evidence for the role of epistasis in evolution. *Genet. Res.* **86**: 89–95.
- MOISEYEVA, I. G., M. N. ROMANOV, A. A. NIKIFOROV, A. A. SEVASTYANOVA and S. K. SEMYENOVA, 2003 Evolutionary relationships of Red Jungle Fowl and chicken breeds. *Genet. Sel. Evol.* **35**: 403–423.
- R DEVELOPMENT CORE TEAM, 2007 R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna. <http://www.R-project.org>.
- RICE, S. N., 2000 A general population genetics theory for the evolution of developmental interactions. *Proc. Natl. Acad. Sci. USA* **99**: 15518–15523.
- RICHARDS, F. J., 1959 A flexible growth curve for empirical use. *J. Exp. Bot.* **10**: 290–300.
- ROGERS, S. R., G. M. PESTI and H. I. MARKS, 1987 Comparison of three nonlinear regression models for describing broiler growth curves. *Growth* **51**: 229–239.
- RUBIN, C. J., H. BRANDSTROM, D. WRIGHT, S. KERJE, U. GUNNARSSON *et al.*, 2007 Quantitative trait loci for BMD and bone strength in an intercross between domestic and wildtype chickens. *J. Bone Miner. Res.* **22**: 375–384.
- SCHÜTZ, K., S. KERJE, Ö. CARLBORG, L. JACOBSSON, L. ANDERSSON *et al.*, 2002 QTL analysis of a Red Junglefowl \times White Leghorn intercross reveals trade-off in resource allocation between behavior and production traits. *Behav. Genet.* **32**: 423–433.
- TANKSLEY, S. D., S. GRANDILLO, T. M. FULTON, D. ZAMIR, Y. ESHED *et al.*, 1996 Advanced backcross QTL analysis in a cross between an elite processing line of tomato and its wild relative *L-pimpinellifolium*. *Theor. Appl. Genet.* **92**: 213–224.
- TURELLI, M., and N. H. BARTON, 2006 Will population bottlenecks and multilocus epistasis increase additive genetic variance? *Evolution* **60**: 1763–1776.
- WRIGHT, D., S. KERJE, K. LUNDSTROM, J. BABOL, K. SCHÜTZ *et al.*, 2006 Quantitative trait loci analysis of egg and meat production traits in a red junglefowl \times White Leghorn cross. *Anim. Genet.* **37**: 529–534.
- YANG, R., Q. TIAN and S. XU, 2006 Mapping quantitative trait loci for longitudinal traits in line crosses. *Genetics* **173**: 2339–2356.
- ZENG, Z. B., T. WANG and W. ZOU, 2005 Modeling quantitative trait loci and interpretation of models. *Genetics* **169**: 1711–1725.

Communicating editor: C. HALEY