Genome-wide association study, replication, and mega-analysis using a dense marker panel in a multi-generational mouse advanced intercross line

Xinzhu Zhou¹, Celine L. St. Pierre², Natalia M. Gonzales³, Riyan Cheng⁴, Apurva Chitre⁴, Greta Sokoloff⁵, Abraham A. Palmer⁴,⁶*

¹ Biomedical Sciences Graduate Program, University of California San Diego, La Jolla, CA, USA;
² Department of Genetics, Washington University School of Medicine, St. Louis, MO, USA
³ Department of Human Genetics, University of Chicago, Chicago, IL, USA
⁴ Department of Psychiatry, University of California San Diego, La Jolla, CA, USA
⁵ Department of Psychological & Brain Sciences, University of Iowa, Iowa City, IO, USA
⁶ Institute for Genomic Medicine, University of California San Diego, La Jolla, CA, USA

*Corresponding author

Email: aapalmer@ucsd.edu (AAP)

Author contributions

XZ imputed genotypes, performed SNP and individual QC, and conducted GWAS in F₃⁴ and F₃⁹-⁴³ AILs under supervision of AAP. AAP also provided computational resources for the analyses in this paper. CS prepared GBS libraries for sequencing, as well as organizing portions of the F₃⁹-⁴³ phenotypes. NMG de-multiplexed GBS sequencing results and performed alignment and variant calling. RC helped with kinship relatedness matrix calculated from AIL pedigree. AC provided technical support for running programs and scripts. GS collected F₃⁹-⁴³ phenotypes, respectively. XZ co-wrote the manuscript with AAP, who designed the study and oversaw data collection.

Data Availability

All relevant data are within the paper and its Supporting Information files. The GeneNetwork accession numbers for genotypes and phenotypes of F₃⁴, F₃⁹-⁴³, and mega-analysis cohort of AIL are currently pending but will be secured prior to final acceptance/publication.
Abstract

Genome-wide association studies (GWAS) in multigenerational outbred populations offer improved mapping precision compared to traditional populations such as F$_2$s. Advanced intercross lines (AILs) are the simplest possible multigenerational intercross; AILs are produced by crossing two inbred strains and then breeding unrelated offspring for additional generations. New recombinations accumulate with each successive generation; these new recombinations provide increased mapping resolution. We used genotyping-by-sequencing (GBS) to re-genotype a cohort of sparsely genotyped F$_34$ LG/J x SM/J AIL mice that were the subject of several prior publications as well as to obtain genotypes for new cohort of AIL mice from the F$_{39-43}$ generations. The denser set of GBS markers allowed us to identify 110 significant loci, 36 of which were novel, for 79 behavioral and physiological traits in the 428 F$_{34}$ mice. Genetic correlations between F$_{34}$ and F$_{39-43}$ were high, though F$_{39-43}$ AILs showed systematically lower SNP-heritability estimates. We explored replication of loci identified in either F$_{34}$ or F$_{39-43}$ in the other cohort for the traits measured in both F$_{34}$ and F$_{39-43}$: locomotor activity, body weight, and coat color. While coat color loci were robustly replicated, we observed only partial replication of associations for locomotor activity and body weight. We then performed a mega-analysis of locomotor activity, body weight by combining F$_{34}$ and F$_{39-43}$ mice (N=1028), which identified four novel loci. Finally, we showed that imputation using the old, sparse marker set and the newer dense marker set had little impact on our results, emphasizing the need of the denser GBS genotypes. The present study provides empirical insights into replication, the utility of denser genotyping and identifies new candidate loci that can be explored in future studies.
Author summary

Using a sufficiently dense marker set is essential for genome-wide association studies. The required SNP density is a function of population history. When considering populations derived from a cross between two or more inbred strains, more recombinations will require a correspondingly greater number of markers. We performed several genome-wide association studies using different marker sets - sparse SNPs from a previously publish dataset, dense SNPs obtained using genotyping-by-sequencing, and imputed versions of both the sparse and dense SNPs - to compare their association results in 34th and 39-43rd generations of an LG/J x SM/J advanced intercross line (AIL). We found that the dense SNP set substantially increased the number of significant results, while the imputed sets did not. To achieve maximum power, we performed a mega-analysis using all the mice and the dense SNP set; this analysis identified several novel loci. Finally, we explored the replicability of associated loci discovered in either cohort by examining the same SNP in the other cohort. We have made these data publicly available on www.g nenetwork.org, providing the community with a unique reference dataset.
Introduction

The use of multi-parental crosses and commercially available outbred populations for genome wide association studies (GWAS) in model organisms such as mice [1–17], rats [18], chickens [19,20], zebrafish [21,22], fruit flies [23–27], C. elegans [28] and various plant species [29–31] has become increasingly popular over the last decade. These mapping populations can further be categorized as multi-parental crosses, which are created by interbreeding two or more inbred strains, and commercially available outbred populations, in which the founders are of unknown provenance. An $F_2$ cross between two inbred strains is the prototypical mapping population; however, $F_2$s provide poor mapping resolution [32]. To improve mapping resolution, Darvasi and Soller [33] proposed the creation of advanced intercross lines (AILs), which are produced by intercrossing $F_2$ mice for additional generations. AILs accumulate additional crossovers with every successive generation, leading to a population with shorter linkage disequilibrium (LD) blocks, which improves mapping precision, albeit at the expense of power [32,34].

The longest running mouse AIL was generated by crossing LG/J and SM/J inbred strains, which were selectively bred for large and small body size. We obtained this AIL in 2006 at generation 33 from Dr. James Cheverud (Jmc: LG,SM-G$_{33}$, Washington University in St. Louis). Since then, we have collected genotype and phenotype information from multiple generations, including $F_{34}$ [16,35–38], $F_{39}$-$F_{43}$ and $F_{50}$-$F_{56}$ [39]. Our previous publications using the $F_{34}$ generation used a custom Illumina Infinium genotyping microarray to obtain genotypes for 4,593 SNPs [35,36]. Although not previously published, we also collected phenotype information from the $F_{39}$-$F_{43}$ generations, including body weight, fear conditioning, locomotor activity in response to methamphetamine, and the light dark test for anxiety.
In the present study, we used genotyping-by-sequencing (GBS), which is a reduced-representation sequencing method [40–42], to obtain a much denser set of SNPs in the F_{34} and F_{39-43} generations for the first time. With this denser set of SNPs, we attempted to identify novel loci in the F_{34}s that were not detected using the sparse SNPs. We also performed a GWAS on using the mice from the F_{39-43} AILs. Because F_{39-43} AILs are descendants of the F_{34}, they are uniquely suited to be a replication population for GWAS in the F_{34} generation. Moreover, F_{39-43} may provide improved resolution and allow for the discovery of novel loci not detected in the F_{34} generation. We also performed a mega-analysis of F_{34} mice and F_{39-43} mice to identify loci that were not identified in either individual dataset. Finally, we explored whether imputation from the array SNPs could have provided the additional coverage we obtained using the denser GBS genotypes. This study investigates the effect of SNP density on association analyses, addresses rarely explored questions about replication in outbred rodent populations, and identifies novel loci associated with a range of physiological and behavioral traits in mice.

Results

We used 214 males and 214 females from generation F_{34} (Aap:LG,SM-G34) and 305 males and 295 females from generations F_{39-43}. We performed association studies on 79 traits in F_{34} AIL and 49 traits in F_{39-43} AIL (S1 Table). F_{34} mice had been previously genotyped on a custom SNP array [35,36]. The average minor allele frequency (MAF) of those 4,593 array SNPs was 0.388 (Fig 1). To obtain a denser set of SNP markers, we used genotyping-by-sequencing in F_{34} and F_{39-43} AIL mice. Since data about the F_{39-43} AIL mice had been collected over the span of approximately two years, we carefully considered the possibility of sample contamination and sample mislabeling [43]. We removed samples based on four major features:
heterozygosity distribution, number of reads aligned to sex chromosomes, discrepancies between pedigree and genetic kinship relatedness, and coat color genotype to phenotype mismatch (see Methods; S1 and S2 Figs). The final SNP sets included 60,392 GBS-derived SNPs in 428 F₃₄ AIL mice, 59,790 GBS-derived SNPs in 600 F₃⁹-₄₃ AIL mice, and 58,461 GBS-derived SNPs that existed in both F₃₄ and F₃⁹-₄₃ AIL mice (S2 Table). The MAF for the GBS SNPs was 0.382 in F₃₄, 0.358 in F₃⁹-₄₃, and 0.370 in F₃₄ and F₃⁹-₄₃ (Fig 1). There were 66 SNPs called from our GBS data that were also present on the genotyping array. The genotype concordance rate, which reflects the sum of errors from both sets of genotypes, was 95.4% (S3 Fig). We found that LD decay rates using F₃₄ array, F₃₄ GBS, F₃⁹-₄₃ GBS, and F₃₄ and F₃⁹-₄₃ GBS genotypes were generally similar to one another, though levels of LD using the GBS genotypes appear to be slightly reduced in the later generations of AILs (S4 Fig).

**GBS genotypes produced more significant associations than array genotypes in F₃₄**

We used a linear mixed model (LMM) as implemented in GEMMA [44] to perform GWAS. We used the leave-one-chromosome-out (LOCO) approach to address the problem of proximal contamination, as previously described [39,45–47]. We performed GWAS using both the sparse array SNPs and the dense GBS SNPs to determine whether additional SNPs would produce additional genome-wide significant associations. Autosomal and X chromosome SNPs were included in all GWAS. We obtained adjusted significance threshold for each SNP set using MultiTrans and SLIDES [48,49]. To select independently associated loci (“lead loci”), we used a LD-based clumping method implemented in PLINK to group SNPs that passed the adjusted genome-wide significance thresholds over a large genomic region flanking the index SNP [50]. Applying the most stringent clumping parameters ($r^2 = 0.1$ and sliding window size = 12,150kb, S3 Table), we identified 110 significant lead loci in 49 out of 79 F₃₄ phenotypes using the GBS

6
SNPs. In contrast, we identified 83 significant lead loci in 45 out of 79 F$_{34}$ phenotypes using the sparse array SNPs. Among F$_{34}$ lead loci, 36 were uniquely identified in the GBS GWAS, whereas 11 were unique to the array GWAS. GBS SNPs consistently yielded more significant lead loci compared to array SNPs regardless of the clumping parameter values (S3 Table), indicating that a dense marker panel was able to detect more association signals compared to a sparse marker panel. Significant lead loci in F$_{34}$ GBS and array are summarized in Table S4.

We considered whether the disparity between the numbers of loci identified by the two SNP sets could be resolved by imputation. We used LG/J and SM/J whole genome sequencing data as reference panels [51] and performed imputation on array and GBS SNPs using Beagle v4.1 [52]. After QC filtering, we obtained 4.3M SNPs imputed from the array SNPs and 4.1M SNPs imputed from the GBS SNPs. More imputed GBS SNPs were filtered out because GBS SNPs were called from genotype probabilities, thus introducing uncertainty in imputed SNPs. We found that imputed array genotypes and imputed GBS genotypes did not meaningfully increase the number of new loci discoveries (Fig S5), indicating that the greater number of genome-wide significant associations obtained using the GBS SNPs could not be overcome by imputation.

Under a polygenic model where a large number of additive common variants contribute to a complex trait, heritability estimates could be higher when more SNPs are considered [53]. Given that there were more GBS SNPs than array SNPs, we used autosomal SNPs to examine whether GBS SNPs would generate higher SNP heritability estimates compared to the sparse array SNPs. Heritability estimates were similar for the two SNP sets, with the exception of agouti coat color, which showed marginally greater heritability for the GBS SNPs (S6 Fig; S5 Table). Our results show that while the denser GBS SNP set was able to identify more genome-wide significant loci, greater SNP density did not improve the polygenic signal.
Genetic correlation and SNP heritability of traits in F_{34} and F_{39-43}

One goal of collecting phenotypic information in the F_{39-43} generations was to evaluate whether or not the genetic signal driving the phenotypes are consistent across cohorts. To this end, we used two approaches to evaluate replication between F_{34} and F_{39-43}. First, we used autosomal SNPs to calculate genetic correlations between the F_{34} and F_{39-43} generations for body weight, coat color, and locomotor activity phenotypes (S6 Table). Locomotor activity on days 1 and 2, albino and agouti coat color, and body weight were highly genetically correlated \( r_G > 0.7 \). In contrast, locomotor activity on day 3 showed a significant but weaker genetic correlation \( r_G = 0.577 \), perhaps reflecting variability in the quality of the methamphetamine injection, which was not given on days 1 and 2. Overall, these results suggest that genetic influences on these traits were broadly similar in the two cohorts. Despite the relatively robust genetic correlations, SNP heritability was consistently lower in the F_{39-43} compared to the F_{34}, possibly a result of increased experimental variance introduced by our extended phenotype collection period (Fig 2; S7 Table).

Replication of loci detected in F_{34} using F_{39-43}

Using the dense GBS SNP set, we identified 27 genome-wide significant loci in 21 out of 49 F_{39-43} phenotypes. A subset of those traits, including coat color, body weight, and locomotor activity, were also phenotyped in the F_{34} AILs (S8 Table). To assess replication, we determined whether loci that were significant in one cohort (either F_{34} or F_{39-43}) would also be significant in the other. We termed the cohort in which a locus was initially discovered as its “discovery set” and the cohort we attempted replication in as the “replication set” (Table 1). Coat color phenotypes (both albino and agouti) are Mendelian traits and thus served as positive control. As expected, all body weight loci replicated. All three body weight loci identified in the F_{34} were
replicated at nominal levels of significance ($p<0.05$) in $F_{39-43}$; similarly, the only body weight
locus identified in $F_{39-43}$ was replicated in $F_{34}$. However, none of the five locomotor activity
associated loci were replicated in the reciprocal (replication) cohorts. We also used the “sign
test” to determine whether the directions of the effect (beta) of the loci were in the same
direction between the discovery and replication cohorts. We found that 11 of 13 loci passed this
much less stringent test of replication. The two loci that did not pass the sign test were the two
locomotor loci “discovered” in $F_{39-43}$ (Table 1).
Table 1. Replication of significant SNPs between F\textsubscript{34} and F\textsubscript{39-43} AIL association analyses. “Discovery set” indicates the AIL generation that significant SNPs were identified. “Replication set” shows the association p-value, \(\beta\) estimates, etc. of the “discovery set” significant SNPs in the replication AIL generation. SNPs that replicated (p<0.05, same sign for the beta) between F\textsubscript{34} and F\textsubscript{39-43} are highlighted in bold italics. Genetic correlations for phenotypes measured in both F\textsubscript{34} and F\textsubscript{39-43} are listed (see also Supplementary Table 6).

| Phenotype                                      | Discovery set                                                                 | Replication set                                                                 |
|-----------------------------------------------|-------------------------------------------------------------------------------|---------------------------------------------------------------------------------|
|                                               | SNP                             | P      | af    | beta | se    | P      | af    | beta | se    |
|                                               |                                 | F\textsubscript{34} GBS |       |       |       | F\textsubscript{3943 GBS} replicate |       |       |       |
| Body weight                                   | chr4.664414508                  | 8.06E-08 | 0.419 | -2.50E-01 | 4.56E-02 | 1.56E-02 | 0.406 | -1.05E-01 | 4.33E-02 |
|                                              | chr6.1405109                    | 5.89E-06 | 0.497 | 2.06E-01 | 4.51E-02 | 2.36E-02 | 0.518 | 9.48E-02 | 4.18E-02 |
|                                              | chr14.79312393                  | 7.53E-06 | 0.514 | -2.01E-01 | 4.45E-02 | 1.44E-02 | 0.566 | -1.04E-01 | 4.26E-02 |
|                                              | chr7.87642045                   | 5.00E-106 | 0.432 | -5.81E-01 | 1.91E-02 | 2.8E-163 | 0.387 | -6.07E-01 | 1.55E-02 |
|                                              | chr2.154464466                  | 9.43E-191 | 0.129 | 9.39E-01 | 1.25E-02 | 5.70E-93 | 0.207 | 7.20E-01 | 2.57E-02 |
|                                              | chr19.21812298                  | 9.28E-07 | 0.461 | -6.90E+02 | 1.39E+02 | 7.72E-01 | 0.52 | -1.94E-02 | 6.74E-02 |
|                                              | chr7.45084416                   | 1.12E-05 | 0.246 | 6.77E+02 | 1.53E+02 | 3.40E-01 | 0.217 | 7.38E-02 | 7.77E-02 |
|                                              | chr8.17410225                   | 4.96E-06 | 0.171 | 7.91E+02 | 1.71E+02 | 4.03E-01 | 0.192 | 7.04E-02 | 8.43E-02 |
|                                              | chr14.82586326                  | 2.63E-06 | 0.658 | -2.09E-01 | 4.43E-02 | 2.87E-05 | 0.575 | -1.89E-01 | 4.50E-02 |
|                                              | chr7.87255356                   | 5.91E-167 | 0.389 | -6.24E-01 | 1.57E-02 | 7.80E-97 | 0.444 | -5.75E-01 | 2.07E-02 |
|                                              | chr2.155091628                  | 1.78E-115 | 0.218 | 7.42E-01 | 2.17E-02 | 1.51E-185 | 0.135 | 8.98E-01 | 1.33E-02 |
|                                              | chr15.67235072                  | 3.21E-06 | 0.47 | 3.14E-01 | 6.63E-02 | 1.69E-01 | 0.522 | -1.72E+02 | 1.25E+02 |
|                                              | chr7.113250866                  | 5.88E-06 | 0.389 | 3.29E-01 | 7.20E-02 | 8.27E-01 | 0.483 | -1.21E+02 | 5.45E+02 |
Due to the high genetic correlations between cohorts (S6 Table), we suspected that a mega-analysis using the combined sample set would allow for the identification of additional loci. This analysis identified four novel genome-wide significant associations that were not discovered in either individual cohort (Fig 3; Table 2; S9 Table).
Table 2. Lead SNPs in F$_{34}$ and F$_{39-43}$ (N=1028) mega-analysis in locomotor activity (day 1 – 3) and body weight. Significant SNPs are clumped using parameters $r^2$=0.1, 12150kb. Rows in bold indicate SNPs that were discovered in the mega-analysis but not in individual F$_{34}$ or F$_{39-43}$ AILs.

| Phenotype                                      | SNP            | P       | af  | Beta    | se     | Discovery set                                      |
|------------------------------------------------|----------------|---------|-----|---------|--------|--------------------------------------------------|
| Locomotor test day 1, total distance travelled in 30min | chr1.122479820 | 8.84E-06 | 0.292 | -2.03E-01 | 4.56E-02 | Combined set                                     |
| Locomotor test day 1, total distance travelled in 30min | chr10.104988207 | 1.94E-07 | 0.373 | 2.16E-01  | 4.13E-02 | Combined set                                     |
| Locomotor test day 2, total distance travelled in 30min | chr1.40907532 | 2.65E-07 | 0.664 | 2.27E-01  | 4.38E-02 | Combined set                                     |
| Locomotor test day 2, total distance travelled in 30min | chr1.123997667 | 3.53E-08 | 0.289 | -2.53E-01 | 4.56E-02 | Combined set                                     |
| Body weight                                     | chr2.157502044 | 1.10E-05 | 0.42 | -1.35E-01 | 3.06E-02 | Combined set                                     |
| Body weight                                     | chr4.66866758 | 1.64E-09 | 0.403 | -1.84E-01 | 3.01E-02 | F$_{34}$ array, F$_{34}$ GBS, combined set       |
| Body weight                                     | chr6.81267890 | 8.61E-07 | 0.507 | 1.40E-01  | 2.83E-02 | F$_{34}$ GBS, combined set                        |
| Body weight                                     | chr14.82672838| 2.06E-10 | 0.623 | -1.90E-01 | 2.96E-02 | F$_{34}$ array, F$_{34}$ GBS, F$_{39-43}$ GBS, combined set |
Outbred model organism populations have become important resources for genetics studies. We used $F_{34}$ and $F_{39-43}$ generations of a LG/J x SM/J AIL to investigate how SNP density influences GWAS results, heritability, and genetic correlations. Using the ~60K SNPs in the $F_{34}$ GBS set, we found 110 loci, of which 36 were unique to the GBS genotypes. Notably, $F_{34}$ array genotypes were able to detect the majority of loci that were identified in $F_{34}$ GBS SNP set, indicating that many but not all loci were successfully ‘tagged’ by the sparse array SNPs. Although the genetic correlations between traits measured in $F_{34}$ and $F_{39-43}$ were high, replication was incomplete, with the locomotor traits showing especially poor replication. Finally, imputation to reference panels increased the number of SNPs available for analysis but did not meaningfully enhance the number of loci we discovered because it did not improve our ability to capture recombination events.

Previous work from our lab used the array genotypes described in this study to map QTLs for various behavioral and physiological traits in 688 $F_{34}$ AILs. With $F_{34}$ GBS SNPs, which were only available for 428 of the initial 688 AIL mice, we were able to replicate many of our previously published loci. For instance, using the array SNPs in the $F_{34}$ generation in conjunction with the LMM software QTLRel, Cheng et al. [35] found a significant locus on chromosome 8 for locomotor day 2 activity that contained only one gene: Csmd1 (CUB and sushi multiple domains 1). Gonzales et al. [39] replicated this finding in $F_{50-56}$ AIL and identified a cis-eQTL mapped to the same region. Csmd1 mutant mice showed increased locomotor activity compared to wild-type and heterozygous mice, indicating that Csmd1 is likely a causal gene for locomotor and related traits [39]. We replicated this locus in our own analysis of the $F_{34}$ array and GBS SNP, using GEMMA rather than QTLRel (S7 Fig). We also replicated a locus on chromosome 17 for distance traveled in the periphery in the open field test (Fig 4; [36,39]), three loci on chromosomes 4, 6, and 14 for body weight (Supplementary Fig 7; [54]), one locus on
chromosome 7 for mean corpuscular hemoglobin concentrations (MCHC, complete blood count; S7 Fig; [55]), and numerous loci on chromosome 4, 6, 7, 8, and 11 for muscle weights (Supplementary Fig 7; [37]). We noticed that while using the same set of F34 array genotypes, some previously published loci were not detected in our GWAS. The most likely explanation is that we had only 428 of the 688 mice used in the previous publications in this study. Methodological differences between prior studies and the current study, such as the use of QTLRel rather than GEMMA and the choice of pedigree rather than genotypes in relatedness calculation, may also lead to lack of complete replication [56].

F39-43 AILs replicated some, but not all, significant loci identified in F34, despite generally high genetic correlations between the two cohorts. Significant loci for coat color, which are monogenic and viewed as positive controls, were consistent between the two AIL cohorts. Loci for body weight were fully replicated between F34 and F39-43 using the p<0.05 cutoff for the replication cohort, whereas loci for locomotor activity were not. Nevertheless, the beta estimates for all but two loci shared the same sign, which provides modest evidence for replication. The circumstances under which the F39-43 data were collected were not ideal, likely contributing to their lower heritability and the limited replication. Unlike the F34 dataset, the F39-43 used multiple technicians to conduct the behavioral tests and occurred over a prolonged period of time in which numerous environmental factors may have changed.

We performed a mega-analysis using F34 and F39-43 AIL mice. The combined dataset increased our power and allowed us to identify four novel genome-wide significant associations that were not detected in either the F34 or the F39-43 cohorts. For example, the mega-analysis identified a locus for body weight on chromosome 2 (Fig S7). Parker et al. [54] identified the same locus using an integrated analysis of LG/J x SM/J F2 and F34 AILs.
Many GWAS use a 1.0~2.0 LOD support interval to approximate the size of the association region (see [57,58]). The LOD support interval, proposed by Conneally et al. [59] and Lander & Botstein [60], is a simple confidence interval method involving converting the p-value of the peak locus into a LOD score, subtracting “drop size” from the peak locus LOD score, and finding the two physical positions to the left and to the right of the peak locus location that correspond to the subtracted LOD score. Although Mangin et al. [61] showed via simulation that the boundaries of LOD support intervals depend on effect size, others observed that a 1.0 ~ 2.0 LOD support interval accurately captures ~95% coverage of the true location of the loci when using a dense set of markers [60,62,63]. In the present study, we considered using LOD support intervals but found that the sparse array SNPs produced misleadingly large support intervals. Various methods have been proposed for calculating confidence intervals in analogous situations (e.g. [12,64]). Rather than adopting a formal method, we compared LocusZoom plots of the same locus region between array SNPs and the GBS SNPs (S7 Fig; [65]). For example, the benefit of the denser SNP coverage is easily observed in the locus on chromosome 7 for “complete blood count, repeat measure of retic parameters, cell hemoglobin concentration mean” (Fig S7). Thus, there are advantages of dense SNP sets that go beyond the ability to discover additional loci (Table 1).

Our study has notable limitations. First, not all $F_{34}$ and $F_{39-43}$ animals that were phenotyped were later genotyped by GBS due to missing tissue samples, which in turn lowered our sample size and reduced the power of association analyses. Second, $F_{39-43}$ traits have been collected by different technicians over the span of several years, which introduced noise and diminished trait heritability (Fig 2).

The present study explored the role of marker density and imputation in GWAS. Furthermore, the combination of denser marker coverage and the addition of 600 $F_{39-43}$ AIL mice allowed us to identify novel loci for locomotor activity, open field test, fear conditioning, light dark
test for anxiety, complete blood count, iron content in liver and spleen, and muscle weight. Our
findings add to the large body of phenotypic/genotypic data available for the LG/SM AIL, which
can be found on GeneNetwork [http://www.genenetwork.org].

Materials and Methods

Animals

All mice used in this study were members of the advanced intercross line (AIL) between
LG/J and SM/J that was originally created by Dr. James Cheverud (Washington University, St.
Louis, MO). The AIL line has been maintained in the Palmer laboratory since generation F$_{33}$.
Age and exact number of animals tested in each phenotype are described in S1 Table. Several
previous publications [16,35–38,54,55] have reported on association analyses of the F$_{34}$ mice
(N=428). No prior publications have described the F$_{39-43}$ generations (N=600). The sample size
of F$_{34}$ mice reported in this study (N=428) is smaller than that in previous publications of F$_{34}$
(N=688) because we only sequenced a subset of F$_{34}$ animals using GBS. With the exception of
coop color and locomotor activities, we quantile normalized all phenotypes. Coat color traits
were coded in binary numbers (albino: 1 = white, 0 = non-white; agouti: 1 = tan, 0 = black,
NA=white). Locomotor activity traits were not quantile transformed in order to follow the
guideline described in Cheng et al. [35] for direct comparison.

F$_{34}$, F$_{39-43}$ Phenotypes

We have previously described the phenotyping of F$_{34}$ animals for locomotor activity [35],
fear conditioning [36], open field [36], coat color, body weight [54], complete blood counts [55],
heart and tibia measurements [37], muscle weight [37]. Iron content in liver and spleen, which
have not been previously reported in these mice, was measured by atomic absorption spectrophotometry, as described in Gardenghi et al. [66] and Graziano, Grady and Cerami [67]. Although the phenotyping of F\textsubscript{39-43} animals has not been previously reported, we used method that were identical to those previously reported for locomotor activity [35], open field [36], coat color, body weight [54], and light/dark test for anxiety [15].

\section*{F\textsubscript{34} AIL Array Genotypes}

\textit{F\textsubscript{34}} animals had been genotyped on a custom SNP array on the Illumina Infinium platform [35,36], which yielded a set of 4,593 SNPs on autosomes and X chromosome that we refer to as ‘array SNPs’.

\section*{F\textsubscript{34} and F\textsubscript{39-43} GBS Genotypes}

\textit{F\textsubscript{34}} and \textit{F\textsubscript{39-43}} animals were genotyped using genotyping-by-sequencing (GBS), which is a reduced-representation genome sequencing method [1,39]. We used the same protocol for GBS library preparation that was described in Gonzales et al [39]. We called GBS genotype probabilities using ANGSD [68]. GBS identified 1,667,920 autosomal and 43,015 X-chromosome SNPs. To fill in missing genotypes at SNPs where some but not all mice had calls, we ran within-sample imputation using Beagle v4.1, which generated hard call genotypes as well as genotype probabilities [52]. After imputation, only SNPs that had dosage $r^2 > 0.9$ were retained. We removed SNPs with minor allele frequency < 0.1 and SNPs with $p < 1.0 \times 10^{-6}$ in the Chi-square test of Hardy–Weinberg Equilibrium (HWE) (S2 Table). All phenotype and GBS genotype data are deposited in GeneNetwork (http://www.genenetwork.org).
QC of individuals

We have found that large genetic studies are often hampered by cross-contamination between samples and sample mix-ups. We used four features of the data to identify problematic samples: heterozygosity distribution, proportion of reads aligned to sex chromosomes, pedigree/kinship, and coat color. We first examined heterozygosity across autosomes and removed animals where the proportion of heterozygosity that was more than 3 standard deviations from the mean (S1 Fig). Next, we sought to identify animals in which the recorded sex did not agree with the sequencing data. We compared the ratio of reads mapped to the X and Y chromosomes. The 95% CI for this ratio was 196.84 to 214.3 in females and 2.13 to 2.18 in males. Twenty-two $F_{34}$ and $F_{39-43}$ animals were removed because their sex (as determined by reads ratio) did not agree with their recorded sex; we assumed this discrepancy was due to sample mix-ups. To further identify mislabeled samples, we calculated kinship coefficients based on the full AIL pedigree using QTLRel. We then calculated a genetic relatedness matrix (GRM) using IBDLD, which estimates identity by descent using genotype data. The comparison between pedigree kinship relatedness and genetic kinship relatedness identified 7 pairs of animals that showed obvious disagreement between kinship coefficients and the GRM. Lastly, we excluded 14 $F_{39-43}$ animals that showed discordance between their recorded coat color and their genotypes at markers flanking Tyr, which causes albinism in mice. The numbers of animals filtered at each step are listed in S2 Table. Some animals were detected by more than one QC step, substantiating our belief that these samples were erroneous.

At the end of SNP and sample filtering, we had 59,561 autosomal and 831 X chromosome SNPs in $F_{34}$, 58,966 autosomal and 824 X chromosome SNPs in $F_{39-43}$, and 57,635 autosomal and 826 X chromosome SNPs in the combined $F_{34}$ and $F_{39-43}$ set (S2 Table). GBS genotype quality was estimated by examining concordance between the 66 SNPs that were present in both the array and GBS genotyping results.
LD decay

Average LD ($r^2$) was calculated using allele frequency matched SNPs (MAF difference < 0.05) within 100,000bp distance, as described in Parker et al. [1].

Imputation to LG/J and SM/J reference panels

$F_{34}$ array genotypes (N=428) and $F_{34}$ GBS genotypes (N=428) were imputed to LG/J and SM/J whole genome sequence data [51] using BEAGLE. For $F_{34}$ array imputation, we used a large window size (100,000 SNPs and 45,000 SNPs overlap). Imputation to reference panels yielded 4.3 million SNPs for $F_{34}$ array and $F_{34}$ GBS imputed sets. Imputed SNPs with $DR^2 > 0.9$, $MAF > 0.1$, HWE p value $> 1.0 \times 10^{-6}$ were retained, resulting in 4.1M imputed $F_{34}$ GBS SNPs and 4.3M imputed $F_{34}$ array SNPs.

Genome-wide association analysis (GWAS)

We used the linear mixed model, as implemented in GEMMA [44], to perform a GWAS that accounted for the complex familial relationships among the AIL mice [35,39]. We used the leave-one-chromosome-out (LOCO) approach to calculate genetic relatedness matrix, which effectively circumvented the problem of proximal contamination [46]. Separate GWAS were performed using $F_{34}$ array genotypes, $F_{34}$ GBS genotypes, and $F_{39-43}$ GBS genotypes. Apart from coat color (binary trait) and locomotor activity, raw phenotypes were quantile normalized prior to analysis. Locomotor activity was not quantile normalized because the trait was reasonably normally distributed already and because we wanted our analysis to match those performed in Cheng et al [35]. Because $F_{34}$ AIL had already been studied using array genotypes [35] and mapped using QTLRel [56], we used the same covariates as described in Cheng et al.
[35] in order to examine whether our array and GBS GWAS would replicate their findings. We included sex and body weight as covariates for locomotor activity traits (see covariates used in [35]) and sex, age, and coat color as covariates for fear conditioning and open field test in F_{34} AILs (see covariates used in [36]). We used sex and age as covariates for all other phenotypes. Covariates for each analysis are shown in S1 Table. Finally, we performed mega-analysis of F_{34} and F_{39-43} animals (N=1,028) for body weight, coat color, and locomotor activity, since these traits were measured in the same way in both cohorts. For the mega-analyses, locomotor activity was quantile normalized after the combination of the two datasets to ensure that data were normally distributed across generations.

Identifying suspicious SNPs

Some significant SNPs in F_{34} GWAS and in the mega-analysis of F_{34} and F_{39-43} were suspicious because nearby SNPs, which would have been expected to be in high LD (a very strong assumption in an AIL), did not have high -log10 values. We only examined SNPs that obtained significant p-values; these examinations revealed that these SNPs had suspicious ratios of heterozygotes to homozygotes calls and had corresponding HWE p-values that were close to our 1.0×10^{-6} threshold (S10 and S11 Tables). To avoid counting these as novel loci, we removed those SNPs prior to summarizing our results as they likely reflected genotyping errors.

Selecting independent significant SNPs

To identify independent “lead loci” among significant GWAS SNPs that surpass the significance threshold, we used the LD-based clumping method in PLINK v1.9. We empirically chose clumping parameters (r^2 = 0.1 and sliding window size = 12,150kb) that gave us a
conservative set of independent SNPs (S3 Table). For the coat color phenotypes, we found that
multiple SNPs remained significant even after LD-based clumping, presumably due to the
extremely significant associations at these Mendelian loci. In these cases, we used a stepwise
model selection procedure in GCTA [69] and performed association analyses conditioning on
the most significant SNPs.

Significance thresholds

We used MultiTrans and SLIDE to set significance thresholds for the GWAS [48,49].
MultiTrans and SLIDE are methods that assume multivariate normal distribution of the
phenotypes, which in LMM models, contain a covariance structure due to various degrees of
relatedness among individuals. We were curious to see whether MultiTrans/SLIDE produces
significance thresholds drastically different from the threshold we obtained from a standard
permutation test (‘naïve permutation’ as per Cheng et al. [46]). We performed 1,000
permutations using the F$_{34}$ GBS genotypes and the phenotypic data from locomotor activity
(days 1, 2, and 3). We found that the 95th percentile values for these permutations were 4.65,
4.79, and 4.85, respectively, which were very similar to 4.85, the threshold obtained from
MultiTrans using the same data. Thus, the thresholds presented here were obtained from
MultiTrans but are similar (if anything slightly more conservative) than thresholds we would have
obtained had we used permutation. Because the effective number of tests depends on the
number of SNPs and the specific animals used in GWAS, we obtained a unique adjusted
significance threshold for each SNP set in each animal cohort (S12 Table).
Genetic correlation and heritability estimates between F_{34} and F_{39-43} phenotypes

Locomotor activity, body weight, and coat color had been measured in both F_{34} and F_{39-43} populations. We calculated both SNP heritability and genetic correlations between F_{34} and F_{39-43} animals using GCTA bivariate GREML analysis [69]. Because F_{39-43} day 1 locomotor activity data were not normally distributed, we quantile normalized locomotor activity data when estimating SNP heritabilities and genetic correlations.

LocusZoom Plots

LocusZoom plots were generated using the standalone implementation of LocusZoom [65], using LD scores calculated from PLINK v.1.9 --ld option and mm10 gene annotation file downloaded from USCS genome browser.

Ethics Statement

All procedures were approved by the Institutional Animal Care and Use Committee (IACUC protocol: S15226) Euthanasia was accomplished using CO_{2} asphyxiation followed by cervical dislocation.

Acknowledgements

We would like to recognize Jackie Lim and Kaitlin Samocha for collecting F_{34} AIL phenotype data and Ryan Walters for collecting F_{39-43} AIL phenotype data. We wish to acknowledge Alex Gileta for input on a draft of this manuscript.
References

1. Parker CC, Gopalakrishnan S, Carbonetto P, Gonzales NM, Leung E, Park YJ, et al. Genome-wide association study of behavioral, physiological and gene expression traits in outbred CFW mice. Nat Genet. 2016;48(8):919.

2. Gatti DM, Svenson KL, Shabalin A, Wu L-Y, Valdar W, Simecek P, et al. Quantitative trait locus mapping methods for diversity outbred mice. G3 Genes Genomes Genet. 2014;4(9):1623–1633.

3. Chesler EJ. Out of the bottleneck: the Diversity Outcross and Collaborative Cross mouse populations in behavioral genetics research. Mamm Genome. 2014;25(1–2):3–11.

4. Churchill GA, Gatti DM, Munger SC, Svenson KL. The diversity outbred mouse population. Mamm Genome. 2012;23(9–10):713–718.

5. Consortium CC. The genome architecture of the Collaborative Cross mouse genetic reference population. Genetics. 2012;190(2):389–401.

6. Talbot CJ, Nicod A, Cherny SS, Fulker DW, Collins AC, Flint J. High-resolution mapping of quantitative trait loci in outbred mice. Nat Genet. 1999;21(3):305.

7. Demarest K, Koyner J, McCaughran J, Cipp L, Hitzemann R. Further characterization and high-resolution mapping of quantitative trait loci for ethanol-induced locomotor activity. Behav Genet. 2001;31(1):79–91.

8. Valdar W, Solberg LC, Gauguier D, Burnett S, Klenerman P, Cookson WO, et al. Genome-wide genetic association of complex traits in heterogeneous stock mice. Nat Genet. 2006;38(8):879.

9. Ghazalpour A, Doss S, Kang H, Farber C, Wen P-Z, Brozell A, et al. High-resolution mapping of gene expression using association in an outbred mouse stock. PLoS Genet. 2008;4(8):e1000149.

10. Svenson KL, Gatti DM, Valdar W, Welsh CE, Cheng R, Chesler EJ, et al. High-resolution genetic mapping using the Mouse Diversity outbred population. Genetics. 2012;190(2):437–447.
11. Yalcin B, Willis-Owen SA, Fullerton J, Meesaq A, Deacon RM, Rawlins JNP, et al. Genetic dissection of a behavioral quantitative trait locus shows that Rgs2 modulates anxiety in mice. Nat Genet. 2004;36(11):1197.

12. Nicod J, Davies RW, Cai N, Hassett C, Goodstadt L, Cosgrove C, et al. Genome-wide association of multiple complex traits in outbred mice by ultra-low-coverage sequencing. Nat Genet. 2016;48(8):912.

13. Carbonetto P, Cheng R, Gyekis JP, Parker CC, Blizard DA, Palmer AA, et al. Discovery and refinement of muscle weight QTLs in B6\times D2 advanced intercross mice. Physiol Genomics. 2014;46(16):571–582.

14. Coyner J, McGuire JL, Parker CC, Ursano RJ, Palmer AA, Johnson LR. Mice selectively bred for High and Low fear behavior show differences in the number of pMAPK (p44/42 ERK) expressing neurons in lateral amygdala following Pavlovian fear conditioning. Neurobiol Learn Mem. 2014;112:195–203.

15. Parker CC, Cheng R, Sokoloff G, Palmer AA. Genome-wide association for methamphetamine sensitivity in an advanced intercross mouse line. Genes Brain Behav. 2012;11(1):52–61.

16. Samocha KE, Lim JE, Cheng R, Sokoloff G, Palmer AA. Fine mapping of QTL for prepulse inhibition in LG/J and SM/J mice using F2 and advanced intercross lines. Genes Brain Behav. 2010;9(7):759–767.

17. Hernandez Cordero Al, Carbonetto P, Riboni Verri G, Gregory JS, Vandenbergh DJ, P Gyekis J, et al. Replication and discovery of musculoskeletal QTLs in LG/J and SM/J advanced intercross lines. Physiol Rep. 2018;6(4).

18. Baud A, Guryev V, Hummel O, Johannesson M, Hermsen R, Stridh P, et al. Genomes and phenomes of a population of outbred rats and its progenitors. Sci Data. 2014;1:140011.

19. Besnier F, Wahlberg P, Rönnegard L, Ek W, Andersson L, Siegel PB, et al. Fine mapping and replication of QTL in outbred chicken advanced intercross lines. Genet Sel Evol. 2011;43(1):3.

20. Johansson M, Henriksen R, Höglund A, Fogelholm J, Jensen P, Wright D. Genetical genomics of growth in a chicken model. BMC Genomics. 2018;19(1):72.

21. Guryev V, Koudijs MJ, Berezikov E, Johnson SL, Plasterk RH, Van Eeden FJ, et al. Genetic variation in the zebrafish. Genome Res. 2006;16(4):491–497.

22. Patowary A, Purkanti R, Singh M, Chauhan R, Singh AR, Swarnkar M, et al. A sequence-based variation map of zebrafish. Zebrafish. 2013;10(1):15–20.

23. Mackay TF, Richards S, Stone EA, Barbadilla A, Ayroles JF, Zhu D, et al. The Drosophila melanogaster genetic reference panel. Nature. 2012;482(7384):173.

24. Vonesch SC, Lamparter D, Mackay TF, Bergmann S, Hafen E. Genome-wide analysis reveals novel regulators of growth in Drosophila melanogaster. PLoS Genet. 2016;12(1):e1005616.
25. King EG, Macdonald SJ, Long AD. Properties and power of the Drosophila Synthetic Population Resource for the routine dissection of complex traits. Genetics. 2012;genetics–112.

26. Kislukhin G, King EG, Walters KN, Macdonald SJ, Long AD. The genetic architecture of methotrexate toxicity is similar in Drosophila melanogaster and humans. G3 Genes Genomes Genet. 2013;g3–113.

27. Marriage TN, King EG, Long AD, Macdonald SJ. Fine-mapping nicotine resistance loci in Drosophila using a multiparent advanced generation inter-cross population. Genetics. 2014;198(1):45–57.

28. Doitsidou M, Jarriault S, Poole RJ. Next-generation sequencing-based approaches for mutation mapping and identification in Caenorhabditis elegans. Genetics. 2016;204(2):451–474.

29. Diouf IA, Derivot L, Bitton F, Pascual L, Causse M. Water Deficit and Salinity Stress Reveal Many Specific QTL for Plant Growth and Fruit Quality Traits in Tomato. Front Plant Sci. 2018;9:279.

30. Cockram J, Mackay I. Genetic Mapping Populations for Conducting High-Resolution Trait Mapping in Plants. 2018;

31. Rishmawi L, Bühler J, Jaegle B, Hülskamp M, Koornneef M. Quantitative trait loci controlling leaf venation in Arabidopsis. Plant Cell Environ. 2017;40(8):1429–1441.

32. Parker CC, Palmer AA. Dark matter: are mice the solution to missing heritability? Front Genet. 2011;2:32.

33. Darvasi A, Soller M. Advanced intercross lines, an experimental population for fine genetic mapping. Genetics. 1995;141(3):1199–1207.

34. Gonzales NM, Palmer AA. Fine-mapping QTLs in advanced intercross lines and other outbred populations. Mamm Genome. 2014;25(7–8):271–292.

35. Cheng R, Lim JE, Samocha KE, Sokoloff G, Abney M, Skol AD, et al. Genome-wide association studies and the problem of relatedness among advanced intercross lines and other highly recombinant populations. Genetics. 2010;

36. Parker CC, Carbonetto P, Sokoloff G, Park YJ, Abney M, Palmer AA. High-resolution genetic mapping of complex traits from a combined analysis of F2 and advanced intercross mice. Genetics. 2014;198(1):103–116.

37. Lionikas A, Cheng R, Lim JE, Palmer AA, Blizard DA. Fine-mapping of muscle weight QTL in LG/J and SM/J intercrosses. Physiol Genomics. 2010;42(1):33–38.

38. Carroll AM, Cheng R, Collie-Duguid ESR, Meharg C, Scholz ME, Fiering S, et al. Fine-mapping of genes determining extrafusal fiber properties in murine soleus muscle. Physiol Genomics. 2017;49(3):141–150.

39. Gonzales NM, Seo J, Hernandez-Cordero AI, Pierre CLS, Gregory JS, Distler MG, et al. Genome wide association study of behavioral, physiological and gene expression traits in a multigenerational mouse intercross. bioRxiv. 2017 Dec 8;230920.
40. Elshire RJ, Glaubitz JC, Sun Q, Poland JA, Kawamoto K, Buckler ES, et al. A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. PloS One. 2011;6(5):e19379.

41. Davey JW, Hohenlohe PA, Etter PD, Boone JQ, Catchen JM, Blaxter ML. Genome-wide genetic marker discovery and genotyping using next-generation sequencing. Nat Rev Genet. 2011;12(7):499.

42. Fitzpatrick CJ, Gopalakrishnan S, Cogan ES, Yager LM, Meyer PJ, Lovic V, et al. Variation in the form of Pavlovian conditioned approach behavior among outbred male Sprague-Dawley rats from different vendors and colonies: sign-tracking vs. goal-tracking. PloS One. 2013;8(10):e75042.

43. Toker L, Feng M, Pavlidis P. Whose sample is it anyway? Widespread misannotation of samples in transcriptomics studies. F1000Research (Internet). 2016 Sep 30 (cited 2018 Jul 11);5. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5034794/

44. Zhou X, Stephens M. Genome-wide efficient mixed-model analysis for association studies. Nat Genet. 2012;44(7):821.

45. Listgarten J, Lippert C, Kadie CM, Davidson RI, Eskin E, Heckerman D. Improved linear mixed models for genome-wide association studies. Nat Methods. 2012;9(6):525.

46. Cheng R, Parker CC, Abney M, Palmer AA. Practical considerations regarding the use of genotype and pedigree data to model relatedness in the context of genome-wide association studies. G3 Genes Genomes Genet. 2013;g3–113.

47. Yang J, Zaitlen NA, Goddard ME, Visscher PM, Price AL. Advantages and pitfalls in the application of mixed-model association methods. Nat Genet. 2014;46(2):100.

48. Joo JWJ, Hormozdiari F, Han B, Eskin E. Multiple testing correction in linear mixed models. Genome Biol. 2016;17(1):62.

49. Han B, Kang HM, Eskin E. Rapid and accurate multiple testing correction and power estimation for millions of correlated markers. PLoS Genet. 2009;5(4):e1000456.

50. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 2007;81(3):559–575.

51. Nikolskiy I, Conrad DF, Chun S, Fay JC, Cheverud JM, Lawson HA. Using whole-genome sequences of the LG/J and SM/J inbred mouse strains to prioritize quantitative trait genes and nucleotides. BMC Genomics. 2015;16(1):415.

52. Browning SR, Browning BL. Rapid and accurate haplotype phasing and missing-data inference for whole-genome association studies by use of localized haplotype clustering. Am J Hum Genet. 2007;81(5):1084–1097.

53. Yang J, Zeng J, Goddard ME, Wray NR, Visscher PM. Concepts, estimation and interpretation of SNP-based heritability. Nat Genet. 2017 Sep;49(9):1304–10.
54. Parker CC, Cheng R, Sokoloff G, Lim JE, Skol AD, Abney M, et al. Fine-mapping alleles for body weight in LG/J × SM/J F<Subscript>2</Subscript> and F<Subscript>34</Subscript> advanced intercross lines. Mamm Genome. 2011 Oct 1;22(9–10):563.

55. Bartnikas TB, Parker CC, Cheng R, Campagna DR, Lim JE, Palmer AA, et al. QTLs for murine red blood cell parameters in LG/J and SM/J F2 and advanced intercross lines. Mamm Genome. 2012 Jun;23(5–6):356–66.

56. Cheng R, Abney M, Palmer AA, Skol AD. QTLRel: an R package for genome-wide association studies in which relatedness is a concern. BMC Genet. 2011 Jul 27;12:66.

57. Logan RW, Robledo RF, Recla JM, Philip VM, Bubier JA, Jay JJ, et al. High-precision genetic mapping of behavioral traits in the diversity outbred mouse population. Genes Brain Behav. 2013;12(4):424–437.

58. Cervino AC, Li G, Edwards S, Zhu J, Laurie C, Tokiwa G, et al. Integrating QTL and high-density SNP analyses in mice to identify Insig2 as a susceptibility gene for plasma cholesterol levels. Genomics. 2005;86(5):505–517.

59. Conneally PM, Edwards JH, Kidd KK, Lalouel J-M, Morton NE, Ott J, et al. Report of the committee on methods of linkage analysis and reporting. Cytogenet Genome Res. 1985;40(1–4):356–9.

60. Lander ES, Botstein D. Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. Genetics. 1989;121(1):373–386.

61. Mangin B, Goffinet B, Rebai A. Constructing confidence intervals for QTL location. Genetics. 1994;138(4):1301–1308.

62. Dupuis J, Siegmund D. Statistical methods for mapping quantitative trait loci from a dense set of markers. Genetics. 1999;151(1):373–386.

63. Manichaikul A, Dupuis J, Sen S, Broman KW. Poor performance of bootstrap confidence intervals for the location of a quantitative trait locus. Genetics. 2006;

64. Sequencing RG, Baud A, Hermesen R, Guryev V, Stridh P, Graham D, et al. Combined sequence-based and genetic mapping analysis of complex traits in outbred rats. Nat Genet. 2013;45(7):767.

65. Pruim RJ, Welch RP, Sanna S, Teslovič TM, Chines PS, Gliedt TP, et al. LocusZoom: regional visualization of genome-wide association scan results. Bioinformatics. 2010;26(18):2336–2337.

66. Gardenghi S, Marongiu MF, Ramos P, Guy E, Breda L, Chadburn A, et al. Ineffective erythropoiesis in β-thalassemia is characterized by increased iron absorption mediated by down-regulation of hepcidin and up-regulation of ferroportin. Blood. 2007;109(11):5027–5035.

67. Graziano JH, Grady RW, Cerami A. The identification of 2, 3-dihydroxybenzoic acid as a potentially useful iron-chelating drug. J Pharmacol Exp Ther. 1974;190(3):570–575.

68. Korneliussen TS, Albrechtsen A, Nielsen R. ANGSD: analysis of next generation sequencing data. BMC Bioinformatics. 2014;15(1):356.
Main figure legends

Fig 1. Minor allele frequency distributions for F$_{34}$ array, F$_{34}$ GBS, F$_{39}$-F$_{43}$ GBS, and F$_{34}$ and F$_{39}$-F$_{43}$ GBS SNP sets. MAF distributions are highly comparable between AIL generations.

Fig 2. Chip-heritability estimates in F$_{34}$ and F$_{39}$-F$_{43}$ AILs. All heritability estimates are highly significant (p < 1.0E-05; see S7 Table).

Fig 3. Manhattan plots comparing F$_{34}$ GBS, F$_{39}$-F$_{43}$ GBS, and mega-analysis on locomotor day 1 test using 57,170 shared SNPs in all AIL generations. Mega-analysis identified a locus on chromosome 10 (chr10:104988207) that was not detected in the F$_{34}$ or F$_{39}$-F$_{43}$ alone, suggesting that mega-analysis enhanced power to detect some loci.

Fig 4. Significant loci on chromosome 17 for open field, distance traveled in periphery in F$_{34}$ AIL. (a) Genome-wide Manhattan plots highlighting the significant loci on chromosome 17 associated with open field, distance traveled in periphery in F$_{34}$ array SNP set (chr17:27427358) and F$_{34}$ GBS SNP set (chr17:27130383). (b) As exemplified in this pair of LocusZoom plots, GBS SNPs defined the boundaries of the loci much more precisely than array SNPs. GBS SNPs that are in high LD ($r^2 > 0.8$, red dots) with lead SNP chr17:27130383 resides between 27.28 - 28.2 Mb. In contrast, too few SNPs are present in the array plot to draw any definitive conclusion about the boundaries or LD pattern in this region. This locus has also been observed in GWAS for a variety of other traits, including locomotor activity, startle and prepulse inhibition [39] and open field test for anxiety [36]. LocusZoom plots for all loci identified in this paper are in Fig S7.

Supporting information
S1 Fig. Autosomal heterozygosity distribution in F_{34}, F_{39-43} AILs. Animals with excessive or insufficient heterozygosity (3 s.d. away from mean) were removed from further analysis. As controls, we have sequenced two F_2 s of LG and SM, four LG mice and four SM mice (see annotated data points with 1 and 0 heterozygosity).

S2 Fig. Kinship coefficients in F_{34} and F_{39-43} AILs calculated from pedigree against genetic relatedness matrix calculated using IBDLD [49]. Each circle represents a pair of animals, which their genetic kinship relatedness on the x-axis and pedigree kinship relatedness on the y-axis. Color signifies relatedness based on AIL pedigree. Blue circles represent identical twins, red full siblings, yellow parent-offspring pairs, grey other relationships. Seven animal pairs that deviate from the pedigree relationship clusters were excluded (see black arrows).

S3 Fig. Heatmap showing F_{34} array and F_{34} GBS genotype concordance in percentages, using 66 shared SNPs. “A” codes for the LG/J allele, and “B” codes for the SM/J allele. “AA” genotype concordance between array and GBS is 24.54%, “AB” 43.23%, “BB” 27.60%.

S4 Fig. LD decay in F_{34} array, F_{34} GBS, F_{39-43} GBS, and F_{34} and F_{39-43} GBS SNP sets.

S5 Fig. Manhattan plots comparing 4,593 F_{34} array, 60.3K F_{34} GBS, 4.3M imputed F_{34} array, and 4.1M imputed F_{34} GBS (N=428) SNPs on day 2 locomotor activity. Adjusted significance thresholds for imputed array and GBS SNPs were estimated using LD pruned SNPs (r^2=0.1, window size=20kb; PLINK v1.9). Notice that even though the imputed sets have more SNPs (the two right panels), they are frequently blocks of many SNPs with almost identical position and LD=1, therefore making it hard to visualize the additional SNPs.

S6 Fig. SNP heritability using F_{34} GBS and F_{34} array SNPs (slope=1).

S7 Fig. LocusZoom for F_{34} array, F_{34} GBS, F_{39-43} GBS, and mega-analysis QTLs.

S1 Table. List of phenotypes used in GWAS.

S2 Table. SNP and individual QC filter table. Numbers of animals and SNPs remained after each step of filtering are shown per GBS SNP set.

S3 Table. Effect of PLINK v1.9 clump-based pruning parameters on number of independent SNPs remained. At all r^2 values examined, a sliding window size of 12150kb was the first smallest window that yield the most stringent number of clumped SNPs in both array and GBS GWAS.

S4 Table. Lead QTL in F_{34} GBS and F_{34} array GWAS studies across phenotypes. Significant SNPs are clumped using parameters r^2=0.1, 12150kb.

S5 Table. F_{34} GBS and array SNP heritability estimates.

S6 Table. F_{34} and F_{39-43} genetic correlations in locomotor activity, coat color, and body weight.

S7 Table. SNP-heritability comparison between F_{34} and F_{39-43} GBS.
S8 Table. Lead QTL in F$_{39-43}$ N=600 GBS GWAS studies across phenotypes. Significant SNPs are clumped using parameters $r^2=0.1$, 12150kb.

S9 Table. Lead QTL in F$_{34}$ and F$_{39-43}$ (N=1028) mega-analysis across phenotypes. Significant SNPs are clumped using parameters $r^2=0.1$, 12150kb.

S10 Table. SNPs in F$_{34}$ GBS set with HWE p-values close to 1.0×10^{-6} cutoff threshold. These SNPs are removed from QTL summary tables.

S11 Table. SNPs in F$_{34}$ and F$_{39-43}$ mega-analysis GBS set with HWE p values close to 1.0×10^{-6} cutoff threshold. These SNPs are removed from QTL summary tables.

S12 Table. Adjusted significance threshold for each SNP set and GWAS cohort.
**a**  
F\text{34} array, open filed, distance traveled in periphery

**b**  
F\text{34} GBS, open filed, distance traveled in periphery

chr17:27427358

chr17:27130383

46 genes omitted