

Low-coverage sequencing in a deep intercross of the Virginia body weight lines provides insight to the polygenic genetic architecture of growth: novel loci revealed by increased power and improved genome-coverage

T. Rönneburg ^{*}, Y. Zan,^{*} C. F. Honaker,[†] P. B. Siegel,[†] and Ö. Carlborg ^{*,1}

^{*}Department of Medical Biochemistry and Microbiology, Uppsala University, Uppsala, Sweden; and [†]Department of Animal and Poultry Sciences, Virginia Polytechnic Institute and State University, Blacksburg VA, USA

ABSTRACT Genetic dissection of highly polygenic traits is a challenge, in part due to the power necessary to confidently identify loci with minor effects. Experimental crosses are valuable resources for mapping such traits. Traditionally, genome-wide analyses of experimental crosses have targeted major loci using data from a single generation (often the F_2) with individuals from later generations being generated for replication and fine-mapping. Here, we aim to confidently identify minor-effect loci contributing to the highly polygenic basis of the long-term, bi-directional selection responses for 56-d body weight in the Virginia body weight chicken lines. To achieve this, a strategy was developed to make use of data from all generations (F_2 – F_{18}) of the advanced intercross line, developed by crossing the low and high selected lines after 40 generations of selection. A cost-efficient low-coverage sequencing based approach was used to obtain high-confidence genotypes in 1Mb bins across 99.3% of the chicken genome for >3,300 intercross individuals. In total, 12 genome-wide significant, and 30 additional suggestive QTL reaching a 10%

FDR threshold, were mapped for 56-d body weight. Only 2 of these QTL reached genome-wide significance in earlier analyses of the F_2 generation. The minor-effect QTL mapped here were generally due to an overall increase in power by integrating data across generations, with contributions from increased genome-coverage and improved marker information content. The 12 significant QTL explain >37% of the difference between the parental lines, three times more than 2 previously reported significant QTL. The 42 significant and suggestive QTL together explain >80%. Making integrated use of all available samples from multiple generations in experimental crosses are economically feasible using the low-cost, sequencing-based genotyping strategies outlined here. Our empirical results illustrate the value of this strategy for mapping novel minor-effect loci contributing to complex traits to provide a more confident, comprehensive view of the individual loci that form the genetic basis of the highly polygenic, long-term selection responses for 56-d body weight in the Virginia body weight chicken lines.

Key words: QTL mapping, low-coverage sequencing, body weight, advanced intercross line

2023 Poultry Science 102:102203

<https://doi.org/10.1016/j.psj.2022.102203>

INTRODUCTION

Quantitative traits remain difficult to analyze and break down into their component loci (Flint and Mott, 2001). Effect sizes of individual loci often explain a very small fraction of the phenotypic variance (Boyle et al., 2017 and references within)—often much smaller than environmental effects—and are regularly dependent on the genetic background (Pettersson et al., 2011; Mackay, 2014; Forsberg et al., 2017; Zan and

Carlborg, 2020). Experimental populations are valuable resources for studying quantitative traits, and by reducing confounding factors such as environmental noise, they have provided a clearer view on the genetic architecture of a wide range of complex traits (Andersson, 2001; Flint and Mott, 2001; Andersson and Georges, 2004). Examples include shank-length in mice (Castro et al., 2019), longevity in *Drosophila melanogaster* (Curtsinger and Khazaeli, 2002), and oil content in corn (Hopkins, 1899; Dudley, 2007).

Although quantitative trait loci (QTL) studies in experimental crosses have high power, resolution is limited due to the extensive linkage disequilibrium (LD) introduced by the crossing design. This meant that, historically, it was possible to use the sparse or very sparse marker maps available at the time without compromising

© 2022 The Authors. Published by Elsevier Inc. on behalf of Poultry Science Association Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

Received December 14, 2021.

Accepted August 24, 2022.

¹Corresponding author: orjan.carlborg@imbim.uu.se

the resolution much further, as resolution was primarily gated by the number of recombinations. Still, this resulted in regions of the genome with less coverage and, in some cases, missing data on small chromosomes and/or chromosomal ends (Mackay, 2001). Similarly, it was not uncommon to have regions lacking markers informative for line origin when using experimental crosses between outbred founders (Andersson, 2001). To increase resolution and facilitate fine-mapping of detected QTL, follow-up studies in additional crosses have been performed. Generally, these have excluded regions outside of previously observed QTL, therefore leaving much of the genome without further study. As a result, these studies may be underpowered or lack the resolution to make inferences on the genetic architecture of the studied traits beyond a few large-effect loci (Flint and Mott, 2001).

More complete dissection of highly polygenic complex traits requires large and powerful studies. In natural populations such as humans, hundreds of thousands of individuals have been used to study highly polygenic model traits such as height (Lango Allen et al., 2010; Yang et al., 2010; Wood et al., 2014). In experimental populations, smaller populations are required to detect even minor-effect loci due to the higher power achieved from, for example, segregation of alleles at intermediate frequencies and greater control over environmental influences. Genotyping and imputation approaches based on low-coverage whole genome sequencing (WGS; Altshuler et al., 2000; Andolfatto et al., 2011; Zhang et al., 2015; Pértille et al., 2016; Whalen et al., 2018; Zan et al., 2019), have provided opportunities to reanalyze existing individuals generated for different purposes to perform integrated analyses, thus enabling greater insights into the contribution of loci with minor effects on the genetic basis of complex traits in existing experimental populations. The increased genome-wide marker coverage provided by WGS-based genotyping technologies also provides an extended coverage of regions outside of the current consensus linkage maps in species such as the chicken, where the major focus has been on the large chromosomes, leaving the microchromosomes largely unexplored (Groenen et al., 2000, 2009; Elferink et al., 2010).

The Virginia body weight lines of chickens were developed by long-term, bi-directional selection for a single trait—body weight at 56 d of age—resulting in a nine-fold difference between the low (LWS) and high (HWS) weight lines after 40 generations of selection (Dunnington and Siegel, 1996; Márquez et al., 2010; Dunnington et al., 2013). Genome-wide comparisons showed that the footprint of selection between LWS and HWS covers hundreds of loci across the genome (Johansson et al., 2010; Lillie et al., 2018, 2019). Efforts to identify which of these loci contribute to the observed responses include a series of experiments utilizing an intercross developed from individuals of generation S₄₁ ($n_{\text{HWS}} = 29$, $n_{\text{LWS}} = 30$). Efforts include genome-wide mapping (Jacobsson et al., 2005; Carlborg et al., 2006; Wahlberg et al., 2009), as well as replication and fine-mapping studies (Besnier et al., 2011; Pettersson et al.,

2011, 2013; Sheng et al., 2015; Brandt et al., 2017; Zan et al., 2017) on different generations in this population. Although these studies agree that the long-term responses are primarily from selection on a highly polygenic genetic architecture where most loci have small effects, the statistical support for individual loci is low. This study aims to overcome this deficiency of statistical power to facilitate the mapping of contributing minor-effect loci with confidence by re-genotyping and performing an integrated analysis of >3,300 individuals from generations F_2 to F_{18} of the intercross of the Virginia body weight lines. Aims were to identify and map new QTL, confirm earlier reported loci, and explain more of the selection responses with individually significant loci than previously reported.

MATERIALS AND METHODS

Materials

All chickens described in this study were White Plymouth Rock originating from the Virginia body weight lines and their intercross. The procedures involving animals used in this experiment were carried out in accordance with the Virginia Tech Animal Care and Use Committee protocols.

The custom code, software versions, and software parameters used in this manuscript can be found in a dedicated Github repository at github.com/CarlborgGenomics/AIL-scan

The sequencing data generated for this project can be found in the NCBI sequence read archive, BioProject PRJNA788343.

Sequencing

Founder-line Sequencing and Variant Calling All 59 HWS and LWS founders of the AIL ($n_{\text{HWS}} = 29$ and $n_{\text{LWS}} = 30$) were whole-genome re-sequenced to ~30X coverage (Guo et al., 2019). The obtained reads were mapped to the newest reference genome (GalGal6a, Genome Reference Consortium, 2018) using *BWA* (version 0.7.17; Li, 2013). Variants were called and filtered using *GATK* (McKenna et al., 2010) according to best-practices recommendations (DePristo et al., 2011; Auwera et al., 2013) modified to accommodate for non-model organisms.

Low-coverage Sequencing of Advanced Intercross Line Individuals In addition to the 837 F_2 individuals sequenced by Zan et al. (2019), all remaining chickens from generations F_2 - F_{18} of the AIL (total, $n_{F_2-F_{18}} = 3,327$, Table S2) were sequenced to ~0.4X coverage using the strategy outlined by Zan et al. (2019). The obtained reads were mapped to the GalGal6 reference genome using *BWA* (version 0.7.17; Li, 2013). Variants were next called using a pipeline implemented using *bcftools* (version 1.9; Li, 2011), *samtools* (version 1.9; Li et al., 2009), *biopython* (version 1.70; Cock et al., 2009), and *cyvcf2* (version 0.10.0; Pedersen and Quinlan, 2017). Obtained variants were merged for each

generation and filtered to only include polymorphisms present in the filtered set of founder genotypes described in the section above. Next, for each individual, only variants informative about the founder line origin (HWS/LWS) were kept and formatted for use as input to the *Stripes* genotyping pipeline (Zan et al., 2019).

Genotype Estimation in the Virginia Body Weight Line AIL Using Deep-Stripes, a Pipeline for Founder-Line Genotype Estimation in Deep Intercross Populations

Deep-Stripes *Deep-Stripes* was used to call founder line origin genotypes in 1 Mb bins across the genome of the F_2 to F_{18} generations of the advanced intercross line (AIL). For this, high-coverage sequence data from the outbred founders of the population and low-coverage sequence data from the intercross individuals, generated as described below, were used.

Deep-Stripes is an extension of *Stripes* (Zan et al., 2019), a pipeline for founder-line genotype estimation using low-coverage sequencing data, adapted from *TIGER* (Rowan et al., 2015) for use in outbred intercross populations. In deep intercross populations from outbred founders, such as the AIL studied here, there is a generally lower and more variable density of founder-line informative markers. Here, we have further extended the *Stripes* pipeline to deep intercross populations, including updates to enhance stability and improve genotype calling quality in later generations.

Deep-Stripes updates are implementations of (a) reverting to hardcoded genotype emission thresholds in cases where the original nonlinear minimization procedure for determining these (Rowan et al., 2015) failed due to uniform ancestry across an entire chromosome, (b) a modified nonlinear minimization procedure improving convergence as well as defaulting to hardcoded parameters after 20 unsuccessful tries to determine the genotype emission thresholds, (c) a modified logic for comparing highly similar beta distributions to make results stable across computing platforms, and (d) automation of multiple rounds of genotype estimation for each individual.

Genotype Quality Control and Filtering In order to detect incoherently called genotypes in low-information areas due to a delay of inferred crossovers to the end of such regions, *Deep-Stripes* implemented genotype estimation in both directions on the chromosome. Genotype estimation in this way with 2 window-sizes (50 and 200 markers) reduces the number of false positive crossovers in marker-dense areas resulting from the increased flat, per-window genotype estimation error rate in high marker density regions. The final genotypes were assigned by transforming the output from each of the four runs (2 directions $\times$ 2 window-sizes) described above to a genotype matrix. These contained estimated founder line genotypes for each individual in even-sized (1 Mb) bins across the genome. The process for this used the 4 matrices as follows: In bins where no

recombination event was inferred, the genotypes were coded as numerical values (1, 0, -1) corresponding to the homozygote for founder-line 1, heterozygote and homozygote for founder-line 2, respectively. Recombination breakpoints were estimated with bp resolution. However, if one or more recombination events were detected, resulting in multiple genotypes being called in a 1 Mb window, the genotype was scored on a continuous scale from 1 to -1 by averaging the founder genotypes scores across the base pairs in the segment. Second, the genotype matrices from the forward and reverse runs were filtered by considering bins where estimates differed by more than one recombination event as uninformative and setting them to missing. This procedure was performed separately for the two window-sizes (50 and 200 markers). Third, the forward- and reverse-filtered genotype matrices obtained using 50 and 200 marker window-sizes were combined. This was done by using the genotype derived using 50 marker sliding windows if the marker density was below 20 markers/Mb and 200 if above. Bins with ambiguous genotypes were set to missing. These were defined as bins with genotype scores in the ranges [0.8, 0.2] and [-0.8, -0.2]. Finally, all bins genotyped in less than 100 individuals were set to missing.

Analysis

QTL Mapping QTL mapping was performed for body weight at 8 wk of age. To correct for generational effects, 8-wk body weights were standardized within each generation. This was done by subtracting the generation mean from each observation and then dividing it by the within-generation standard deviation. The genome scan was performed using the ‘scanone’ function from the package ‘qtl’ (Broman et al., 2003) in R (R Core Team, 2013) using Haley-Knott regression (Knott and Haley, 1992) with sex as a covariate. Significance threshold was obtained using permutations ($n = 10,000$), resulting in a 5% genome-wide significance of a Logarithm of Odds (LOD) score of 4.01.

Identification of Suggestive QTL using a FDR Approach Suggestive QTL were here defined as those that reached a 10% false-discovery rate (FDR) threshold. First, LOD scores for markers across the genome were transformed into P -values (Peirce et al., 2006). They were then evaluated against a significance threshold adjusted for multiple testing using the Benjamini Hochberg procedure with a false discovery rate of 10%, as implemented in *statsmodels* (Benjamini and Hochberg, 1995; Seabold and Perktold, 2010).

Estimation of the Genetic Effect and Residual Variance Explained by the Mapped QTL To estimate the residual variance explained by the QTL, we corrected for the sex effect using a linear model and fit either 1) all 12 significant or 2) all 42 suggestive and significant QTL using the *fitqtl* function in *r/qtl* on the residuals. This was done on the same individuals used in the QTL scan. Estimates for the residual variance explained by each

QTL were obtained by fitting all significant and suggestive QTL jointly and using the *SSv3* drop-one-term anova in the *fitqtl* function. Estimates for the effect on body weight in grams for each QTL were determined by fitting each QTL individually with sex as a covariate (Table S3). Due to the standardization of the phenotypes, the estimates were multiplied with the population-standard deviation to obtain an estimate in grams. The sum of these estimates was then expressed as a fraction of the between-line difference. The same was also done by fitting all QTL jointly (Table S4).

$$y = SEX + \sum_{i=n}^1 QTL_i \quad (1)$$

Estimating the Effect of Increased Genotyping Information Content on Statistical Power in QTL Analyses The information content (IC) was calculated across the genome in the set of F_2 individuals that were common to this study and the earlier one of the F_2 generation (Wahlberg et al., 2009). The measure used was defined as (Knott et al., 1998):

$$IC = VAR(a) + 2VAR(d) \quad (2)$$

This was calculated at individual marker locations, as well as every cM, across the genome using the a and d indicator regression variables from Wahlberg et al. (2009). For the dataset from this study, the a and d indicator variables were calculated from the genotype estimates in each 1Mb bin as (Knott and Haley, 1992):

$$a = P(Homozygote_{HWS}) - P(Homozygote_{LWS}) \quad (3)$$

$$d = P(Heterozygote) \quad (4)$$

The information content was compared at the physical locations (Mb; Genome Reference Consortium, 2018) of the genotyped markers in Wahlberg et al. (2009) and across all tested locations in the 2 studies (every cM/Mb; Wahlberg et al., 2009; this study).

RESULTS AND DISCUSSION

Increased Coverage in the Genome-Wide Scan via Genotyping by Low-Coverage Sequencing

After sequencing and variant calling, genotypes were estimated for $n_{F_2-F_{18}} = 3,327$ AIL individuals that passed quality control. The average density of informative markers across the genome was 102 markers/Mb. Because the founder lines are outbred and the individual AIL offspring descend from different founders, the number of informative SNPs varied among individuals and decreased over generations, as more ancestors contributed to the genotype of each individual. Founder-line genotypes were obtained for all of the 1,058 1Mb bins defined on the 33 largest chromosomes, identifying an average of 74 recombination breakpoints per Mb across all individuals before filtering. The accuracy of this imputation approach was previously validated on a subset of the individuals against experimentally determined SNP genotypes by Zan et al. (2019; see Figure 3) Compared to the previous genome-wide scan performed in the F_2 population by Wahlberg et al. (2009), an additional 100 Mb were covered (+27%) with markers, encompassing 2 small chromosomes (Chr31, Chr33), several previously uncovered scaffolds/unplaced segments and chromosome ends (Figure 1). Further, the information content at the tested locations across the genome also improved from an average of 0.77 (Wahlberg et al., 2009) to an average of 0.90 (Figure 3, panel F).

Since the founder lines were outbred, only a small fraction of the total markers genotyped for each individual are informative for line origin, averaging between 1.4 and 8% ($4.7 \pm 3.3\%$, on average >100,000 markers, Figure S1). While these are more than sufficient to impute founder-line haplotypes with confidence, the trade-off for this type of imputation is that it does not account for variation within the outbred lines, which limits detection power for QTL that do not differ between the lines or segregate in only one line. While

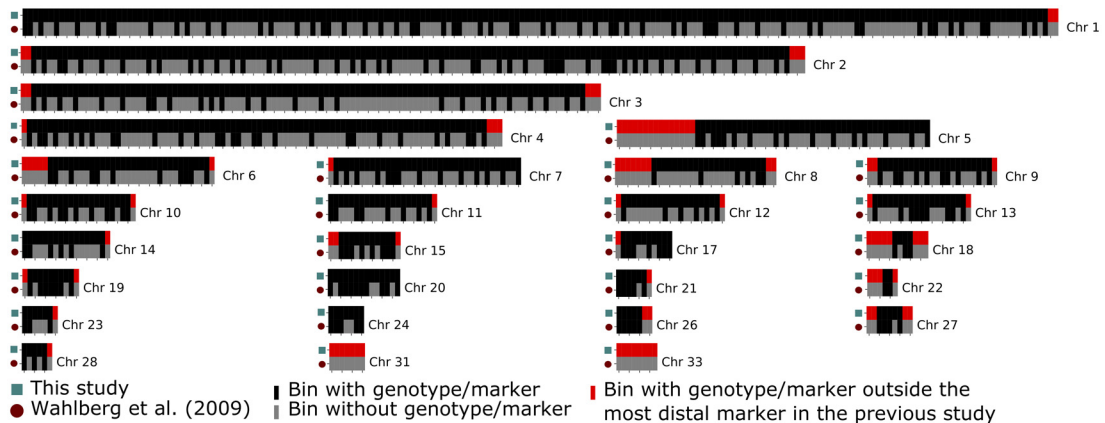


Figure 1. Genome coverage by the use of low-coverage sequencing data compared to an earlier F_2 genome scan (Wahlberg et al., 2009) based on 372 SNP and microsatellite markers. Black/grey colors indicate 1 Mb bins with/without genotypes and red highlights chromosome ends with new genotypes outside the outermost markers in Wahlberg et al. (2009).

this trade-off for confidence and power in detecting QTLs is a limitation, it is due to the imputation, and could be supplemented by alternative approaches on the same sequencing data for subsequent finemapping.

More Individuals and Increased Genome-Coverage Facilitate Detection of New QTL

Comparisons to Earlier Mapped QTL from Genome Scanning of the AIL As illustrated in Figure 2, the 2 +1 genome-wide significant and suggestive QTL in the most recent genome-scan of the F_2 generation of the Virginia line AIL (Table 1; Wahlberg et al., 2009) were also detected here. Further, nine additional QTL for this trait were also mapped with genome-wide significance (Figure 2; Table 1). At a 10% FDR-adjusted threshold for significance, a total of 42 QTL were identified (Figure 2; Table S3), that is, 30 QTL in addition to the 12 genome-wide significant QTL.

Overlap With Suggestive Regions From Fine-Mapping Analyses of the Virginia Line AIL

Several fine-mapping studies (Besnier et al., 2011; Sheng et al., 2015; Zan et al., 2017) have been performed in 1) 10 QTL-regions with significant and suggestive associations to 56-d body weights in the F_2 generation (Wahlberg et al., 2009) and 2) 99 selective sweep regions identified in comparisons of the HWS and LWS Virginia lines (Johansson et al., 2010). Here, 2 loci (located on Chr10, Mb 9, and Chr23, Mb 6) that were not part of the significant or suggestive QTL in the original F_2 genome-scan (Wahlberg et al., 2009) reached genome-wide significance when analyzed across the entire AIL (Figure 2, Table 1). Further, when comparing the 42 QTL significant at a 10% FDR threshold to 99 regions under selection evaluated by Sheng et al. (2015), 41 of these overlapped.

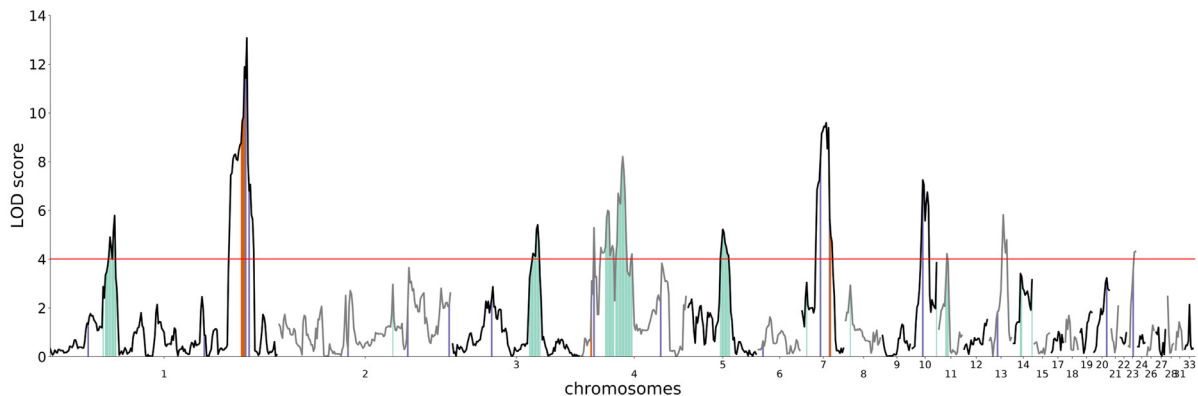


Figure 2. Genome-wide QTL scan for 56-d body weight in generations F to F_{18} of the Virginia body weight lines AIL. The y-axis shows the statistical support for QTL as LOD scores and the x-axis the genomic location in Mb bins. The solid red horizontal line shows the genome-wide significance thresholds. Red vertical segments show the most recent earlier reported associations in genome-wide scans in F_2 (Wahlberg et al., 2009), blue vertical segments indicate associations in the fine-mapping analyses in F_{15} (Zan et al., 2017). Green vertical segments indicate new suggestive QTL without a previous association within 15 Mb.

More Individuals and Increased Marker Density Dissects a QTL on Chromosome 4 into Multiple, Independent Associated Regions

Previously, 2 regions on Chromosome 4 were implicated for body weight, with one of them, *Growth6*, reaching genome-wide significance for 56-d body weight in the latest genome scan of the F_2 generation (Wahlberg et al., 2009). The population used here provided sufficient power to replicate the previously reported QTL and confirm the association of another region, *Growth7*, that was previously reported as associated with body weight at different ages and growth traits (Figure 3, panel B, compares F_2 individuals from Wahlberg et al. (2009; dashed lines) to the AIL population (solid lines)).

In addition, the approach used here provided sufficient resolution to partition *Growth7* into three independent peaks (4/23, 4/36, 4/70 Chr/Mb). Figure 3, panel B, provides a comparison of the full AIL population (solid lines) with *deep-Stripes* inferred genotypes (empty squares, green) to the genotypes from the SNP and micro-satellite based marker panel used in Wahlberg et al. (full circles, pink) as an illustration of how the additional resolution facilitated separation of the 4/23 and 4/36 loci.

Large Contributions by Mapped QTL to the Selection Response

The AIL was produced by intercrossing chickens from HWS and LWS after 40 generations of selection (S_{41}), and the founders for the intercross differed more than 8-fold (1,341 g) in 56-d body weight (Table 1). Estimated from the AIL F_2 - F_{18} , the 12 QTL reaching genome-wide significance explain 501.4 g (37.4%) of the difference between the parental lines (Table 1). This compares to the 159 g (12%) explained by the 2 significant QTL in Wahlberg et al. (2009). Together, the significant QTL mapped here/by Wahlberg et al. (2009) explained 8.3/

Table 1. Significant QTL for 56-day body weight and overlaps with earlier reported significant and suggestive QTL in this population.

Chromosome	Position (Mb)	a(SE) ¹	%var ²	LOD ³	QTL	Previous
1	56	19.6 (4.0)	0.43	5.79		
1	171	31.6 (4.1)	1.21	13.08	Growth1	Zan (2017), Wahlberg (2009)
3	74	20.3 (4.2)	0.31	5.41	Growth4	
4	11	20.7 (4.4)	0.28	5.29	Growth6	Zan (2017), Wahlberg (2009)
4	23	20.4 (4.0)	0.07	6	Growth7	
4	36	25.6 (4.2)	0.45	8.2	Growth7	
5	30	20.9 (4.3)	0.31	5.22	Growth8	
7	21	25.6 (4.0)	1.26	9.6	Growth9	Zan (2017), Wahlberg (2009)
10	9	22.3 (3.9)	0.27	7.25		Zan (2017)
11	7	13.2 (4.0)	0.32	4.22		
13	12	19.6 (3.9)	0.32	5.83	Growth10	
23	6	10.9 (4.3)	0.47	4.32		Zan (2017)
		a = 250.7				
		2a = 501.4				

¹Estimated effect size.²Percent of total variance explained.³Logarithm of Odds.

5.2% of the residual phenotypic variance in the AIL F_8 to F_{18} populations (Table 1). The 42 markers derived from the FDR-approach explain 14.6% of the residual phenotypic variance. This corresponds to 1,130 g (84.3%) of the difference between the founder lines using individually fitted estimates or 793 g (59%) of the difference between the founder lines using estimates from the joint fit.

Concordance With Previous Estimates of Effect Sizes

The difference in effect size estimates for *Growth1* on Chromosome 1 between Wahlberg et al. (2009) and the approach used here were within the range of the standard error (34.2 ± 9.2 g and 31.6 ± 4.1 g, respectively). The effect size estimated here for *Growth6* on Chromosome 4 (20.7 ± 4.4 g) was lower than the previous estimate (36.3 ± 8.3 g). However, the current approach found 3 significant and 2 suggestive QTL on

Chromosome 4. Of these, 2 significant and one suggestive QTL were within a QTL region (*Growth7*) reported in the first analysis of the F_2 generation (Jacobsson et al., 2005), but later showed no significant association with 56-d body weight in the extended analysis of the same individuals with a more comprehensive marker set by Wahlberg et al. (2009). In this study, these 3 + 2 QTL explained a total of 91.7 g suggesting a larger total contribution by Chromosome 4 than the earlier studies. In contrast, *Growth9* on Chromosome 7 had a reduced effect size estimate (25.6 ± 4.0 g), compared to the 43.2 ± 9 g estimated by Wahlberg et al. (2009).

Additional Power in the Analyses Facilitates Detection of New QTL

A Novel QTL on Chromosome 1 Revealed by Combining Data Across Generations A genome-wide significant QTL was mapped to 27 Mb on Chromosome 1 (Figure 3a; Table 1). This region was covered by

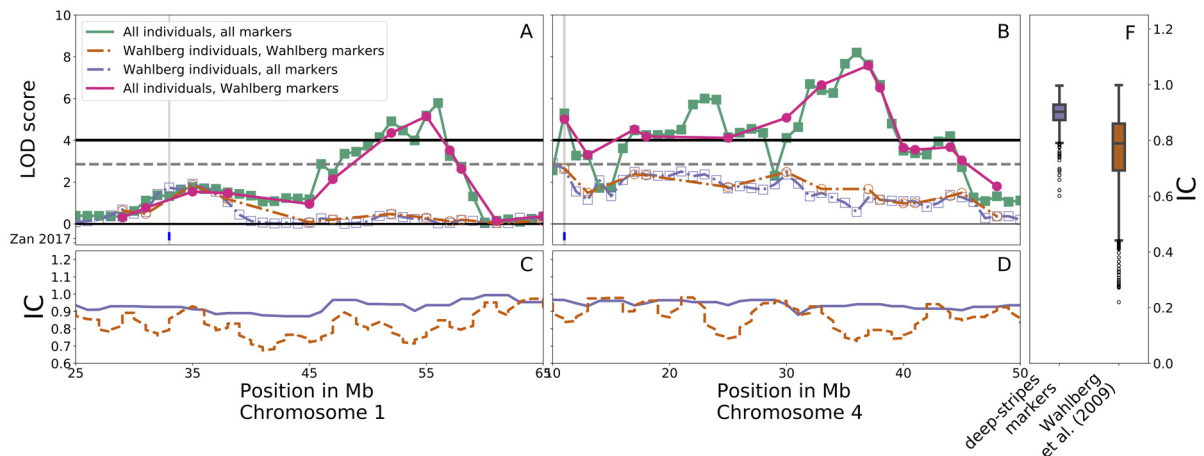


Figure 3. Comparisons to previous marker-sets from Wahlberg et al. (2009). Upper left and middle panels show the LOD score across selected peak regions on chromosome 1 (25–65 Mb) and chromosome 4 (10–50 Mb), using different selections of markers and individuals. Dashed/solid lines show the QTL scans using the F_2 individuals from Wahlberg et al. (2009). All individuals, solid round/empty square markers show QTL scans using markers from Wahlberg et al. (2009) / low coverage data. Bottom left and middle panels display information content across the same regions on Chromosome 1 and 4 for the markers used in Wahlberg et al. (2009) (orange, dashed) and the *deep-Stripes* markers (periwinkle, solid) across the individuals used by Wahlberg et al. (2009). The right panel summarizes information content for the same sets of markers and individuals, but across the 30 largest chromosomes.

multiple SNP and micro-satellite markers in the previous genome scans (Wahlberg et al., 2009; Figure 1). However, as these markers segregated in the founder lines, the estimation of founder-line QTL genotypes was less precise than the current approach utilizing low-coverage sequencing data (Figure 3c). To evaluate the potential contribution by either improved marker density or larger number of individuals to an increase in statistical power, the genome scan was performed on different subsets of the data. The first subset was F_2 individuals from Wahlberg et al. (2009), and second, the subset of bins containing SNP and microsatellite markers from the same study. Third, all combinations with the individuals and markers from this study were used. The results showed (Figure 3) that finding this QTL was driven by the integration of all generations in the AIL rather than the increase in marker density.

New QTL Revealed in Regions With Poor Marker Coverage in the Genome

The low coverage sequencing approach implemented here improved the genome-wide marker coverage in this population. In particular, the coverage was improved by reaching further out to the ends of almost all chromosomes, specifically 5, 6, and 8 (Figure 1). As a result, a suggestive QTL for 56-d body weight was observed on Chromosome 8 where an additional 7 Mb (0–7 Mb) were covered on its distal end (Figure 4b). The QTL peak is located on the end of the chromosome and did not extend into the part of the chromosome covered in the earlier studies (Wahlberg et al., 2009). Its effect is relatively small, and hence explains a modest amount of the variance in 56-d body weight (Table S1). Its peak location was further located approximately 20 cM outside of the most distal marker in the linkage map used by Wahlberg et al. (2009). Figure 4 illustrates how the

combination of better regional coverage and gain in power from merging individuals from multiple generations in the AIL resulted in its detection.

Doubling the Number of Genome-Wide Significant QTL Identified Triples the Amount of the Founder-Line Difference Explained for the Selected Trait

In this study, the intercross population was bred from the Virginia body weight selection lines. These two pedigreed populations were divergently selected long-term for a single trait, 56-d body weight, for 40 generations before the intercross was formed. Previous whole-genome analyses of this intercross were challenged to detect genome-wide significant QTL, finding only 1 (Jacobsson et al., 2005) or 2 (Wahlberg et al., 2009). Subsequent fine-mapping analyses focused on significant and suggestive QTL regions and selective-sweep regions (Johansson et al., 2010; Besnier et al., 2011; Pettersson et al., 2011, 2013; Sheng et al., 2015; Zan et al., 2017) Results from these studies suggested that the genetic architecture of 56-d body weight in this population was highly polygenic and that individual loci contributed small marginal effects.

The increased power in our study obtained by increasing quality and coverage of markers, as well as increasing population-size by integrating across generational data, facilitated detection of 10 additional genome-wide significant QTL associated with 56-d body weight. None had been previously identified as significant. The residual phenotypic variance explained by the mapped QTL increased from 5.3% by the 2 genome-wide significant QTL of Wahlberg et al. (2009) to 8.3% by the 12 genome-wide significant QTL detected here. The effect sizes of the QTL estimated in the population analyzed

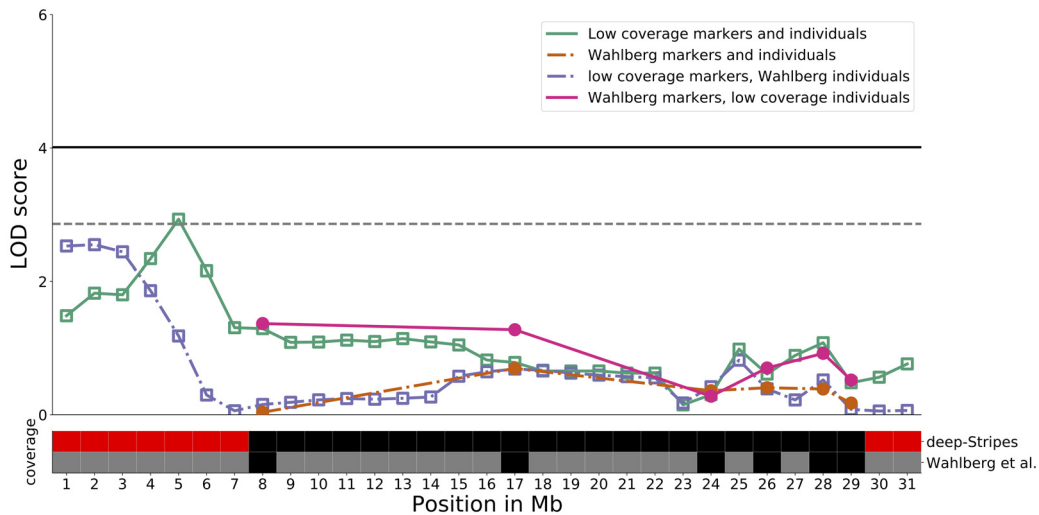


Figure 4. A novel QTL due to increased coverage on Chromosome 8. Upper panel: QTL scan for 56-d body weight across Chromosome 8, using the individuals from the original F_2 genome scan (Wahlberg et al., 2009, dashed lines) and all AIL individuals (solid lines) utilizing only the marker positions from Wahlberg et al. (2009); filled circles) or all low-coverage markers (empty squares). Lower panel: Chromosome 8 sectioned into 1 Mb bins, indicating whether markers are present (black, red) or not present (grey) for both low coverage data (upper row) and the marker set from the original F_2 scan by Wahlberg et al. (2009; lower row). Red highlights chromosome ends with new genotypes outside the outermost markers in Wahlberg et al. (2009).

here were often lower than previous estimates, with the exception of *Growth1*. The reduced effect size for *Growth9* was not surprising, given that Chromosome 7, and *Growth9* in particular, were involved in epistatic interactions (Pettersson et al., 2011) and is likely a complex region. For *Growth6*, while the estimate is lower than that in Wahlberg et al. (2009), the sum of all significant and suggestive QTL on Chromosome 4 was larger (91.7 g compared to 79.1 g) than the estimate for the 2 earlier mapped QTL on Chromosome 4 (*Growth6* and *Growth7*; Jacobsson et al., 2005). Our study thus confirms previous suggestive evidence obtained in studies of related growth phenotypes.

Additionally, our study enabled fine-mapping them into multiple independent regions previously associated with 56-d body weight found by selection-scans (Zan et al., 2017). In total, out of 11 markers significant in Zan et al. (2017), all but 2 were situated in QTLs passing the 10% FDR threshold here. Further, the considerable overlap between the QTL from this study and the previously identified sweep regions from Johansson et al. (2010), supports the hypothesis that the sweeps result from selection acting on QTL rather than being the result of drift.

Overall, the effects of the significant QTL contributed 37.4% of the founder-line difference (Table S2), which is >3 times the founder-line difference explained by the significant QTL in Wahlberg et al. (2009). The QTL significant at a 10% FDR-threshold explained more than 84% of the difference between the founder lines. A caveat with this estimation is the large fraction of the genome covered by these markers, and hence the approach used here to estimate the effect size of individual loci likely contributed to a slight overestimation from residual linkage between QTL and tagging of linked loci. The estimates might also be affected by a “winners curse,” as the effect sizes were estimated using the same individuals as in the QTL mapping. We consider a likely range of the estimate to be between this estimate and that of the joint fit (59%). This due to the joint fit likely being an underestimate, for the same reasons as described above.

Further, 30.3% of the bins on the covered autosomal chromosomes were statistically associated with the phenotype after lenient adjustment for multiple testing. This means that any overlap between the extended QTL and the selective sweep regions should be interpreted with caution. Still, this supports the interpretation of a high polygenicity of the investigated trait, and it is reasonable to believe that these regions were undergoing true and detectable selection, given the strong single trait artificial selection regime in the selected lines (Siegel, 1962; Dunnington and Siegel, 1996; Márquez et al., 2010; Dunnington et al., 2013).

Added Value by Increased Genome-Wide Marker Coverage

Genotyping by sequencing has provided opportunities to both increase marker density and decrease the cost

for genotyping. Strategies based on low-coverage sequencing approaches open cost-efficient opportunities to re-analyze existing experimental populations. Here, we demonstrated one such approach targeted to deep outbred intercrosses. This approach increased the overall marker coverage of the chicken genome from 93% (Wahlberg et al., 2009), as estimated as coverage of the autosomal genetic map, to 1,058 Mb (99.3% of 1,065 Mb; Genome Reference Consortium, 2018) with a mean density of 102 line-informative markers per Mb.

Beyond Markers that are fixed and therefore always informative for line-origin, outbred founder lines have many more markers that can be line-informative since they are fixed between the HWS and LWS founders contributing to the ancestry of an individual, despite still segregating within the lines. However, as the number of these markers decreases with the increase in unique ancestors contributing to each individual in later generations, the depth of generations one can investigate with this can be limited by the number of founding individuals as well as the size of each generation. Thus, it is important that the founder lines are sufficiently divergent to provide a minimum resolution via fixed markers. However, because the nonphysical window size of the imputation process does not sacrifice higher resolution in divergent regions, it is likely that any cross between lines with a significant genetic component to their phenotypic divergence will have sufficient coverage with informative markers in regions of interest. For the AIL used in this study, it is likely that the resolution is not limited by the marker coverage, but rather by both the 1 Mb binning approach and the 50/200 marker sliding window imputation of founder genotypes. This is because there were enough accumulated recombinations in the later generations that were lost through these approaches. While these were chosen for robustness and provided an appropriate resolution/power trade-off across all individuals, modifying these parameters could provide added opportunities for fine mapping.

The increased marker density also increased the information content throughout the genome, which further increased power in the QTL analyses. In particular, coverage was extended at the ends of the larger chromosomes. This led to locating a novel suggestive QTL on Chromosome 8 that was primarily due to increase in coverage, though the increase in power helped elevate and define the peak of the QTL. It was seen in the full marker set using only the F_2 population, but without a definitive peak. This is likely due to a lack of recombination events distal to the peak in the F_2 population, though the LOD scores for the outermost 3 Mb were above the 5% permutation threshold for chromosome-wide significance.

In addition, two additional small linkage groups were also covered. The minor regions of the genome that remain are present in scaffolds containing too few markers for reliable genotype estimation using the *Stripes* pipeline. Additional work beyond this study is needed to estimate the genotype in these regions using alternative genotyping or bioinformatics approaches before they can be included in QTL mapping studies.

Integrating AIL Data

Integrating data across an AIL population provides value to the QTL scan by helping to map new QTL and refining the resolution and explanatory power of existing ones. This is particularly so when intermediate generations already exist due to previous attempts in fine mapping. At less than 1 EUR/sample for library preparation and 10 EUR/sample for library preparation and sequencing (Zan et al., 2019), our approach is cost-effective in enhancing the statistical power to dissect complex traits from potentially any experimental population or selection experiment.

For AIL populations with smaller intermediate generations and multiple siblings or half sibs, correcting generation with a mixed or fixed effect model will likely result in overcorrection due to the correlation between generations and kinship. Standardization using z-scores provides an acceptable trade-off between overcorrecting and confidence in accounting for generational batch effects.

In conclusion, this study represents a more comprehensive study of the individual loci forming the genetic basis of the highly polygenic, long-term selection responses on 56-d body weight in the Virginia chicken lines, to date. It contributes not only to our current understanding of the genetic basis of body weight in chickens, but also provides a solid methodological foundation to further investigate the genetic architecture of complex traits in populations with similar designs.

ACKNOWLEDGMENTS

We thank Lars Rönnegård for the helpful comments and discussions regarding the stability of the minimization algorithm. The computations and data handling were enabled by resources in projects SNIC 2017/7-53, SNIC 2018-3-170 and SNIC 2020-5-14 provided by the Swedish National Infrastructure for Computing (SNIC) at UPPMAX, partially funded by the Swedish Research Council through grant agreement no. 2018-05973. The work was supported by the Swedish Research Council (grants 349-2005-8628, 621-2012-4634, 2017-3726 and 2018-5991) and FORMAS (grants 2013-450 and 2017-415).

The Funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

DISCLOSURES

The authors have declared that no competing interests exist.

SUPPLEMENTARY MATERIALS

Supplementary material associated with this article can be found in the online version at [doi:10.1016/j.psj.2022.102203](https://doi.org/10.1016/j.psj.2022.102203).

REFERENCES

- Altshuler, D., V. J. Pollara, C. R. Cowles, W. J. Van Etten, J. Baldwin, L. Linton, and E. S. Lander. 2000. An SNP map of the human genome generated by reduced representation shotgun sequencing. *Nature* 407:513–516.
- Andersson, L. 2001. Genetic dissection of phenotypic diversity in farm animals. *Nat. Rev. Genet.* 2:130–138.
- Andersson, L., and M. Georges. 2004. Domestic-animal genomics: deciphering the genetics of complex traits. *Nat. Rev. Genet.* 5:202–212.
- Andolfatto, P., D. Davison, D. Erezylmaz, T. T. Hu, J. Mast, T. Sunayama-Morita, and D. L. Stern. 2011. Multiplexed shotgun genotyping for rapid and efficient genetic mapping. *Genome Res.* 21:610–617.
- Auwerda, G. A. V. der, M. O. Carneiro, C. Hartl, R. Poplin, G. del Angel, A. Levy-Moonshine, T. Jordan, K. Shakir, D. Roazen, J. Thibault, E. Banks, K. V. Garimella, D. Altshuler, S. Gabriel, and M. A. DePristo. 2013. From FastQ data to high-confidence variant calls: the genome analysis toolkit best practices pipeline. *Curr. Protoc. Bioinforma.* 43:11.10.1–11.10.33, doi:10.1002/0471250953.bi1110s43.
- Benjamini, Y., and Y. Hochberg. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B Methodol.* 57:289–300.
- Besnier, F., P. Wahlberg, L. Rönnegård, W. Ek, L. Andersson, P. B. Siegel, and Ö. Carlborg. 2011. Fine mapping and replication of QTL in outbred chicken advanced intercross lines. *Genet. Sel. Evol.* 43:3.
- Boyle, E. A., Y. I. Li, and J. K. Pritchard. 2017. An expanded view of complex traits: from polygenic to omnigenic. *Cell* 169:1177–1186.
- Brandt, M., M. Ahsan, C. F. Honaker, P. B. Siegel, and Ö. Carlborg. 2017. Imputation-based fine-mapping suggests that most QTL in an outbred chicken advanced intercross body weight line are due to multiple, linked loci. *G3 GenesGenomesGenetics* 7:119–128.
- Broman, K. W., H. Wu, S. Sen, and G. A. Churchill. 2003. R/qtl: QTL mapping in experimental crosses. *Bioinforma. Oxf. Engl.* 19:889–890.
- Carlborg, Ö., L. Jacobsson, P. Åhgren, P. Siegel, and L. Andersson. 2006. Epistasis and the release of genetic variation during long-term selection. *Nat. Genet.* 38:418–420.
- Castro, J. P. L., M. N. Yancoskie, M. Marchini, S. Belohlavy, L. Hiramatsu, M. Kučka, W. H. Beluch, R. Naumann, I. Skuplik, J. Cobb, N. H. Barton, C. Rolian, and Y. F. Chan. 2019. An integrative genomic analysis of the Longshanks selection experiment for longer limbs in mice. *eLife* 8:e42014.
- Cock, P. J. A., T. Antao, J. T. Chang, B. A. Chapman, C. J. Cox, A. Dalke, I. Friedberg, T. Hamelryck, F. Kauff, B. Wilczynski, and M. J. L. de Hoon. 2009. Biopython: freely available Python tools for computational molecular biology and bioinformatics. *Bioinformatics* 25:1422–1423.
- Curtsinger, J. W., and A. A. Khazaeli. 2002. Lifespan, QTLs, age-specificity, and pleiotropy in *Drosophila*. *Mech. Ageing. Dev.* 123:81–93.
- DePristo, M. A., E. Banks, R. Poplin, K. V. Garimella, J. R. Maguire, C. Hartl, A. A. Philippakis, G. del Angel, M. A. Rivas, M. Hanna, A. McKenna, T. J. Fennell, A. M. Kernysky, A. Y. Sivachenko, K. Cibulskis, S. B. Gabriel, D. Altshuler, and M. J. Daly. 2011. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat. Genet.* 43:491–498.
- Dudley, J. W. 2007. From means to QTL: the illinois long-term selection experiment as a case study in quantitative genetics. *Crop Sci.* 47:S-20-S-31.
- Dunnington, E. A., C. F. Honaker, M. L. McGilliard, and P. B. Siegel. 2013. Phenotypic responses of chickens to long-term, bidirectional selection for juvenile body weight—Historical perspective. *Poult. Sci.* 92:1724–1734.
- Dunnington, E. A., and P. B. Siegel. 1996. Long-term divergent selection for eight-week body weight in white plymouth rock chickens. *Poult. Sci.* 75:1168–1179.
- Elferink, M. G., P. van As, T. Veenendaal, R. P. Crooijmans, and M. A. Groenen. 2010. Regional differences in recombination hotspots between two chicken populations. *BMC Genet.* 11:1–10.

- Flint, J., and R. Mott. 2001. Finding the molecular basis of quantitative traits: successes and pitfalls. *Nat. Rev. Genet.* 2:437–445.
- Forsberg, S. K. G., J. S. Bloom, M. J. Sadhu, L. Kruglyak, and Ö. Carlborg. 2017. Accounting for genetic interactions improves modeling of individual quantitative trait phenotypes in yeast. *Nat. Genet.* 49:497–503.
- Genome Reference Consortium. 2018. Chicken genome - Assembly GRC6a. Available at https://www.ncbi.nlm.nih.gov/assembly/GCF_000002315.5/ (verified 17 November 2020).
- Groenen, M. A. M., H. H. Cheng, N. Bumstead, B. F. Benkel, W. E. Briles, T. Burke, D. W. Burt, L. B. Crittenden, J. Dodgson, J. Hillel, S. Lamont, A. P. de Leon, M. Soller, H. Takahashi, and A. Vignal. 2000. A consensus linkage map of the chicken genome. *Genome Res.* 10:137–147.
- Groenen, M. A. M., P. Wahlberg, M. Foglio, H. H. Cheng, H.-J. Megens, R. P. M. A. Crooijmans, F. Besnier, M. Lathrop, W. M. Muir, G. K.-S. Wong, I. Gut, and L. Andersson. 2009. A high-density SNP-based linkage map of the chicken genome reveals sequence features correlated with recombination rate. *Genome Res.* 19:510–519.
- Guo, Y., M. Lillie, Y. Zan, J. Beranger, A. Martin, C. F. Honaker, P. B. Siegel, and Ö. Carlborg. 2019. A genomic inference of the White Plymouth Rock genealogy. *Poult. Sci.* 98:5272–5280.
- Hopkins, C. 1899. Improvement in the chemical composition of the corn kernel. *J. Am. Chem. Soc.* 21:1039–1057.
- Jacobsson, L., H.-B. Park, P. Wahlberg, R. Fredriksson, M. Perez-Enciso, P. B. Siegel, and L. Andersson. 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.
- Johansson, A. M., M. E. Pettersson, P. B. Siegel, and Ö. Carlborg. 2010. Genome-wide effects of long-term divergent selection. *PLoS Genet.* 6:e1001188.
- Knott, S. A., and C. S. Haley. 1992. Maximum likelihood mapping of quantitative trait loci using full-sib families. *Genetics* 132:1211–1222.
- Knott, S. A., L. Marklund, C. S. Haley, K. Andersson, W. Davies, H. Ellegren, M. Fredholm, I. Hansson, B. Hoyheim, K. Lundström, M. Moller, and L. Andersson. 1998. Multiple marker mapping of quantitative trait loci in a cross between outbred wild boar and large white pigs. *Genetics* 149:1069–1080.
- Lango Allen, H., K. Estrada, G. Lettre, S. I. Berndt, M. N. Weedon, F. Rivadeneira, C. J. Willer, A. U. Jackson, S. Vedantam, S. Raychaudhuri, T. Ferreira, A. R. Wood, R. J. Weyant, A. V. Segrè, E. K. Speliotes, E. Wheeler, N. Soranzo, J.-H. Park, J. Yang, D. Gudbjartsson, N. L. Heard-Costa, J. C. Randall, L. Qi, A. Vernon Smith, R. Mägi, T. Pastinen, L. Liang, I. M. Heid, J. Luan, G. Thorleifsson, T. W. Winkler, M. E. Goddard, K. Sin Lo, C. Palmer, T. Workalemahu, Y. S. Aulchenko, Å. Johansson, M. Carola Zillikens, M. F. Feitosa, T. Esko, T. Johnson, S. Ketkar, P. Kraft, M. Mangino, I. Prokopenko, D. Absher, E. Albrecht, F. Ernst, N. L. Glazer, C. Hayward, J.-J. Hottenga, K. B. Jacobs, J. W. Knowles, Z. Kutalik, K. L. Monda, O. Polasek, M. Preuss, N. W. Rayner, N. R. Robertson, V. Steinthorsdottir, J. P. Tyrer, B. F. Voight, F. Wiklund, J. Xu, J. Hua Zhao, D. R. Nyholt, N. Pellikka, M. Perola, J. R. B. Perry, I. Surakka, M.-L. Tammesoo, E. L. Altmaier, N. Amin, T. Aspelund, T. Bhangale, G. Boucher, D. I. Chasman, C. Chen, L. Coin, M. N. Cooper, A. L. Dixon, Q. Gibson, E. Grundberg, K. Hao, M. Juhani Juntila, L. M. Kaplan, J. Kettunen, I. R. König, T. Kwan, R. W. Lawrence, D. F. Levinson, M. Lorentzon, B. McKnight, A. P. Morris, M. Müller, J. Suh Ngwa, S. Purcell, S. Rafelt, R. M. Salem, E. Salvi, S. Sanna, J. Shi, U. Sovio, J. R. Thompson, M. C. Turchin, L. Vandenput, D. J. Verlaan, V. Vitart, C. C. White, A. Ziegler, P. Almgren, A. J. Balmforth, H. Campbell, L. Citterio, A. De Grandi, A. Dominiczak, J. Duan, P. Elliott, R. Elosua, J. G. Eriksson, N. B. Freimer, E. J. C. Geus, N. Glorioso, S. Haiqing, A.-L. Hartikainen, A. S. Havulinna, A. A. Hicks, J. Hui, W. Igl, T. Illig, A. Jula, E. Kajantie, T. O. Kilpeläinen, M. Koiranen, I. Kolcic, S. Koskinen, P. Kovacs, J. Laitinen, J. Liu, M.-L. Lokki, A. Marusic, A. Maschio, T. Meitinger, A. Mulas, G. Paré, A. N. Parker, J. F. Peden, A. Petersmann, I. Pichler, K. H. Pietiläinen, A. Pouta, M. Ridderstråle, J. I. Rotter, J. G. Sambrook, A. R. Sanders, C. Oliver Schmidt, J. Sinisalo, J. H. Smit, H. M. Stringham, G. Bragi Walters, E. Widen, S. H. Wild, G. Willemsen, L. Zagato, L. Zgaga, P. Zitting, H. Alavere, M. Farrall, W. L. McArdle, M. Nelis, M. J. Peters, S. Ripatti, J. B. J. van Meurs, K. K. Aben, K. G. Ardlie, J. S. Beckmann, J. P. Beilby, R. N. Bergman, S. Bergmann, F. S. Collins, D. Cusi, M. den Heijer, G. Eiriksdottir, P. V. Gejman, A. S. Hall, A. Hamsten, H. V. Huikuri, C. Iribarren, M. Kähönen, J. Kaprio, S. Kathiresan, L. Kiemeny, T. Kocher, L. J. Launer, T. Lehtimäki, O. Melander, T. H. Mosley Jr, A. W. Musk, M. S. Nieminen, C. J. O'Donnell, C. Ohlsson, B. Oostra, L. J. Palmer, O. Raitakari, P. M. Ridker, J. D. Rioux, A. Rissanen, C. Rivolta, H. Schunkert, A. R. Shuldiner, D. S. Siscovick, M. Stumvoll, A. Tönjes, J. Tuomilehto, G.-J. van Ommen, J. Viikari, A. C. Heath, N. G. Martin, G. W. Montgomery, M. A. Province, M. Kayser, A. M. Arnold, L. D. Atwood, E. Boerwinkle, S. J. Chanock, P. Deloukas, C. Gieger, H. Grönberg, P. Hall, A. T. Hattersley, C. Hengstenberg, W. Hoffman, G. Mark Lathrop, V. Salomaa, S. Schreiber, M. Uda, D. Waterworth, A. F. Wright, T. L. Assimes, I. Barroso, A. Hofman, K. L. Mohlke, D. I. Boomsma, M. J. Caulfield, L. Adrienne Cupples, J. Erdmann, C. S. Fox, V. Gudnason, U. Gyllenstein, T. B. Harris, R. B. Hayes, M.-R. Jarvelin, V. Mooser, P. B. Munroe, W. H. Ouwehand, B. W. Penninx, P. P. Pramstaller, T. Quertermous, I. Rudan, N. J. Samani, T. D. Spector, H. Völzke, H. Watkins, J. F. Wilson, L. C. Groop, T. Haritunians, F. B. Hu, R. C. Kaplan, A. Metspalu, K. E. North, D. Schlessinger, N. J. Wareham, D. J. Hunter, J. R. O'Connell, D. P. Strachan, H.-E. Wichmann, I. B. Borecki, C. M. van Duijn, E. E. Schadt, U. Thorsteinsdottir, L. Peltonen, A. G. Uitterlinden, P. M. Visscher, N. Chatterjee, R. J. F. Loos, M. Boehnke, M. I. McCarthy, E. Ingelsson, C. M. Lindgren, G. R. Abecasis, K. Stefansson, T. M. Frayling, and J. N. Hirschhorn. 2010. Hundreds of variants clustered in genomic loci and biological pathways affect human height. *Nature* 467:832–838.
- Li, H. 2011. A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinforma. Oxf. Engl.* 27:2987–2993.
- Li, H. 2013. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *ArXiv13033997 Q-Bio* Available at <http://arxiv.org/abs/1303.3997> (verified 17 November 2020).
- Li, H., B. Handsaker, A. Wysoker, P. B. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, and R. Durbin. 2009. 1000 Genome Project Data Processing Subgroup. 2009. The sequence alignment/map format and SAMtools. *Bioinforma. Oxf. Engl.* 25:2078–2079.
- Lillie, M., C. F. Honaker, P. B. Siegel, and Ö. Carlborg. 2019. Bidirectional selection for body weight on standing genetic variation in a chicken model. *G3 Genes Genomes Genet.* 9:1165–1173.
- Lillie, M., Z. Y. Sheng, C. F. Honaker, L. Andersson, P. B. Siegel, and Ö. Carlborg. 2018. Genomic signatures of 60 years of bidirectional selection for 8-week body weight in chickens. *Poult. Sci.* 97:781–790.
- Mackay, T. F. C. 2001. The genetic architecture of quantitative traits. *Annu. Rev. Genet.* 35:303–339.
- Mackay, T. F. C. 2014. Epistasis and quantitative traits: using model organisms to study gene–gene interactions. *Nat. Rev. Genet.* 15:22–33.
- Márquez, G. C., P. B. Siegel, and R. M. Lewis. 2010. Genetic diversity and population structure in lines of chickens divergently selected for high and low 8-week body weight. *Poult. Sci.* 89:2580–2588.
- McKenna, A., M. Hanna, E. Banks, A. Sivachenko, K. Cibulskis, A. Kernysky, K. Garimella, D. Altshuler, S. Gabriel, M. Daly, and M. A. DePristo. 2010. The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* 20:1297–1303.
- Pedersen, B. S., and A. R. Quinlan. 2017. cyvcf2: fast, flexible variant analysis with Python. *Bioinformatics* 33:1867–1869.
- Peirce, J. L., H. Li, J. Wang, K. F. Manly, R. J. Hitzemann, J. K. Belknap, G. D. Rosen, S. Goodwin, T. R. Sutter, R. W. Williams, and L. Lu. 2006. How replicable are mRNA expression QTL? *Mamm. Genome* 17:643–656.
- Pérille, F., C. Guerrero-Bosagna, V. H. da Silva, C. Boschiero, J. de R. da S. Nunes, M. C. Ledur, P. Jensen, and

- L. L. Coutinho. 2016. High-throughput and cost-effective chicken genotyping using next-generation sequencing. *Sci. Rep.* 6:26929.
- Pettersson, M., F. Besnier, P. B. Siegel, and Ö. Carlborg. 2011. Replication and explorations of high-order epistasis using a large advanced intercross line pedigree. *PLoS Genet.* 7:e1002180.
- Pettersson, M. E., A. M. Johansson, P. B. Siegel, and Ö. Carlborg. 2013. Dynamics of adaptive alleles in divergently selected body weight lines of chickens. *G3 Genes Genomes Genet.* 3:2305–2312.
- R Core Team, R. C. T. 2013. R: a language and environment for statistical computing.
- Rowan, B. A., V. Patel, D. Weigel, and K. Schneeberger. 2015. Rapid and inexpensive whole-genome genotyping-by-sequencing for crossover localization and fine-scale genetic mapping. *G3 Genes Genomes Genet.* 5:385–398.
- Seabold, S., and J. Perktold. 2010. Statsmodels: econometric and statistical modeling with Python. *Proc. 9th Python Sci. Conf* (pp. 92–96)92–96.
- Sheng, Z., M. E. Pettersson, C. F. Honaker, P. B. Siegel, and Ö. Carlborg. 2015. Standing genetic variation as a major contributor to adaptation in the Virginia chicken lines selection experiment. *Genome Biol.* 16:219.
- Siegel, P. B. 1962. Selection for body weight at eight weeks of age: 1. Short term response and Heritabilities. *Poult. Sci.* 41:954–962.
- Wahlberg, P., Ö. Carlborg, M. Foglio, X. Tordoir, A.-C. Syvänen, M. Lathrop, I. G. Gut, P. B. Siegel, and L. Andersson. 2009. Genetic analysis of an F2 intercross between two chicken lines divergently selected for body-weight. *BMC Genomics* 10:248.
- Whalen, A., R. Ros-Freixedes, D. L. Wilson, G. Gorjanc, and J. M. Hickey. 2018. Hybrid peeling for fast and accurate calling, phasing, and imputation with sequence data of any coverage in pedigrees. *Genet. Sel. Evol.* 50:67.
- Wood, A. R., T. Esko, J. Yang, S. Vedantam, T. H. Pers, S. Gustafsson, A. Y. Chu, K. Estrada, J. Luan, Z. Kutalik, N. Amin, M. L. Buchkovich, D. C. Croteau-Chonka, F. R. Day, Y. Duan, T. Fall, R. Fehrmann, T. Ferreira, A. U. Jackson, J. Karjalainen, K. S. Lo, A. E. Locke, R. Mägi, E. Mihailov, E. Porcu, J. C. Randall, A. Scherag, A. A. E. Vinkhuyzen, H.-J. Westra, T. W. Winkler, T. Workalemahu, J. H. Zhao, D. Absher, E. Albrecht, D. Anderson, J. Baron, M. Beekman, A. Demirkan, G. B. Ehret, B. Feenstra, M. F. Feitosa, K. Fischer, R. M. Fraser, A. Goel, J. Gong, A. E. Justice, S. Kanoni, M. E. Kleber, K. Kristiansson, U. Lim, V. Lotay, J. C. Lui, M. Mangino, I. M. Leach, C. Medina-Gomez, M. A. Nalls, D. R. Nyholt, C. D. Palmer, D. Pasko, S. Pechlivanis, I. Prokopenko, J. S. Ried, S. Ripke, D. Shungin, A. Stancáková, R. J. Strawbridge, Y. J. Sung, T. Tanaka, A. Teumer, S. Trompet, S. W. van der Laan, J. van Setten, J. V. Van Vliet-Ostaptchouk, Z. Wang, L. Yengo, W. Zhang, U. Afzal, J. Ärnlöv, G. M. Arscott, S. Bandinelli, A. Barrett, C. Bellis, A. J. Bennett, C. Berne, M. Blüher, J. L. Bolton, Y. Böttcher, H. A. Boyd, M. Bruinenberg, B. M. Buckley, S. Buyske, I. H. Caspersen, P. S. Chines, R. A. Clarke, S. Claudi-Boehm, M. Cooper, E. W. Daw, P. A. De Jong, J. Deelen, G. Delgado, J. C. Denny, R. Dhonukshe-Rutten, M. Dimitriou, A. S. F. Doney, M. Dörr, N. Eklund, E. Eury, L. Folkersen, M. E. Garcia, F. Geller, V. Giedraitis, A. S. Go, H. Grallert, T. B. Grammer, J. Gräßler, H. Grönberg, L. C. P. G. M. de Groot, C. J. Groves, J. Haessler, P. Hall, T. Haller, G. Hallmans, A. Hannemann, C. A. Hartman, M. Hassinen, C. Hayward, N. L. Heard-Costa, Q. Helmer, G. Hemani, A. K. Henders, H. L. Hillege, M. A. Hlatky, W. Hoffmann, P. Hoffmann, O. Holmen, J. J. Houwing-Duistermaat, T. Illig, A. Isaacs, A. L. James, J. Jeff, B. Johansen, A. Johansson, J. Jolley, T. Juliusdottir, J. Junttila, A. N. Kho, L. Kinnunen, N. Klopp, T. Kocher, W. Kratzer, P. Lichtner, L. Lind, J. Lindström, S. Lobbens, M. Lorentzon, Y. Lu, V. Lyssenko, P. K. E. Magnusson, A. Mahajan, M. Maillard, W. L. McArdle, C. A. McKenzie, S. McLachlan, P. J. McLaren, C. Menni, S. Merger, L. Milani, A. Moayyeri, K. L. Monda, M. A. Morken, G. Müller, M. Müller-Nurasyid, A. W. Musk, N. Narisu, M. Nauck, I. M. Nolte, M. M. Nöthen, L. Oozageer, S. Pilz, N. W. Rayner, F. Renstrom, N. R. Robertson, L. M. Rose, R. Roussel, S. Sanna, H. Scharnagl, S. Scholtens, F. R. Schumacher, H. Schunkert, R. A. Scott, J. Sehmi, T. Seufferlein, J. Shi, K. Silventoinen, J. H. Smit, A. V. Smith, J. Smolonska, A. V. Stanton, K. Stirrups, D. J. Stott, H. M. Stringham, J. Sundström, M. A. Swertz, A.-C. Syvänen, B. O. Tayo, G. Thorleifsson, J. P. Tyrer, S. van Dijk, N. M. van Schoor, N. van der Velde, D. van Heemst, F. V. A. van Oort, S. H. Vermeulen, N. Verweij, J. M. Vonk, L. L. Waite, M. Waldenberger, R. Wennauer, L. R. Wilkens, C. Willenborg, T. Wilsgaard, M. K. Wojczynski, A. Wong, A. F. Wright, Q. Zhang, D. Arveiler, S. J. L. Bakker, J. Beilby, R. N. Bergman, S. Bergmann, R. Biffar, J. Blangero, D. I. Boomsma, S. R. Bornstein, P. Bovet, P. Brambilla, M. J. Brown, H. Campbell, M. J. Caulfield, A. Chakravarti, R. Collins, F. S. Collins, D. C. Crawford, L. A. Cupples, J. Danesh, U. de Faire, H. M. den Ruijter, R. Erbel, J. Erdmann, J. G. Eriksson, M. Farrall, E. Ferrannini, J. Ferrières, I. Ford, N. G. Forouhi, T. Forrester, R. T. Gansevoort, P. V. Gejman, C. Gieger, A. Golay, O. Gottesman, V. Gudnason, U. Gyllenstein, D. W. Haas, A. S. Hall, T. B. Harris, A. T. Hattersley, A. C. Heath, C. Hengstenberg, A. A. Hicks, L. A. Hindorf, A. D. Hingorani, A. Hofman, G. K. Hovingh, S. E. Humphries, S. C. Hunt, E. Hyppönen, K. B. Jacobs, M.-R. Jarvelin, P. Jousilahti, A. M. Jula, J. Kaprio, J. J. P. Kastelein, M. Kayser, F. Kee, S. M. Keinänen-Kiukkaanniemi, L. A. Kiemeny, J. S. Kooner, C. Kooperberg, S. Koskinen, P. Kovacs, A. T. Kraja, M. Kumari, J. Kuusisto, T. A. Lakka, C. Langenberg, L. Le Marchand, T. Lehtimäki, S. Lupoli, P. A. F. Madden, S. Männistö, P. Manunta, A. Marette, T. C. Matise, B. McKnight, T. Meitinger, F. L. Moll, G. W. Montgomery, A. D. Morris, A. P. Morris, J. C. Murray, M. Nelis, C. Ohlsson, A. J. Oldehinkel, K. K. Ong, W. H. Ouwehand, G. Pasterkamp, A. Peters, P. P. Pramstaller, J. F. Price, L. Qi, O. T. Raitakari, T. Rankinen, D. C. Rao, T. K. Rice, M. Ritchie, I. Rudan, V. Salomaa, N. J. Samani, J. Saramies, M. A. Sarzynski, P. E. H. Schwarz, S. Sebert, P. Sever, A. R. Shuldiner, J. Sinisalo, V. Steinthorsdottir, R. P. Stolk, J.-C. Tardif, A. Tönjes, A. Tremblay, E. Tremoli, J. Virtamo, M.-C. Vohl, P. Amouyel, F. W. Asselbergs, T. L. Assimes, M. Bochud, B. O. Boehm, E. Boerwinkle, E. P. Bottinger, C. Bouchard, S. Cauchi, J. C. Chambers, S. J. Chanock, R. S. Cooper, P. I. W. de Bakker, G. Dedoussis, L. Ferrucci, P. W. Franks, P. Froguel, L. C. Groop, C. A. Haiman, A. Hamsten, M. G. Hayes, J. Hui, D. J. Hunter, K. Hveem, J. W. Jukema, R. C. Kaplan, M. Kivimäki, D. Kuh, M. Laakso, Y. Liu, N. G. Martin, W. März, M. Melbye, S. Moebus, P. B. Munroe, I. Njølstad, B. A. Oostra, C. N. A. Palmer, N. L. Pedersen, M. Perola, L. Pérusse, U. Peters, J. E. Powell, C. Power, T. Quertermous, R. Rauramaa, E. Reinmaa, P. M. Ridker, F. Rivadeneira, J. I. Rotter, T. E. Saaristo, D. Saleheen, D. Schlessinger, P. E. Slagboom, H. Snieder, T. D. Spector, K. Strauch, M. Stumvoll, J. Tuomilehto, M. Uusitupa, P. van der Harst, H. Völzke, M. Walker, N. J. Wareham, H. Watkins, H.-E. Wichmann, J. F. Wilson, P. Zanen, P. Deloukas, I. M. Heid, C. M. Lindgren, K. L. Mohlke, E. K. Speliotes, U. Thorsteinsdottir, I. Barroso, C. S. Fox, K. E. North, D. P. Strachan, J. S. Beckmann, S. I. Berndt, M. Boehnke, I. B. Borecki, M. I. McCarthy, A. Metspalu, K. Stefansson, A. G. Uitterlinden, C. M. van Duijn, L. Franke, C. J. Willer, A. L. Price, G. Lettre, R. J. F. Loos, M. N. Weedon, E. Ingelsson, J. R. O'Connell, G. R. Abecasis, D. I. Chasman, M. E. Goddard, P. M. Visscher, J. N. Hirschhorn, and T. M. Frayling. 2014. Defining the role of common variation in the genomic and biological architecture of adult human height. *Nat. Genet.* 46:1173–1186.
- Yang, J., B. Benyamin, B. P. McEvoy, S. Gordon, A. K. Henders, D. R. Nyholt, P. A. Madden, A. C. Heath, N. G. Martin, G. W. Montgomery, M. E. Goddard, and P. M. Visscher. 2010. Common SNPs explain a large proportion of the heritability for human height. *Nat. Genet.* 42:565–569.
- Zan, Y., and Ö. Carlborg. 2020. Dynamic genetic architecture of yeast response to environmental perturbation shed light on origin of cryptic genetic variation. *PLoS Genet.* 16:e1008801.
- Zan, Y., T. Payen, M. Lillie, C. F. Honaker, P. B. Siegel, and Ö. Carlborg. 2019. Genotyping by low-coverage whole-genome sequencing in intercross pedigrees from outbred founders: a cost-efficient approach. *Genet. Sel. Evol.* 51:44.
- Zan, Y., Z. Sheng, M. Lillie, L. Rönnegård, C. F. Honaker, P. B. Siegel, and Ö. Carlborg. 2017. Artificial selection response

- due to polygenic adaptation from a multilocus, multiallelic genetic architecture. *Mol. Biol. Evol.* 34:2678–2689.
- Zhang, X., P. Pérez-Rodríguez, K. Semagn, Y. Beyene, R. Babu, M. A. López-Cruz, F. San Vicente, M. Olsen, E. Buckler, J.-L. Jannink, B. M. Prasanna, and J. Crossa. 2015. Genomic prediction in biparental tropical maize populations in water-stressed and well-watered environments using low-density and GBS SNPs. *Heredity* 114:291–299.