A complex multi-locus, multi-allelic genetic architecture underlying the long-term selection-response in the Virginia body weight line of chickens

Yanjun Zan^{1*}, Zheya Sheng^{2,3*}, Lars Rönnegård^{3,5,6}, Christa F. Honaker⁴, Paul B. Siegel⁴ and Örjan Carlborg^{1†}

Affiliations:

¹Department of Medical Biochemistry and Microbiology, Uppsala University, Uppsala, Sweden

²Key Laboratory of Agricultural Animal Genetics, Breeding and Reproduction of Ministry of Education, College of Animal Science and Technology, Huazhong Agricultural University, Wuhan, People's Republic of China.

³Department of Clinical Sciences, Swedish University of Agricultural Sciences, Uppsala, Sweden

⁴Department of Animal and Poultry Sciences, Virginia Polytechnic Institute and State University, Blacksburg, VA, USA

⁵Department of Animal Breeding and Genetics, SLU, Uppsala, Sweden

⁶Section of Statistics, School of Technology and Business Studies, Dalarna University, Sweden

Author emails:

YZ: Yanjun.Zan@imbim.uu.se

ZS: zheya.sheng@mail.hzau.edu.cn

LR: Lars.Ronnegard@slu.se, lrn@du.se

CFH: chonaker@vt.edu

PBS: pbsiegel@vt.edu

†ÖC: Orjan.Carlborg@imbim.uu.se

*These authors contributed equally

†To whom correspondence should be addressed: orjan.carlborg@imbim.uu.se

1 **Short running title**

2 Multi-locus adaptation in the Virginia chicken lines

3 **Key words**

4 Multi-locus and multi allelic, genetic architecture, long term selection response, Virginia
5 lines, body-weight

6 **Corresponding author**

7 Örjan Carlborg

8 Department of Medical Biochemistry and Microbiology, Uppsala University

9 BMC, Box 582, SE-751 23 Uppsala, Sweden

10 Phone: +46-18-4714592

11 Email: orjan.carlborg@imbim.uu.se

12

13

14

15

16

17

18

19

20

21

22

23

24

25

26

27

28

ABSTRACT

The ability of a population to adapt to changes in their living conditions, whether in nature or captivity, often depends on polymorphisms in multiple genes across the genome. In-depth studies of such polygenic adaptations are difficult in natural populations, but can be approached using the resources provided by artificial selection experiments. Here, we dissect the genetic mechanisms involved in long-term selection responses of the Virginia chicken lines, populations that after 40 generations of divergent selection for 56-day body weight display a nine-fold difference in the selected trait. In the F₁₅ generation of an intercross between the divergent lines, 20 loci explained more than 60% of the additive genetic variance for the selected trait. We focused particularly on seven major QTL and found that only two fine-mapped to single, bi-allelic loci; the other five contained linked loci, multiple alleles or were epistatic. This detailed dissection of the polygenic adaptations in the Virginia lines provides a deeper understanding of genome-wide mechanisms involved in the long-term selection responses. The results illustrate that long-term selection responses, even from populations with a limited genetic diversity, can be polygenic and influenced by a range of genetic mechanisms.

1 INTRODUCTION

2 Experimentally adapted populations are powerful resources for dissecting the genetic basis of
3 adaptation (Hill 2005). These closed populations, subjected to long-term artificial selection
4 for clearly defined adaptive traits, will accumulate adaptive genetic variations at a more rapid
5 rate than natural populations. By mapping individual loci contributing to selection response in
6 these populations, it is possible to gain fundamental insights on the genetic architecture of
7 complex traits and their contributions to adaptation and evolution. As highlighted in Churchill
8 (Churchill 2016), efforts to map quantitative trait loci (QTL) have often revealed an almost
9 overwhelming complexity. Fine-mapping efforts have been challenged by the highly
10 polygenic architecture of traits, where each of the contributing loci often encompass
11 additional complexities, including multiple tightly linked variants with small or large,
12 additive or epistatic genetic effects (; Brandt et al. 2016; Holland 2007; Laurie et al. 2004;
13 Mott and Flint 2013). Therefore, there is a shortage of studies that provide a deeper
14 understanding about the genetic architectures contributing to the adaptive response to long-
15 term selection (Burke 2012).

16
17 Modern genomics allows cost-efficient, in-depth characterization of genetic variation across
18 many loci, or entire genomes, in large populations. This provides opportunities to design
19 studies that make efficient use of selected populations to gain a genome-wide perspective on
20 the genetic architecture of long-term selection response (Burke 2012; Johansson et al. 2010;
21 Pettersson et al. 2013) or to dissect the complexity within and across the loci that contribute
22 to adaptations for highly polygenic traits (Brandt et al. 2016; Chan et al. 2012; Sheng et al.
23 2015). Here, we describe the simultaneous fine-mapping of nine QTL contributing to long-
24 term selection responses in the Virginia chicken lines. These lines have been under bi-
25 directional selection for 56-day body weight (Dunnington and Siegel 1996; Dunnington et al.
26 2013; Márquez et al. 2010). After 40 generations of selection, the High- (HWS) and Low-
27 (LWS) weight selected lines differed 9-fold for the selected trait. In the current generation
28 (S59) the lines differ by more than 16-fold with many loci contributing to this difference

1 (Besnier et al. 2011; Carlborg et al. 2006; Jacobsson et al. 2005; Johansson et al. 2010;
2 Pettersson et al. 2011; Sheng et al. 2015; Wahlberg et al. 2009) with likely contributions by
3 allelic heterogeneity (Besnier *et al.* 2011; Brandt *et al.* 2016) and epistasis (Pettersson et al.
4 2011).

5

6 Reported here is our most recent progress in dissection of the genetic architecture underlying
7 the extreme adaptations in the Virginia body weight chicken lines. Seven major QTL, mapped
8 and confirmed in earlier reports (Besnier et al. 2011; Brandt et al. 2016; Jacobsson et al.
9 2005; Wahlberg et al. 2009), were fine-mapped using data from the F₁₅ generation of an
10 Advanced Intercross Line bred from HWS and LWS founders from generation 40. We
11 performed a multi-locus, association-based fine-mapping analysis, and explored the changes
12 in allele frequencies across the QTL during ongoing selection in the parental lines. Revealed
13 was the genetic basis of several QTL where earlier statistical analyses suggested more
14 complex underlying genetic architectures than a single causal bi-allelic locus (Besnier et al.
15 2011; Brandt et al. 2016). In particular, examples are provided of how tight linkage, multiple
16 alleles from the founder-lines, and epistatic interactions between fine-mapped loci contribute
17 to the QTL effects. In total, seven QTL were fine-mapped into 10 contributing loci, and
18 another 10 loci earlier reported elsewhere in the genome were confirmed (Sheng et al. 2015).
19 Together, these 20 loci explain more than 60% of the additive genetic variance in the
20 population, illustrating how this study provides one of the more comprehensive dissections of
21 a complex adaptive trait in animals. By developing and using new statistical approaches to
22 analyse this novel resource population, we provide insights on the genetic mechanisms
23 contributing to long-term responses to selection that could guide future work to further
24 understanding of the genetic basis of adaptation and evolution.

25

26 RESULTS

27 **Seven of nine QTL were replicated in the F₁₅ generation of the Advanced Intercross**
28 **Line**

Our fine-mapping analysis of nine QTL affecting 56-day body weight in the Virginia chicken lines (Besnier et al. 2011; Jacobsson et al. 2005; Wahlberg et al. 2009;) used genotypic and phenotypic data from the F₁₅ generation of an Advanced Intercross between HWS and LWS founders from generation 40. A multi-locus backward-elimination analysis with bootstrapping to correct for population structure allowed scanning for independent associations across 216 markers in the nine QTL, and 218 markers from selective-sweeps located elsewhere in the genome (Sheng et al. 2015). In total, 24 SNP markers with statistically independent associations to 56-day body-weight were identified at 20% False Discovery Rate (FDR). Of these, 13 were located in seven of the nine QTL (Table 1), and 11 in selective-sweeps elsewhere in the genome (Supplemental Table S1).

11

Table 1. In the nine fine-mapped QTL, 13 SNP markers are associated with 56-day body-weight in generation F₁₅ of the Virginia Advanced Intercross Line. Outside of these QTL, associations to 11 additional markers are detected (Supplemental Table S1).

GGA ^a	Pos ^b	Marker	QTL ^c	Freq		RMIP	
				HWS/LWS40 ^d	P value ^e	5/20% ^f	Sign ^g
1	168703806	rs14916997	G1	1/0 [*]	4.2e-07	0.53/0.55	0.05
1	171538472	rs316102705	G1	1/0.03	1.8e-01	0.36 /0.47	0.20
2	60532699	rs313044994	G2	1/0	1.9e-03	0.62 /0.73	0.05
2	112081184	rs15143460	G3	1/0	1.2e-04	0.84/0.91	0.05
3	33619052	rs15321683	G4	0.9/0 [*]	1.7e-05	0.53/0.57	0.05
3	33880949	rs14339371	G4	1/0.45 [*]	1.9e-02	0.41/0.58	0.20
4	1469998	rs14417942	G6	1/0.05	6.5e-02	0.32 /0.53	0.20
4	11718474	rs314557363	G6	0.98/0	7.1e-02	0.41/0.52	0.20
7	15660310	rs14610961	G9	0.88/0 [*]	4.0e-04	0.41/0.53	0.20
7	16252812	rs15853763	G9	0.65/0	6.2e-03	0.34/0.47	0.20

7	23930693	rs316269048	G9	0.88/0	5.4e-03	0.40/0.59	0.20
7	24032601	rs315233740	G9	1/0*	2.1e-04	0.60/0.74	0.05
20	12756772	rs14280503	G12	1/0*	2.9e-03	0.46/0.60	0.05

^aGGA: *Gallus Gallus Autosome*, ^bLocation in November 2011 (*galGal4*) assembly; ^cQTL mapped and replicated in previous studies (Besnier et al. 2011; Jacobsson et al. 2005; Wahlberg et al. 2009), *G1-G12: Growth1 - Growth12* named as in (Jacobsson et al. 2005); ^dEstimated allele frequencies in the HWS/LWS founders, respectively, from individual SNP chip genotypes or, if this was not available, pooled sequencing data (labelled with *); ^eSignificance from a likelihood ratio test comparing model with or without the tested marker (See Materials and methods); ^fResample Model Inclusion Probability for the marker at 5/20 % FDR thresholds, respectively; ^gFDR threshold reached with an RMIP >0.46 threshold as suggested in Valdar et al (Valdar et al. 2009).

Fine-mapping of the seven replicated QTL

We tested for associations to each of the SNP markers within each of the fine mapped QTL in the AIL F₁₅ generation. This analysis was performed to obtain QTL profiles in this generation for comparison with earlier results from the F₂ generation (Wahlberg et al. 2009). In this analysis, all markers detected at 20% FDR in multi-locus bootstrapping analysis were fitted as fixed effects to account for the multi-locus genetic architecture of 56-day body-weight in this population. The QTL profiles revealed that two QTL, *Growth2* on GGA2 and *Growth3* on GGA2; QTL names as in (Jacobsson et al. 2005), could be fine-mapped to single, bi-allelic loci (Figure 1; Table 1). More complex genetic architectures involving multiple linked loci, segregation of multiple alleles in at least one of the founder lines, or both were revealed in the other five QTL (*Growth1* on GGA1, *Growth4* on GGA3, *Growth6* on GGA4, *Growth9* on GGA7 and *Growth12* on GGA20; Table 1; Figure 2). Described in the sections below are more detailed results from the fine-mapping analyses.

Two QTL fine-mapped to single locus with fixed alleles in the founder lines: The multi-locus bootstrap-based backward-elimination analysis revealed only one 56-day body weight associated marker in each of the two QTL on GGA2 (*Growth2*, *Growth3*; Table 1). This result is consistent with that of the QTL-scan across *Growth2* that identifies a well-defined region with a peak at 60.5Mb (*galGal4*) to the marker retained in the backward-elimination

1 analysis (Figure 1A). The fine-mapped association signal in the F_{15} is located ~5 Mb away
 2 from, and slightly outside of, the QTL-peak identified in the original F_2 analysis (Wahlberg et
 3 al. 2009) (59.4 Mb vs 55Mb in F_{15} and F_2 , respectively; Figure 1A). The additive genetic
 4 effects at the top associated markers in the F_2 and F_{15} are, however, similar (less than 7g
 5 difference; Table 2).
 6

7
 8 **Figure 1.** Statistical support curves from QTL-scans across the QTL Growth2 (A) and Growth3 (B) on GGA2. These QTL were
 9 originally detected in the F_2 intercross generation (purple line) and fine-mapped here in the F_{15} generation (blue line) of an AIL.
 10 The lines connect the significances at the locations of genotyped markers in the two populations, from Wahlberg et al (Wahlberg
 11 et al. 2009) and this study, and the red diamonds highlight the marker retained in the multi-locus bootstrap-based backward-
 12 elimination analysis in the F_{15} generation (Table 1).
 13
 14

Table 2. Estimates of the additive effects for the markers closest to the fine-mapped QTL-peaks in the F_2 (Wahlberg et al. 2009) and F_{15} intercross generations (this study).

Locus ¹	GGA ²	QTL	Markers ³ (F_{15}/F_2)	Position ⁴ (F_{15}/F_2 Mb)	F_2 Additive ⁵ ($a \pm SE$)	F_{15} Additive ⁶ ($a \pm SE$)
1	1	Growth1	rs14916997/rs15502284	168.7/170.2	30.0 $\pm$ 8.5	26.2 $\pm$ 5.8
2	1	Growth1	rs316102705/rs15502284	171.5/170.2	30.0 $\pm$ 8.5	-6.9 $\pm$ 5.9
3	2	Growth2	rs313044994/rs15105085	60.5/59.4	13.3 $\pm$ 8.5	20.9 $\pm$ 7.1
4	2	Growth3	rs15143460/rs16107541	112.1/113.1	17.4 $\pm$ 8.5	21.4 $\pm$ 6.3
5	3	Growth4	rs15321683/rs15321683	33.6/33.6	12.8 $\pm$ 8.6	21.1 $\pm$ 5.6
			rs14339371/rs15321683	33.9/33.6	12.8 $\pm$ 8.6	9.9 $\pm$ 6.5
6	4	Growth6	rs14417942/ADL0143	1.5/1.4	17.1 $\pm$ 8.5	10.1 $\pm$ 6.0
7	4	Growth6	rs314557363/rs14428592	11.7/11.3	29.0 $\pm$ 8.4	10.6 $\pm$ 6.1
8	7	Growth9	rs14610961/rs15852323	15.7/15.5	14.8 $\pm$ 10.5	11.8 $\pm$ 5.9
			rs15853763/rs15852323	16.3/15.5	14.8 $\pm$ 10.5	0.8 $\pm$ 5.7
9	7	Growth9	rs316269048/ADL0279	23.9/23.2	37.5 $\pm$ 8.6	1.8 $\pm$ 5.6
			rs315233740/ADL0279	24.0/23.2	37.5 $\pm$ 8.6	12.7 $\pm$ 5.5
10	20	Growth12	rs14280503/MC3R_PYRO	12.8/12.4	18.1 $\pm$ 8.4	15.1 $\pm$ 5.7

¹Locus: Physically closed markers identified in the bootstrap-based backward-elimination analysis (<1 Mb apart) are clustered into a locus; ²Gallus gallus Autosome; ³Markers for which additive estimates are compared. In the F_{15} , the markers were from the bootstrap-based, backward-elimination analysis and in the F_2 the marker in Wahlberg et al (Wahlberg et al. 2009) with the shortest physical distance (Mb galGal4) to the respective F_{15} markers; ⁴Location of markers in the Nov. 2011 (galGal4) assembly; ⁵Additive genetic effect $\pm$ Standard Error estimated using a model including the genotype for the tested marker and the sex of the bird in the F_2 generation (Wahlberg et al. 2009); ⁶Additive genetic effect $\pm$ Standard Error estimated in F_{15} generation.

The QTL-scan across *Growth3* in the F_{15} generation identifies a single association peak at 112.1 Mb to the same marker detected in the multi-locus analysis (Figure 1B). The association signal is located approximately 2 Mb from the top associated marker in the F_2

(Figure 1B). In *Growth3*, the peak marker in the F_2 population was the closest one among all other peak markers to the fine-mapped location in the F_{15} generation, and the additive effects in the two generations differ only marginally (4g; Table 2). Thus, *Growth2* and *Growth3* QTL could be fine-mapped to narrow chromosomal regions using the data from the AIL F_{15} generation (Figure 1). The 1 LOD drop-off confidence intervals for *Growth2* and *Growth3* were 0.5 Mb (60.2-60.7) and 1.9 Mb (112.0-113.9), respectively.

Two linked loci are revealed in four fine-mapped QTL: In four of the fine-mapped QTL (*Growth1* on *GGA1*, *Growth6* on *GGA4*, *Growth9* on *GGA7* and *Growth12* on *GGA20*; Figure 2; Table 1), significant associations were detected to distant markers (> 2.5 Mb apart) in either the QTL-scan, the backward-elimination analysis or both in the F_{15} generation data. In addition to this, pairs of physically close markers (< 1 Mb apart) were detected in three of these fine-mapped loci in two QTL (*Growth4* on *GGA3* and *Growth9* on *GGA7*) in the backward elimination analysis. Below, we describe the dissection of these associations in more detail.

Figure 2. Statistical support curves from the QTL-scans across five QTL associated with 56-day body weight in the F_2 intercross generation (purple) (Wahlberg et al. 2009) and fine mapped in the F_{15} generation of the AIL (blue; this study). The plots show (A) Growth1 on GGA1, (B) Growth4 on GGA3, (C) Growth6 on GGA4, (D) Growth9 on GGA7, and (E) Growth 12 on GGA20. The red diamonds indicate the significant markers in the bootstrap-based backward-elimination analysis in the F_{15} generation (Table 1).

A major QTL, *Growth1*, was detected on GGA1 in the F_2 intercross (Jacobsson et al. 2005; Wahlberg et al. 2009). The QTL-scan in the F_{15} generation of the AIL suggests that two loci, tagged by the markers *rs14916997* ($p = 2.1 \times 10^{-6}$; likelihood ratio test) at 168.7 Mb and *rs14924102* ($p = 9.5 \times 10^{-5}$; likelihood ratio test) at 175.6 Mb, contribute to the F_2 QTL (Figure 2A). The association to the marker *rs14916997* at 168.7 Mb was also detected in the multi-locus bootstrap-based backward elimination analysis ($p = 4.2 \times 10^{-7}$; likelihood ratio test; 5% FDR level). Although the multi-locus backward-elimination analysis also detected a second marker in this QTL, it was not the same one as in the QTL-scan, but to marker *rs316102705* at 171.5 Mb ($p = 1.8 \times 10^{-1}$; likelihood ratio test; 20% FDR; Table 1). A

possible explanation for why an association is present in the QTL-scan, but not in the bootstrap-based backward-elimination analysis, is that its effects are dependent on the presence of alleles at other loci. This epistasis hypothesis is further explored in a separate section below.

5

The *Growth6* QTL extended across a long region of GGA4 in the F₂ analysis (Jacobsson et al. 2005; Wahlberg et al. 2009) (Figure 2C). In the AIL F₁₅, the QTL-scan (Figure 2C) and the multi-locus bootstrap-based backward-elimination analysis (Table 1) identify the same two regions to be associated with 56-day body-weight in the AIL F₁₅ generation. In generation 40, the first region is centered around a marker at the proximal end of the QTL (*rs14417942* at 1.5 Mb; $p = 6.5 \times 10^{-2}$; likelihood ratio test) that is fixed for one allele in the HWS, and nearly fixed for the opposite allele in LWS (Table 1). At generation 50, this major LWS allele was fixed. The second association was to a marker located at the distal end of the QTL (*rs314557363* at 11.7 Mb; $p = 7.1 \times 10^{-2}$; likelihood ratio test) that was fixed for alternative alleles in the HWS and LWS lines at generation 40 (Table 1). The F₂ *Growth6* QTL extends across these two fine-mapped loci in the F₂, with the main peak in the middle (Figure 2C). The fine-mapping in the F₁₅ therefore indicates that *Growth6* was a ghost-QTL (Knott and Haley 1992) in the F₂ generation caused by the extended LD in this population.

19

The most significant QTL in the F₂, *Growth9*, is located on GGA7 (Figure 2D). In the AIL F₁₅, the QTL-scan and bootstrap-based backward-elimination analyses identified, 8 Mb apart, the same two independent loci in this QTL. The backward-elimination analysis detected associations to two markers in each of these loci (Figure 2D; Table 1). The proximal fine-mapped locus (16 Mb GGA7; galGal4) is located on the border of *Growth9*, whereas the distal locus (24 Mb GGA7; galGal4) is located in the middle of the QTL. The first peak at 16 Mb included associations to *rs14610961* (15.6 Mb; $p = 4.0 \times 10^{-4}$; likelihood ratio test) and *rs15853763* (16.2 Mb; $p = 6.2 \times 10^{-3}$; likelihood ratio test). In generation 40, both of these markers were segregating in HWS (Table 1). The second peak at 24 Mb included associations

1 to *rs316269048* (23.9 Mb; $p = 5.4 \times 10^{-3}$; likelihood ratio test) and *rs315233740* (24.0 Mb; p
2 $= 2.1 \times 10^{-4}$; likelihood ratio test). The first of these markers segregated in HWS in generation
3 40 (Table 1). A possible explanation for why associations are detected to multiple physically
4 close markers (< 1 Mb apart) is that together they tag haplotypes with different effects that
5 segregate at these loci. We describe how this hypothesis was tested for the loci where such
6 associations were detected in a separate section below.

7

8 The *Growth12* QTL was mapped to GGA20 in the F_2 population (Jacobsson et al. 2005;
9 Wahlberg et al. 2009). The QTL-scan in the AIL F_{15} generation detected associations to
10 markers in two linked loci on in this locus: *rs14278283* (10.5 Mb; $p = 1 \times 10^{-3}$; likelihood
11 ratio test) and *rs14280503* (12.8 Mb; $p = 3 \times 10^{-3}$; likelihood ratio test) (Figure 2E). Only the
12 association to *rs14280503* (5% FDR; Table 1) was significant in the bootstrap-based,
13 backward-elimination analysis. Epistasis is a possible explanation for this finding and this
14 was (for the locus detected in *Growth1*) tested in a separate section below.

15

16 **Epistatic interactions between markers associated with 56-day body-weight:** As
17 mentioned above, epistasis is a possible explanation for why some associations detected in the
18 QTL-scan are not detected in the bootstrap-based backward-elimination analysis. This could
19 be a result if the marginal effects of epistatic loci are dependent both on the allele-frequency
20 at the locus itself and on the allele-frequency at other interacting loci elsewhere in the
21 genome. As a consequence of this, the marginal additive effects of epistatic loci will be more
22 sensitive to the bootstrapping procedure and hence potentially selected in fewer of the
23 resampled populations. We therefore tested for epistatic interactions between the loci that
24 were significant in the single marker association scan, but not in the bootstrap-based
25 backward-elimination analysis. The pair *rs14924102* (in *Growth1*) and *rs14278283* (in
26 *Growth12*) displayed a significant epistatic interaction ($p = 8.9 \times 10^{-3}$). Their additive effects
27 were also significant when fitted together with the 24 markers retained in the backward-
28 elimination with bootstrapping analysis ($p = 3 \times 10^{-3}$ for *rs14924102* and $p = 2.2 \times 10^{-4}$ for

1 *rs14278283*; Table 1, Table S1). The allele-frequencies at the marker *rs14924102* had,
2 however, drifted from 0.5 during the breeding of the AIL ($p = 0.13$ in AIL F₁₅). This resulted
3 in a small number of observations in the two-locus genotype-classes including the minor-
4 allele (A) at this locus ($n = [3,11,8]$ for the [AAAA, AATA, AATT] genotypes). Hence,
5 although there was a significant statistical interaction between these two loci, this result
6 should be interpreted with caution: the epistatic interaction is due to the few individuals with
7 the minor-allele homozygote genotype at this locus having opposite effects to those in the
8 more frequent genotypes at this locus (AAAA and AATT; Supplemental Figure S1).

9

Explorations of three fine-mapped loci in two QTL that segregate for multiple alleles:

11 The backward-elimination analysis detected associations to pairs of physically close markers
12 (< 1 Mb apart) at three loci in two QTL (*Growth4* and *Growth9*). In *Growth4* on GGA3 the
13 associated markers are located 0.2 Mb apart and in *Growth9* on GGA7 they are located 0.6
14 and 0.1 Mb apart, respectively (Table 1). Because at least one marker in each pair segregate
15 within one of the founder lines (Table 1), in practice, their two-locus genotypes act as multi-
16 allelic markers in the AIL. This allows tagging the haplotypes that segregate within the
17 founder-lines to these markers and estimating their individual haplotype effects.

18

19 The proximal locus in the *Growth9* QTL on GGA7 was fine-mapped to two SNP markers 593
20 kb apart: *rs14610961* (15.6 Mb; $p = 4.0 \times 10^{-4}$; Table 1) and *rs15853763* (16.3 Mb; $p = 6.2 \times$
21 10^{-3} ; Table 1). Both markers segregated for two alleles in HWS. Identified were three
22 haplotypes across these two loci in the F₁₅ population that segregated at frequencies > 0.1
23 ($freq_{AA} = 0.35$, $freq_{GA} = 0.10$, and $freq_{GG} = 0.55$). In the founders of this pedigree, that were
24 from generation 40 of the selected lines, the LWS was fixed for GG, whereas HWS segregates
25 for AA, AG and GG at frequencies $freq_{AA} = 0.65$, $freq_{AG} = 0.175$ and $freq_{GG} = 0.175$,
26 respectively (Supplemental Figure S2; Supplemental Table S2). Fitting a two-locus
27 haplotype-ANOVA to these markers, while correcting for the effects of the other 22 selected
28 markers (Table 1; Supplemental Table S1), did not reveal any significant associations

1 between these multi-marker genotypes and 56-day body weight in the F_{15} ($p = 0.08$). That
2 these marker haplotypes do, however, not fully capture the allelic complexity in the HWS and
3 LWS lineages at generation 40 and 53 (Supplemental Figure S2), suggests that the region
4 should be evaluated further using more informative markers to understand which haplotypes
5 contribute to the multi-marker association at this locus.

6

7 The distal locus in the *Growth9* QTL was fine-mapped to two SNPs located 102 kb apart on
8 GGA7: *rs316269048* (23.9 Mb, $p = 4.0 \times 10^{-4}$; Table 1) and *rs315233740* (24 Mb, $p = 6.2 \times$
9 10^{-3} ; Table 1). The SNP *rs316269048* segregates for two alleles in HWS. The other SNP
10 *rs315233740* is fixed for alternative alleles between HWS and LWS. Three haplotypes
11 segregate at frequencies > 0.1 in the AIL F_{15} population ($freq_{AA} = 0.40$, $freq_{CA} = 0.10$, and
12 $freq_{CT} = 0.48$). In the founder lines at generation 40, the LWS is fixed for the CT haplotype,
13 whereas HWS segregates for CT and AA at $freq_{CT} = 0.125$ and $freq_{AA} = 0.875$, respectively
14 (Supplemental Figure S3; Supplemental Table S2). The CA haplotype in the F_{15} , that was not
15 observed in the founders, is likely a recombinant between the CT and AA haplotypes. The
16 two-locus haplotype-ANOVA detected a significant association to 56-day weight in the F_{15}
17 AIL population at this locus ($p = 0.008$). However, as the CT haplotype tags three different
18 haplotypes that segregate at this locus in LWS at generation 40 (Supplemental Figure S3;
19 Supplemental Table S2), and few individuals in the F_{15} carried each haplotype, the testing of
20 individual haplotype effects were underpowered. Additional data are needed to explore in
21 further detail the allelic heterogeneity at this locus.

22

23 The *Growth4* QTL on GGA3 was fine-mapped to two SNP markers 260 kb apart: *rs15321683*
24 (33.6 Mb, $p = 1.7 \times 10^{-5}$; Table 1) and *rs14339371* (33.8 Mb, $p = 1.9 \times 10^{-2}$; Table 1, Figure
25 2B). The marker *rs15321683* was fixed for alternative alleles in HWS and LWS at generation
26 40. The marker *rs14339371* segregated for two alleles in LWS. Three haplotypes segregated
27 at this locus in the F_{15} generation at frequencies > 0.1 ($freq_{AG} = 0.59$, $freq_{GA} = 0.29$, and $freq_{GG}$
28 $= 0.12$). The AG haplotype was almost fixed in HWS ($p = 0.98$), whereas GA and GG

segregated in LWS at ($freq_{GA} = 0.61$ and $freq_{GG} = 0.29$, respectively; Table 3) at generation 40. The two-locus haplotype ANOVA detected a significant association to 56-day weight in the F_{15} AIL population in *Growth4* ($p = 0.001$). The individual haplotype frequencies were sufficiently high in *Growth4* to allow an association analysis to the individual haplotypes. This analysis revealed that the segregating haplotypes have different effects on 56-day body-weight in the AIL F_{15} generation (Table 3). The major LWS haplotype at generation 40 (GA) decreased body-weight by 14.8g ($p = 0.039$; Table 3; Figure 3A) compared to the nearly fixed HWS haplotype (AG). The minor LWS haplotype (GG) decreased 56-day weight even more (-44.3g; $p = 0.0003$; Table 3; Figure 3A).

Table 3. Estimated effects and allele-frequencies for the haplotypes segregating in *Growth4* on *GGA3*

Haplotype	Haplotype effects in F_{15} AIL		Haplotype frequencies			
	a ± SE ¹	p ²	F_{15}	HWS40	LWS40	LWS53
AG (HWS)	0	-	0.59	0.98	0.05	0
GA (LWS)	-14.8 ± 6.9	0.039	0.29	0	0.61	0.28
GG (LWS)	-44.3 ± 12.2	0.00031	0.12	0	0.29	0.62

¹Additive effect of the haplotype (g) ± Standard Error; ²Significance for the additive effect of the haplotype

We continued to explore the high-density SNP haplotypes in the HWS and LWS lineages across the region of *Growth4* that contained the two significantly associated linked SNP markers (GGA4: 32.1–35.9 Mb) in detail. The major two-locus HWS haplotype (AG) and the minor LWS haplotype (GG) at generation 40 each tag well-defined haplotypes across the region (Figure 3B; Figure S4). The major two-marker LWS-haplotype (GA) at generation 40, however, tags two different haplotypes. During selection for decreased body-weight in LWS from generation 40 to 53, the frequency of the haplotype with the lowest effect on 56-day weight (GG) increased (Figure 3B, Supplemental Figure S4). By generation 53, it has become the major haplotype in the LWS line (Table 2; Supplemental Figure 3B). This suggests an

- 1 ongoing selection at this QTL at generation 40, leading to fixation of the haplotype with the
2 least effect on 56-day body-weight by generation 53.

3
4 **Figure 3.** Schematic illustration of the haplotype-structure around markers rs15321683 (33.6 Mb; galGal4) and rs14339371
5 (33.8 Mb; galGal4) on GGA3, and how the founder-haplotypes that they tag contribute to 56-day body-weight in the AIL F₁₅
6 generation. (A) At this locus, the individuals segregate for three major haplotypes that have distinct effects on 56-day body
7 weight in the AIL F₁₅ generation. The figure illustrates the regressions obtained for individuals carrying different haplotypes in
8 the AIL F₁₅ generation (the error-bars indicate the standard-errors of the estimates). (B) A schematic illustration of the major
9 haplotypes segregating in the region from 32.9-34.6 Mb on GGA3 (a complete visualization of the haplotype structure based on
10 all SNPs genotyped on the 60K SNP chip is provided in Supplemental Figure S4). The HWS line is fixed for a weight-increasing
11 haplotype AG, while the LWS line segregates for two weight-decreasing haplotypes (GA, GG) in the QTL Growth4 on GGA3 at
12 generation 40. The haplotype in the LWS lineage with the strongest body weight decreasing effect tags a single haplotype that
13 increases in frequency during selection from generation 40 to 53.

Variance explained by the fine-mapped loci

We estimated how much of the additive genetic variance for 56-day body weight in the AIL F_{15} population that could be explained by the associated markers in the fine-mapped QTL (Table 1) and those identified in selective-sweep regions elsewhere in the genome (Sheng et al. 2015) (Supplemental Table S1). Together 24 markers, representing 20 loci, retained in the backward-elimination analysis (20 % FDR) explained 27.8% of the residual phenotypic variance, corresponding to 60.5% of the additive genetic variance in the population.

8

9

DISCUSSION

In this study, we have continued to dissect the genetic architecture underlying the highly polygenic trait, 56-day body-weight, in the Virginia chicken population that has been subjected to long-term, divergent single-trait selection. With a focus on the F_{15} generation of an AIL between the lines, we fine-mapped and explored the genetics of nine major QTL that were mapped in the F_2 (Jacobsson et al. 2005; Wahlberg et al. 2009) and replicated in the F_2 to F_8 generations (Besnier et al. 2011; Brandt et al. 2016) The increased resolution in the F_{15} population allowed us to fine-map several of the major QTL to contributions by linked loci, and sometimes multiple haplotypes within and epistatic interactions across these loci. These findings further emphasize the important contribution by many loci with small additive effects to adaptation (Jacobsson et al. 2005; Lango Allen et al. 2010; Wood et al. 2014), where some of them are tightly linked (Iraqi *et al.* 2000). We also found that one of the original QTL detected in the F_2 population (*Growth6* on GGA4) was likely a ghost-QTL (Haley and Knott 1992).

23

Our backward-elimination based mapping approach revealed associations to markers that were physically close but segregated for multiple alleles in at least one of the founder lines. By utilizing haplotype information on the founders, we confirmed that these markers tag segregating haplotypes with distinct effects on the traits and provide insights to the potential importance of allelic heterogeneity (Forsberg et al. 2015; McClellan and King 2010; Yano et

1 al. 2016) for the response to long-term selection in the Virginia lines. Because the current
2 population is rather small, the haplotype-frequencies have drifted away from 0.5 during the
3 breeding of the deep intercross population (minor allele frequency ranging from 0.03 to 0.5,
4 with median of 0.39). In addition, the markers in this study were not initially selected to tag
5 the segregating haplotypes, resulting in low statistical power to test or estimate the effects of
6 all individual haplotypes. Further studies to explore these multi-allelic loci with more
7 informative SNP markers would help disentangle the likely important contributions by allelic
8 heterogeneity in the fine-mapped loci to the long-term selection response.

9

10 Previous studies of the Virginia body-weight selected populations revealed and replicated
11 extensive epistatic interactions among QTL alleles (Carlborg et al. 2006; Le Rouzic et al.
12 2008). Unfortunately, the allele-frequencies were too imbalanced and the number of
13 individuals too small in this F_{15} dataset to provide sufficient statistical power for an extensive
14 exploration of epistasis. Therefore, we limited testing of epistasis to explore the possibility
15 that two of the loci found to be sensitive to the genetic background in the bootstrap analysis
16 are involved in an interaction.

17

18 The 24 markers in the fine-mapped QTL and selective sweeps outside QTL regions explained
19 27.8 % of the residual phenotypic variance or 60.5% of the additive genetic variance in the
20 F_{15} population. This is considerably more than the 13% explained by the single significant
21 and 12 suggestive QTL in the original F_2 line-cross analysis (Jacobsson et al. 2005). In the F_2 -
22 F_8 generations of the AIL, the five most significant QTL explained 10% of the residual
23 phenotypic variance (Besnier et al. 2011). Our dissection of this adaptive trait in this long-
24 term selection experiment has thus revealed loci that explain nearly two thirds of the additive
25 genetic variance in this population. Although a large fraction of this variance is contributed
26 by loci outside the QTL (Sheng et al. 2015), the linked and segregating loci revealed in the
27 fine-mapping of the QTL suggest that much genetic variance has been released due to
28 recombination break-up of unfavorable linkages in the selected lines during the breeding of

1 the AIL. One example of this is the mapping of transgressive alleles in the F_{15} (e.g. between
2 *rs316102705* and *rs14916997* in *Growth1* on GGA1; Table 2). These unfavorable linkages
3 will decrease the genetic variance explained by the QTL in the F_2 where the LD extends over
4 both these loci.

5

6 For the QTL that were fine-mapped to a single bi-allelic locus, there was an overall
7 agreement between the additive effects estimated in the F_2 and F_{15} generations (Table 2). This
8 was consistent with expectations that there should be less bias in the estimates of QTL effects
9 when sufficiently large QTL mapping populations are used (Beavis 1998). For the QTL with
10 more complex genetic architectures, including those with linked loci and multiple alleles, the
11 deviation was larger, as the allele substitution effect at a single marker is not sufficient for
12 representing such effects (Table 2). The degree to which these estimates differ will depend,
13 for example, on the linkage phase of the linked loci and the frequencies of the respective
14 alleles at the multi-allelic locus. Our results show that overall the marginal effects estimated
15 for the F_2 QTL agree with the joint estimated effects for the fine-mapped loci and alleles in
16 the F_{15} generation. Further work is needed to explore whether the remaining deviations are
17 due to still unresolved complexities for these loci; for example incomplete tagging of
18 segregating alleles in the founders, multi-allelic genetic architecture, and/or epistatic
19 interactions with other loci.

20

21 The ultimate aim for studies designed to dissect the genetic architecture of a complex trait is
22 to find the causal genes and mutations underlying the trait. Although the use of a deep
23 intercross population has allowed us to fine-map the major QTL segregating in this
24 population to 1-2 Mb resolution, the limitation of long LD blocks in linkage mapping analysis
25 still limited our ability to achieve this aim. Additional research is necessary to identify
26 haplotypes segregating within and across the selected lines and AIL, to reduce the number of
27 candidate genes in the 24 identified loci. Regardless, this provided valuable genome-wide
28 insights to the variety of genetic mechanisms that contribute to the genetic variance for a

1 single trait for which dramatic adaptations have emerged during long-term directional
2 selection.

3

4 In conclusion, this study illustrates that the strong response to long-term divergent selection in
5 the Virginia body weight lines rely on a highly polygenic and complex genetic architecture.
6 Most of the original major QTL in this population fine-map to contributions by linked, and
7 sometimes multi-allelic or epistatic loci. By continuing the in-depth characterizations of the
8 highly polygenic genetic architecture of this adaptive trait, we provide insights for future
9 research aiming to understand the variety of genetic mechanisms that together are likely to
10 contribute to adaptation and evolution. It demonstrates that, even in relatively small, closed
11 populations, selection response is likely to involve a wide variety of the complex genetic
12 mechanisms that together contribute to quantitative trait variation and adaptation.

13

14

MATERIALS AND METHODS

15

Ethics statement

16

All procedures involving animals used in this experiment were carried out in accordance with
17 the Virginia Tech Animal Care and Use Committee protocols.

18

19

Animals, phenotyping and DNA extraction

20

The chickens used in this study were from the F₁₅ generation of an Advanced Intercross Line
21 bred from the Virginia high (HWS) and low (LWS) body weight selected lines. The HWS and
22 LWS lines were founded in 1957 from a common base population obtained by crossing seven
23 partially inbred lines of White Plymouth Rock chickens. They have since then been subjected
24 to bi-directional selection for high or low 56-day body weight, respectively. For further
25 details on the Virginia lines, see Dunnington et al (Dunnington and Siegel 1996; Dunnington
26 et al. 2013). An Advanced Intercross Line (AIL) was founded from generation 40 of the HWS
27 and LWS, where the sex-average 56-day body-weights were 1,412 g (SE: ± 36 g) and 170 g
28 (SE: ± 5 g), respectively (Jacobsson et al. 2005). This AIL has previously been used in the

fine-mapping of QTL using the F₂-F₈ AIL generations (Besnier et al. 2011; Brandt et al. 2016; Pettersson et al. 2011). In the AIL F₁₅ generation, 907 individuals were hatched and 56-day body weights and genotypes measured on 825 (Genotype and phenotype data are available as supplemental data 1 and 2). These birds were used in an association analyses to evaluate contributions by genome-wide selective sweeps to 56-day body-weight the F₁₅ generation (Sheng et al. 2015). This study is based on the same individual F₁₅ individuals as in Sheng et al (Sheng et al. 2015).

Marker selection and genotyping

In total, 434 SNP markers, genotyped using a GoldenGate genotyping assay (Illumina Inc; performed at the SciLifeLab SNP&SEQ Technology Platform at Uppsala University), were included in this study. Of these, 252 markers were genotyped and used in an earlier study of the AIL F₁₅ generation (Sheng et al. 2015) to target 99 regions in the genome where the HWS and LWS lines were fixed for alternative alleles after 50 generations of selection. These 99 regions cover about 138 Mb of the 1.2 Gb in the chicken reference genome (galGal4). For this study, we genotyped an additional 182 SNP markers selected to provide dense marker coverage across nine QTL known to affect body weight in the F₂ (Carlborg et al. 2006; Jacobsson et al. 2005; Wahlberg et al. 2009) and F₂-F₈ intercross populations (Besnier et al. 2011; Brandt et al. 2016; Pettersson et al. 2011). Together the 434 markers covered approximately 300 Mb of the chicken genome (galGal4), of which 170 Mb represent targeted QTL regions (in total 216 markers), and 130 Mb (218 markers) selective-sweep regions (Sheng et al. 2015). The Nov 2011 (galGal4) chicken genome assembly was used for comparing physical locations of mapping results from the F₂ (Wahlberg et al. 2009) and the F₁₅ intercross generations.

Coding of SNP marker alleles in the AIL F₁₅ generation

The GoldenGate assay reports SNP marker alleles on a [A, T, C, G] basis. Before the statistical analysis, we re-coded the marker genotypes in individuals as [-1, 0, 1]. The coding

1 -1/1 was used to represent homozygotes for the most common alleles in the LWS/HWS
2 founders, respectively. The code 0 was used to represent heterozygotes. The LWS/HWS
3 origin was determined by comparing the F_{15} genotypes to those of the HWS/LWS
4 F_0 founders. By using this coding, positive estimates of additive effects in the association
5 analysis indicate that predominant allele in HWS increased weight, and negative additive
6 effects that the predominant allele in LWS increased weight.

7

8 Multi-locus association analysis using a backward-elimination strategy with 9 bootstrapping to correct for population structure

10 The statistical analyses were designed to simultaneously fine-map nine QTL regions
11 contributing to 56-day body-weight in the Virginia body weight lines, while also accounting
12 for the effects of other regions across the genome. This was to appropriately account for the
13 highly polygenic genetic architecture of body-weight in this population (Besnier et al. 2011;
14 Jacobsson et al. 2005; Johansson et al. 2010; Pettersson et al. 2013; Sheng et al. 2015), where
15 many of the genotyped markers both in and outside the QTL regions should contribute to the
16 trait. As all individuals in the AIL were progeny of dams of the same age, hatched on the
17 same date, and reared separate from their parents, the environmental contributions to
18 between-family means in the F_{15} population should be minimal. We therefore assume that a
19 large portion of the differences in family means are due to the joint effects of the markers in
20 the genotyped QTL and selective sweep regions, rather than non-genetic effects. Population
21 structure might, however, still be of concern in a deep-intercross population (Peirce et al.
22 2008; Cheng et al. 2010). We therefore also validate our results using a bootstrap-based
23 approach developed to account for the possible effects of population structure in general
24 deep-intercross populations, including AILs (Valdar et al. 2009).

25

26 Not known in advance was how many genetic markers in the fine-mapped QTL contributed to
27 56-day body-weight in the AIL F_{15} generation. Based on previous findings (Besnier et al.
28 2011; Brandt et al. 2016), we expected that at least some QTL would contain multiple linked

1 loci. The number of loci in the final model could therefore vary substantially, and to account
2 for this we implement our analysis with an adaptive model selection criterion controlling the
3 False Discovery Rate (FDR) (Abramovich et al. 2006; Gavrilov et al. 2009) developed for
4 this purpose. Starting with the 434 genotyped SNP markers, we then selected a final set of
5 markers using a two-step backward-elimination strategy applied on a standard linear model
6 with 56-day body-weight as response variable. A 20 % FDR among the selected loci was used
7 as the termination criterion.

8

9 In the first step, a linear model including the SNP markers to be tested, together with the fixed
10 effects of sex of the individual and the genotypes of the 16 sweeps that were earlier found to
11 be associated with 56-day weight in this population (Sheng et al. 2015). The 16 earlier
12 mapped sweeps were included in the model to capture the major polygenic effect in this
13 population. To avoid over parameterizing the linear model, the 434 markers were divided into
14 eight pre-screening sets including approximately 50 markers each. These sets were selected
15 to, as far as possible, include markers from the same chromosomes. A backward-elimination
16 analysis with the adaptive FDR criterion was performed on each set separately. All markers
17 that reached the 20% FDR in either of these analyses were kept for the further joint analyses
18 described in the next section.

19

20 In the second step of the analysis, all markers selected in the pre-screening were analyzed
21 jointly using the bootstrap based method of Valdar et al (Valdar et al. 2009). This analysis
22 was used to identify the loci that contribute to 56-day body weight in this population and that
23 was robust to the possible effects of population structure. Here, a RMIP (Resample Model
24 Inclusion Probability) was calculated for each marker tested. A final model was selected
25 where only markers with $RMIP > 0.46$, the threshold suggested for an AIL generation
26 F_{18} (Valdar et al. 2009), was included. All of these analyses were implemented in custom
27 scripts using the statistical software R (R Core Team 2015).

28

1 Estimation of the additive effects and significances of the markers associated with 56- **2 day body-weight in the F₁₅ population**

3 The general linear model used to estimate additive effects and significances of the loci
4 associated with 56-day body-weight in this population can be formulated as

5

$$6 \quad Y = X_1\beta_1 + e \quad (1)$$

7

8 Here, Y is a column vector containing the 56-day body-weight of the 825 F₁₅ individuals. β_1
9 is a vector with the estimate of the fixed effect of sex and the additive effect of all markers
10 tested. X_1 is the design matrix including the coding for the sex of the birds and the genotypes
11 of the markers coded as [-1,0,1]. e is the normally distributed residual. The markers included
12 to control for the background genetic effects will vary depending on the analysis as described
13 in the sections below.

14

15 **Analyses of markers selected in the backward-elimination analysis:** There are few
16 recombination events within each QTL in the AIL pedigree which limits the resolution in the
17 fine-mapping analysis. In the backward-elimination analysis, we therefore consider markers
18 located closer than 1 Mb from each other as representing the same fine-mapped locus. The
19 effects of all markers in these loci were estimated using the model described above (model 1)
20 by defining X_l to include only the genotypes of the tested marker from that locus and the
21 SNPs in the remaining loci. This meant that when estimating the effect of a marker located
22 inside a fine-mapped locus tagged by more than 1 marker, the remaining markers inside this
23 locus were excluded from the analysis. The significance for each marker was obtained using a
24 likelihood ratio test comparing regression models (model 1) with and without the tested
25 marker.

26

27 **Single marker association scans in the replicated QTL:** To obtain statistical support

1 profiles for the QTL with significant associations in the backward-elimination analysis, we
2 performed single marker association scans across all markers in these QTL. The same
3 regression model (model 1) was used, but X_1 here include the genotype of the tested marker
4 together with the sex effect and the additive effects of all SNPs selected in the backward-
5 elimination analysis.

6

7 **Effects of the QTL detected in the F_2 population:**

8 The additive effects and significances for QTL identified in the F_2 population were extracted
9 at the location of the marker closest to the reported QTL-peak from the results in (Wahlberg
10 et al. 2009).

11

12 **Calculation of founder-line allele-frequencies for the markers associated with body-** 13 **weight in the F_{15} population**

14 Previously, we genotyped 20 individuals from each of the HWS and LWS lines from
15 generation 40 using a 60k SNP chip (Jacobsson et al. 2005; Pettersson et al. 2013). Some of
16 the SNPs evaluated here were included in that study and for those SNPs those data were used
17 to calculate the allele-frequencies in the founder-lines. For the remaining SNPs, the allele-
18 frequencies in the founder-lines were estimated using sequence data (~30X coverage) from
19 two pools including 29 HWS founders and 30 LWS founders, respectively (Supplemental
20 data 3 and 4).

21

22 **Testing for pairwise epistatic interaction between loci**

23 It is possible that association peaks that are significant in the scan across the QTL in the
24 original data, but not in the bootstrap-based backward-elimination analysis, could be lost due
25 to sensitivity to allelic background at other loci (i.e. epistasis). We therefore tested for
26 epistasis between such peaks, using a likelihood ratio test, by evaluating the significance of
27 the pairwise interactions in model 2:

28

$$Y = \mu + a_1\alpha_1 + a_2\alpha_2 + d_1\delta_1 + d_2\delta_2 + a_1a_2I_{\alpha\alpha} + a_1d_2I_{\alpha\delta} + d_1a_2I_{\delta\alpha} + d_1d_2I_{\delta\delta} + e \quad (2)$$

2

3 where Y is the residual from a regression model (model 1) fitting all significant markers
 4 selected in the bootstrap-based backward-elimination analysis together with the sex of the
 5 chickens. μ is the reference point, α_1 and α_2 (δ_1 and δ_2) are the additive allele-substitution
 6 effects (dominance deviations) at the two loci, $I_{\alpha\alpha}$, $I_{\alpha\delta}$, $I_{\delta\alpha}$ and $I_{\delta\delta}$ are the additive by additive,
 7 additive by dominance, dominance by dominance and dominance by dominance interaction
 8 effects, and e is the normally distributed residual. The a^* and d^* are the NOIA indicator
 9 regression variables for the two loci. These indicator regression variables are allele-frequency
 10 weighted codings of the genotypes at the evaluated loci and calculated as detailed in
 11 (Alvarez-Castro and Carlborg 2007). Using the orthogonal statistical NOIA parameterization
 12 of the model, we tested for epistasis by comparing the full model (model 2) to the reduced
 13 model without the interaction effects using a likelihood ratio test. All models were fitted
 14 using the *lm* function in R with design matrices created by the *multilinearRegression* function
 15 in *noia* R-package (Alvarez-Castro and Carlborg 2007; Le Rouzic and Álvarez-Castro 2008).
 16 P values were calculated using the *lrtest* function in the R-package *lmtest* (Zeileis and
 17 Hothorn 2002).

18

19 Analyses of segregating haplotypes

20

21 **Haplotype estimation in individuals from the founder-lines and AIL F₁₅ generation:** We
 22 inferred the haplotypes for $n = 20$ individuals from each of generations 40 (founders) and 53
 23 from the HWS and LWS lines (in total 80 birds), and all 825 birds with phenotypes from the
 24 AIL F₁₅ generation. This was done using the software *fastPHASE* (Scheet and Stephens 2006)
 25 with default parameters “fastPHASE -Z input-file”. For this, we used 60k SNP chip
 26 genotypes generated for the generation 40 and 53 HWS and LWS individuals in two earlier
 27 studies (Johansson et al 2010) and the new 434 SNP markers genotypes in the AIL F₁₅

1 individuals. Some of the SNP markers genotyped in the AIL F_{15} were not included on the 60k
2 SNP chip and could therefore not be unambiguously assigned to HWS or LWS haplotypes
3 based on this haplotyping. Instead, they were assigned to a founder haplotype based on the
4 agreement between their allele-frequencies, estimated from the pooled whole-genome re-
5 sequencing data described above, and those of a nearby SNP from the 60k SNP-chip
6 genotyping data.

7
8 **Haplotype-based association analysis:** A haplotype-based association analysis was
9 performed in regions of the QTL where the bootstrap-based backward-elimination analysis,
10 and the haplotype estimation, suggested that multiple alleles were segregating in one, or both,
11 of the founder lines. These analyses used a one-way ANOVA (model 1) where Y is a column
12 vector including the residuals from a regression model fitted on the phenotypes of all
13 individuals with the genotypes of all markers selected in the backward-elimination analysis
14 located outside of the evaluated locus and sex as fixed effects. X is the design matrix with one
15 column for the reference-point and multiple columns, and each contain a haplo-genotype. β is
16 a column vector with the estimates for the reference point and the deviations from the
17 reference point for each haplo-genotype, and e is the normally distributed residual.

18
19 **Estimation of genetic effects of individual haplotype effects:** For loci where the haplotype-
20 based association analysis indicated that multiple haplotypes were segregating in one, or both,
21 of the founder lines, also performed was an analysis to disentangle the effects of the
22 individual haplotypes. For a locus with multiple haplotypes (denoted here as $A, B, C, \dots, N$),
23 we first identified the common haplotypes ($p > 0.1$) and only included them in the analysis.
24 Second, a strata was formed that only contained homozygous and heterozygous individuals
25 for particular haplotype combinations (e.g. AA, AB, BB). Third, using these strata we
26 estimated the effects of the individual haplotypes, in turn (e.g. the effects of A and B were
27 estimated in the strata containing individuals with the haplotype combinations AA, AB and
28 BB) using a regression model (model 1) parameterized as follows. Y here contains the residual

for the birds in an analyzed strata from a regression model where the genotype of all markers located outside of the tested locus, and sex of the birds, were fitted as fixed effects in a regression on the 56-day body-weight. X is here the design matrix where the haplo-genotype for individuals homozygous for the LWS derived haplotype, heterozygous for LWS/HWS derived haplotype and homozygous for the HWS derived haplotype are coded as [0,1,2], respectively. e is the normally distributed residual. This analysis was repeated for the all the major haplotypes for which sufficiently large numbers of individuals were genotyped and phenotyped ($n > 50$) in at least two of the three haplo-genotype classes in the AIL F₁₅ generation. The significances for the associations obtained as the p-value from a Wald test calculated using the *lm* function in R.

DATA ACCESS

All raw data will be released upon publication of the final manuscript.

ACKNOWLEDGEMENTS

We thank Leif Andersson for initiating the intercross experiment with PBS and sharing the QTL mapping data from the F₂ intercross. Stefan Marklund, Andreas Lundberg and Mats Pettersson are acknowledged for their valuable contributions during selection of polymorphisms in the selective-sweep regions for genotyping, development of primers for these, and for interacting with the genotyping centre. We also thank Mats Pettersson for useful input on the analysis of the data. Genotyping was performed by the SNP&SEQ Technology Platform in Uppsala, which is part of Science for Life Laboratory at Uppsala University and is supported as a national infrastructure by the Swedish Research Council (VR-RFI). This study was funded by the Swedish Research Council (VR grant ID 2012-4634) and the Swedish Research Council for Environment, Agricultural Sciences and Spatial Planning (Formas grant IDs 2010-643 and 2013-450).

AUTHOR CONTRIBUTIONS

1 ÖC and PBS initiated the study and designed the project; PBS developed the Virginia chicken
2 lines; PBS designed, planned and bred the Virginia Advanced Intercross Line; PBS and CFH
3 designed, planned, bred, bled, phenotyped, and extracted DNA from the F₁₅ Virginia
4 Advanced Intercross Line population; ZS performed the quality control of the genotype data;
5 ÖC and LR designed the statistical analyses; ZS, YZ and ÖC contributed analysis scripts; ÖC,
6 YZ and ZS performed the data analyses and summarized the results. ÖC and YZ wrote the
7 manuscript. All authors read and approved the final manuscript.

8

9

DISCLOSURE DECLARATION

10 The authors declare no competing interests.

11

Legends for supplemental figures and tables:

12 **Figure S1.** Two-locus genotype-phenotype plot. The Y-axis shows the residual
13 phenotypes from a regression model including the 24 associated markers in the multi-
14 locus association analysis. The two markers representing the loci, rs14924102
15 (*Growth1*) and rs14278283 (*Growth12*), were selected as they were the ones with the
16 strongest marginal associations to 56-day body weight in the respective QTL.
17

18

19 **Figure S2.** Visualization of the haplotypes inferred across the polymorphic 60k SNP-chip
20 markers in the region from 13.4-18.0 Mb on GGA7 in the HWS and LWS lineages at
21 generations 40 and 53. In every panel, the rows represent an individual chromosome
22 and the columns the marker genotype at a SNP, where different genotypes are plotted in
23 in different colours.

24

25 **Figure S3.** Visualization of the haplotypes inferred across the polymorphic 60k SNP-chip
26 markers in the region from 22.6-25.3 Mb on GGA7 in the HWS and LWS lineages at
27 generations 40 and 53. In every panel, the rows represent an individual chromosome

1 and the columns the marker genotype at a SNP, where different genotypes are plotted in
2 in different colours.

3

4 **Figure S4.** Visualization of the haplotypes inferred across the polymorphic 60k SNP-chip
5 markers in the region from 32.9-34.6 Mb on GGA3 in the HWS and LWS lineages at
6 generations 40 and 53. In every panel, the rows represent an individual chromosome
7 and the columns the marker genotype at a SNP, where different genotypes are plotted in
8 in different colours.

9

10 **Table S1.** The 11 SNP markers that are associated with 56-day body-weight, and located
11 outside the nine fine-mapped QTL regions, in generation F₁₅ of the Advanced Intercross
12 Line between founders from generation 40 of the High- (HWS) and Low (LWS) body-
13 weight selected Virginia chicken lines.

14

15 **Table S2.** Inferred haplotypes and their allele-frequencies in the two 56-day body
16 weight associated regions in the QTL *Growth9* on GGA7.

17

18 REFERENCES

- 19 Abramovich F, Benjamini Y, Donoho DL, Johnstone IM. 2006. Adapting to unknown
20 sparsity by controlling the false discovery rate. *Ann Stat* **34**: 584–653.
- 21 Alvarez-Castro MJM, Carlborg O. 2007. A Unified Model for Functional and Statistical
22 Epistasis and Its Application in Quantitative Trait Loci Analysis. *Genetics* **176**: 1151–
23 1167.
- 24 Beavis WD. 1998. QTL Analyses: Power, Precision, and Accuracy. In *Molecular Dissection*
25 *of Complex Traits*, Vol. 3 of, pp. 145–162.
- 26 Besnier F, Wahlberg P, Rönnegård L, Ek W, Andersson L, Siegel PB, Carlborg O. 2011. Fine
27 mapping and replication of QTL in outbred chicken advanced intercross lines. *Genet Sel*
28 *Evol* **43**: 3.
- 29 Brandt M, Ahsan M, Honaker CF, Siegel PB, Carlborg O. 2016. *Imputation-based fine-*

1 *mapping suggests that most QTL in an outbred chicken Advanced Intercross Line are*
2 *due to multiple, linked loci.* G3 (Bethesda). doi:10.1534/g3.116.036012.

3 Burke MK. 2012. How does adaptation sweep through the genome? Insights from long-term
4 selection experiments. *Proc R Soc B Biol Sci* **279**: 5029–5038.

5 Carlborg O, Jacobsson L, Ahgren P, Siegel P, Andersson L. 2006. Epistasis and the release of
6 genetic variation during long-term selection. *Nat Genet* **38**: 418–420.

7 Chan YF, Jones FC, McConnell E, Bryk J, Bünger L, Tautz D. 2012. Parallel selection
8 mapping using artificially selected mice reveals body weight control loci. *Curr Biol* **22**:
9 794–800.

10 Cheng R, Lim JE, Samocha KE, Sokoloff G, Abney M, Skol AD, Palmer AA. 2010.
11 Genome-wide association studies and the problem of relatedness among advanced
12 intercross lines and other highly recombinant populations. *Genetics* **185**: 1033–1044.

13 Churchill GA. 2016. Eric Lander and David Botstein on mapping quantitative traits. *Genetics*
14 **203**: 1–3.

15 Dunnington E a, Honaker CF, McGilliard ML, Siegel PB. 2013. Phenotypic responses of
16 chickens to long-term, bidirectional selection for juvenile body weight--historical
17 perspective. *Poult Sci* **92**: 1724–1734.

18 Dunnington EA, Siegel PB. 1996. Long-term divergent selection for eight-week body weight
19 in white Plymouth rock chickens. *Poult Sci* **75**: 1168–1179.

20 Forsberg SKG, Andreatta ME, Huang XY, Danku J, Salt DE, Carlborg Ö. 2015. The Multi-
21 allelic Genetic Architecture of a Variance-Heterogeneity Locus for Molybdenum
22 Concentration in Leaves Acts as a Source of Unexplained Additive Genetic Variance ed.
23 G.P. Copenhaver. *PLoS Genet* **11**: e1005648. doi: 10.1371/journal.pgen.1005648.

24 Gavrilov Y, Benjamini Y, Sarkar SK. 2009. An adaptive step-down procedure with proven
25 FDR control under independence. *Ann Stat* **37**: 619–629.

26 Haley CS, Knott SA. 1992. A simple regression method for mapping quantitative trait loci in
27 line crosses using flanking markers. *Heredity (Edinb)* **69**: 315–324.

28 Hill WG. 2005. A Century of Corn Selection. *Science* **307**: 683–684.

1 Holland JB. 2007. Genetic architecture of complex traits in plants. *Curr Opin Plant Biol* **10**:
2 156–161.

3 Iraqi F, Clapcott SJ, Kumari P, Haley CS, Kemp SJ, Teale AJ. 2000. Fine mapping of
4 trypanosomiasis resistance loci in murine advanced intercross lines. *Mamm Genome* **11**:
5 645–648.

6 Jacobsson L, Park H-BB, Wahlberg P, Fredriksson R, Perez-Enciso M, Siegel PB, Andersson
7 L. 2005. Many QTLs with minor additive effects are associated with a large difference
8 in growth between two selection lines in chickens. *Genet Res* **86**: 115–25.

9 Johansson AM, Pettersson ME, Siegel PB, Carlborg Ö. 2010. Genome-Wide Effects of Long-
10 Term Divergent Selection ed. B. Walsh. *PLoS Genet* **6**: e1001188. doi:
11 10.1371/journal.pgen.1001188.

12 Lango Allen H, Estrada K, Lettre G, Berndt SI, Weedon MN, Rivadeneira F, Willer CJ,
13 Jackson AU, Vedantam S, Raychaudhuri S, et al. 2010. Hundreds of variants clustered
14 in genomic loci and biological pathways affect human height. *Nature* **467**: 832–838.

15 Laurie CC, Chasalow SD, LeDeaux JR, McCarroll R, Bush D, Hauge B, Lai C, Clark D,
16 Rocheford TR, Dudley JW. 2004. The genetic architecture of response to long-term
17 artificial selection for oil concentration in the maize kernel. *Genetics* **168**: 2141–2155.

18 Le Rouzic A, Álvarez-Castro JM. 2008. Estimation of genetic effects and genotype-
19 phenotype maps. *Evol Bioinforma* **2008**: 225–235.

20 Le Rouzic A, Álvarez-Castro JM, Carlborg Ö. 2008. Dissection of the genetic architecture of
21 body weight in chicken reveals the impact of epistasis on domestication traits. *Genetics*
22 **179**: 1591–1599.

23 Márquez GC, Siegel PB, Lewis RM. 2010. Genetic diversity and population structure in lines
24 of chickens divergently selected for high and low 8-week body weight. *Poult Sci* **89**:
25 2580–2588.

26 McClellan J, King MC. 2010. Genetic heterogeneity in human disease. *Cell* **141**: 210–217.

27 Mott R, Flint J. 2013. Dissecting Quantitative Traits in Mice. *Annu Rev Genomics Hum Genet*
28 **14**: 421–439.

1 Peirce JL, Broman KW, Lu L, Chesler EJ, Zhou G, Airey DC, Birmingham AE, Williams
2 RW. 2008. Genome Reshuffling for Advanced Intercross Permutation (GRAIP):
3 Simulation and Permutation for Advanced Intercross Population Analysis ed. W.E.
4 Crusio. *PLoS One* **3**: e1977.doi:10.1371/journal.pone.0001977

5 Pettersson M, Besnier F, Siegel PB, Carlborg Ö. 2011. Replication and explorations of High-
6 Order epistasis using a large advanced intercross line pedigree ed. J.M. Akey. *PLoS*
7 *Genet* **7**: e1002180. doi: 10.1371/journal.pgen.1002180

8 Pettersson ME, Johansson AM, Siegel PB, Carlborg O. 2013. Dynamics of adaptive alleles in
9 divergently selected body weight lines of chickens. *G3 (Bethesda)* **3**: 2305–12.

10 R Core Team. 2015. R Core Team, 2015 R: A Language and Environment for Statistical
11 Computing. R Found. Stat. Comput. Vienna Austria. URL: <https://www.R-project.org/>.

12 Scheet P, Stephens M. 2006. A Fast and Flexible Statistical Model for Large-Scale Population
13 Genotype Data: Applications to Inferring Missing Genotypes and Haplotypic Phase. *Am*
14 *J Hum Genet* **78**: 629–644.

15 Sheng Z, Pettersson ME, Honaker CF, Siegel PB, Carlborg Ö. 2015. Standing genetic
16 variation as a major contributor to adaptation in the Virginia chicken lines selection
17 experiment. *Genome Biol* **16**: 219.

18 Valdar W, Holmes CC, Mott R, Flint J. 2009. Mapping in structured populations by resample
19 model averaging. *Genetics* **182**: 1263–1277.

20 Wahlberg P, Carlborg O, Foglio M, Tordoir X, Syvänen A-C, Lathrop M, Gut IG, Siegel PB,
21 Andersson L. 2009. Genetic analysis of an F(2) intercross between two chicken lines
22 divergently selected for body-weight. *BMC Genomics* **10**: 248.

23 Wood AR, Esko T, Yang J, Vedantam S, Pers TH, Gustafsson S, Chu AY, Estrada K, Luan J,
24 Kutalik Z, et al. 2014. Defining the role of common variation in the genomic and
25 biological architecture of adult human height. *Nat Genet* **46**: 1173–86.

26 Yano K, Yamamoto E, Aya K, Takeuchi H, Lo P, Hu L, Yamasaki M, Yoshida S, Kitano H,
27 Hirano K, et al. 2016. Genome-wide association study using whole-genome sequencing
28 rapidly identifies new genes influencing agronomic traits in rice. *Nat Genet* **48**: 927–

1 934.

2 Zeileis A, Hothorn T. 2002. Diagnostic checking in regression relationships. *R News* **2**: 7–10.

3

Supplemental Figure 1

Supplemental Figure 2

Supplemental Figure 3

Supplemental Figure 4

Supplemental Table S1. The 11 SNP markers located outside the nine fine-mapped QTL regions that were associated with 56-day body-weight in generation F_{15} of the Advanced Intercross Line between founders from generation 40 of the High- (HWS) and Low (LWS) body-weight selected Virginia chicken lines.

GGA ^a	Pos ^b	Marker	Freq HWS/LWS40 ^c	a ± SE ^d	P value ^e	RMIP 5/20% ^f	Sign ^g
1	33147325	rs13847968	0.68/0*	-6.0±6.2	3.0 x 10 ⁻¹	0.33/0.46	0.20
1	87189315	rs13899455	1/0.08	-15.8±5.7	3.3e-03	0.53/0.64	0.05
1	132884851	rs13942473	0.95/0	14.3±7.0	2.6e-02	0.57/0.70	0.05
2	148418221	rs15158686	1/0	24.0±5.8	4.3e-06	0.66/0.76	0.05
4	69059781	rs314510951	1/0	9.6±5.4	5.1e-02	0.37/0.53	0.20
4	82310849	rs14498744	1/0	-21.2±5.9	7.7e-05	0.74/0.85	0.05
6	3292946	rs13562096	1/0	12.7±5.9	3.4e-05	0.60/0.70	0.05
6	3417847	rs15758084	1/0	1.9±5.9	5.0e-04	0.60/0.70	0.05
10	8504427	rs312978808	1/0	10.6±5.6	2.9e-02	0.40/0.58	0.20
13	5988992	rs13726878	0.95/0	14.7±5.9	7.4e-03	0.46/0.62	0.20
23	4599103	rs15205573	1/0.09	15.1±6.2	7.2e-03	0.58/0.76	0.05

^aGGA: Gallus Gallus Autosome, ^bPos: November 2011 (galGal4) assembly; ^cFreq HWS/LWS: Estimated allele frequency in the HWS/LWS founders using individual SNP chip genotypes or, if this was not available, pooled sequencing data (labeled with *); ^dAdditive genetic effect ± Standard Error; ^eSignificance for additive genetic effect in model including all loci significant at 20 % FDR; ^fResample Model Inclusion Probability at 5/20 % FDR threshold; ^gFDR threshold at which the marker was selected with RMIP > 0.46[11].

Supplemental Table S2. *Haplotypes and their corresponding allele-frequencies for two regions in Growth9*

QTL region	Haplotypes	Haplotype frequencies		
		F15	HWS40	LWS40
Growth 9.1	AA (HWS1)	0,35	0,65	0
	GA (HWS2)	0,09	0	0
	AG (HWS3)	0	0,175	0
	GG (LWS1)	0,55	0,175	1
Growth 9.2	AA (HWS1)	0,4	0,875	0
	CA (HWS2)	0,1	0,125	0
	CT (LWS1)	0,48	0	1