Mapping Liver Fat Female-dependent Quantitative Trait Loci in Collaborative Cross mice

Hanifa J. Abu-Toamih Atamni[1*], Maya Botzman[2*], Richard Mott[3], Irit Gat-Viks[2] and Fuad A. Iraqi[1]

[1]Sackler Faculty of Medicine, Tel-Aviv University, Israel.

[2]Faculty of Life Sciences, Tel-Aviv University, Israel.

[3]University of Oxford, Oxford, UK

Hanifa J. Abu-Toamih Atamni: hanifajalal@hotmail.com

Maya Botzman: mayabotzman@gmail.com

Richard Mott: r.mott@ucl.ac.uk

Irit Gat-Viks: iritgv@post.tau.ac.il

Fuad A. Iraqi: fuadi@post.tau.ac.il

***have equal contribution**

Corresponding author:

Prof. Fuad A. Iraqi
Department of Clinical Microbiology and Immunology
Sackler Faculty of Medicine
Tel Aviv University, Ramat Aviv
Tel Aviv 69978
Israel
Email: fuadi@post.tau.ac.il

Running title: Mapping female associated QTL for Liver Fat in CC mice

Abstract

Background & Aims: Non-alcoholic fatty liver disease (NAFLD) is the most common cause of chronic liver disease in the western world, with spectrum from simple steatosis to non-alcoholic steatohepatitis, which can progress to cirrhosis. NAFLD developments known to be affected by host genetic background. Herein, we emphasize the power of Collaborative Cross (CC) mouse for dissecting this complex trait, and revealing Quantitative Trait Loci (QTL) controlling hepatic fat accumulation in mice.

Methods: 168 female and 338 male mice from 24 and 37 CC lines, respectively age of 18 to 20 weeks old, maintained on standard rodent diet, since weaning. Hepatic fat content was assessed, using dual DEXA scan in the liver. Using the available high-density genotype markers of the CC line, QTL mapping associated with percentage liver fat accumulation was performed.

Results: Our results revealed significant fatty liver accumulation QTL that were specifically, mapped in females. Two significant QTLs on chromosomes 17 and 18, with genomic intervals 3Mb and 2Mb, respectively, were mapped. A third QTL, with a less significant *P* value, was mapped to chromosome 4, with genomic interval of 2Mb. These QTLs were named *Flal*1-*Flal*3, referring to *Fatty Liver Accumulation Locus* 1-3, for the QTLs on chromosomes 17, 18 and 4, respectively. Unfortunately, no QTL were mapped with males. Searching the mouse genome database suggested several candidate genes involved in hepatic fat accumulation.

Conclusions: Our results show that susceptibility to hepatic fat accumulations is a complex trait, controlled by multiple genetic factors in female mice, but not in male.

Five Keywords: NAFLD, High genetic diverse mouse population, Standard rodent diet, QTL mapping, Candidate genes.

Introduction

Non-alcoholic fatty liver disease (NAFLD) is a spectrum of liver disorders associated with hepatic steatosis that is not due to significant alcohol consumption or other secondary causes (Masuoka and Chalasani 2013). This disorder encompasses a wide range of diseases, from simple steatosis, which is relatively benign, to hepatic inflammation, hepatocyte injury, and fibrosis, a syndrome referred to as non-alcoholic steatohepatitis (NASH) that can progress to cirrhosis (Masuoka and Chalasani 2013; Adams et al. 2005; Ekstedt et al. 2006; Ratziu and Poynard 2006). NAFLD has been diagnosed by the use of several tools including determination of levels of liver enzymes; alanine aminotransferase (ALT) and aspartate aminotransferase (AST) as non-invasive indicators of NAFLD; as well as histology and imaging (CT, MRI, US). Although elevated ALT is generally associated with histological NASH, some of those patients with normal ALT levels may also have NAFLD and even advanced fibrosis. Therefore, ALT activity alone cannot be used to rule out significant liver disease in patients suspected of having NAFLD, especially in those with hepatomegaly or Type 2 Diabetes (T2D), a metabolic disorder that is associated with elevated levels of blood glucose and insulin resistance (Vernon et al. 2011).

In developed countries such as the United States, NAFLD has become the most common cause of chronic liver disease. The rate of NAFLD is increasing likely due to the rising prevalence of associated conditions in adults as well as children such as obesity and T2D, which together with dyslipidemia are the most important risk factors for NAFLD (Caldwell et al. 1999; Vardi et al. 2007). It has been projected that, within the next two decades, NASH will become the predominant cause of cirrhosis requiring orthotopic liver transplantation (Charlton 2004; Charlton et al. 2011).

While NAFLD is attributable to over-nutrition and sedentary life style (Cusi 2012; Larter et al. 2010), not all over-weight/obese people develop NAFLD and additional factor of individual susceptibility is required. Family studies, comparisons of NAFLD frequency between ethnic groups, and genome-wide association studies (GWAS) indicate that genetic predisposition underlines such individual susceptibility to NAFLD (Loomba et al. 2012; Romeo et al. 2008; Valenti et al. 2010), and also its severity (Chalasani et al. 2010). Ethnicity has a significant impact on the prevalence of NAFLD. In the Dallas Heart Study, the prevalence of hepatic steatosis was 45% in

Hispanics, 33% in non-Hispanic Caucasians, and 24% in African Americans (Browning et al. 2004).

It remains a puzzle why some individuals with NAFLD have advanced histological features and develop cirrhosis whereas others with comparable risk factor profile have simple steatosis with minimal or no disease progression. A genetic basis for inter-individual phenotypic variability is strongly speculated, but genetic studies are very limited to small number of individuals with histologically characterized NASH (Targher et al. 2006).

A huge drawback in traditional linkage analysis is its low mapping resolution that rarely leads to gene discovery. A novel and promising mouse genetic reference population for high resolution mapping and subsequently identifying genes underlying the QTL is Collaborative Cross (CC) mouse genetic reference population (GRP).

The CC population is created by a community (Churchill et al. 2004) effort of the complex trait consortium (CTC, www.complextrait.org). This unique genetic resource will eventually comprise a set of approximately 300-400 RIL that will be created by full reciprocal 8-way matings of 8 divergent strains of mice: A/J, C57BL/6J, 129S1/SvImJ, NOD/LtJ, NZO/HiLtJ, CAST/Ei, PWK/PhJ, and WSB/EiJ. Controlled randomization and minimization of selection during the breeding process will recombine the natural genetic variation present in these inbred strains. The result will be a unique collection of RIL exhibiting a large phenotypic and genetic diversity, and bringing the tremendous genetic variation potential of the mouse inbred lines to phenotypic expression (Churchill et al. 2004; Keane et al. 2011). Full details of CC lines status and their power of mapping QTL with host susceptibility to complex traits are presented (Iraqi et al. 2008; Durrant et al. 2011; Aylor et al. 2011; Philip et al. 2011; Mott et al. 2000; Iraqi et al. 2012; Welsh et al. 2012). Here, we show significant achievements of using of this unique mouse reference population to identify QTL and suggest candidate genes underlying these mapped QTL, which are associated with host susceptibility to NAFLD by assessing percentage of liver fat accumulation phenotype.

Materials and Methods

Collaborative cross mouse lines

In this study, a total of 168 female mice from different 24 lines (in average 7 mice per line) and and 338 males from 37 lines (in average 9.1 mice per line) of new developed CC mouse population were studied. The mice were developed and maintained at the Small Animal Facility at Sackler Faculty of Medicine, Tel Aviv University (TAU). The CC lines were at inbreeding generations F_{10}-F_{18}, minimum 90% homozygosity by extensive high-density genotyping. Full details of the development of these CC lines are given in previous reports (Iraqi et al. 2008; Iraqi et al. 2012; Welsh et al. 2012). All experimental mice and protocols were approved by the Institutional Animal Care and Use Committee (IACUC) of TAU with numbers (M-10-073 and M-14-007), which adheres to Israeli guidelines and follows the NIH/USA animal care and use protocols. Mice were housed on hardwood chip bedding in open-top cages under 12 hours light/dark cycle at 21-23^{0}C, given tap water and standard rodent chow diet *ad libitum* since weaning day (3 weeks old) until the age of 20 weeks old.

Dual-energy X-ray absorptiometry (DEXA) scan

DEXA scans were performed using the Lunar PIXImus Densitometer (GE Medical Systems) at the I. Meier Segals Garden for Zoological research, Tel Aviv University. With an image area of 80 mm x 65 mm, the PIXImus scan image reveals precise data about body composition for lean/fat and bone tissues. Following the full body DEXA scans, mice were dissected and liver DEXA scans were performed to assess hepatic fattiness. PIXImus Densitometer was calibrated before each testing using a quality control phantom with known values (BMD=0.0609g/cm^2; and %Fat=11.8%) following the manufactures instructions.

Statistical Analysis

Data analysis was performed using a statistical software package SPSS version 22. One-way ANOVA was carried out for testing the significance of the difference between the recorded traits among the tested CC lines. Phenotype recorded data of the

different CC lines were analyzed by ANOVA and a *P* value of 0.05 or less was considered significant.

Estimation of Heritability and Genetic coefficient of variation

Heritability (H^2) was estimated as the proportion of phenotypic variation explained by differences between CC lines in the ANOVA, i.e. $H^2=V_g/(V_g+V_e)$, as detailed in (Iraqi et al. 2014). The heritability statistic estimates the proportion of observed phenotypic variation that is due to genetic factors. However, it does not tell us whether the absolute amount of genetic variation generated by these genetic factors (the "genetic component of variation") is great or small. In our previous study, we have discussed and addressed in full details and the estimation of the genetic coefficient variation for any given trait (Iraqi et al. 2014). Briefly, for our data, Genetic Coefficient of Variation (CV_G) was estimated as:

SD_G/Mean

Where,

SD_G= the broad-sense genetic standard deviation among CC lines = $V_G^{0.5}$

Mean = mean trait value across all CC lines.

CC lines genotyping data

High molecular genomic DNA of the CC lines were initially genotyped with the mouse diversity array (MDA), which consists of 620,000 SNPs (Yang et al. 2009). After about two 4-generation intervals of inbreeding, all the CC lines were re-genotyped by mouse universal genotype array (MUGA-7,500 markers) and finally with MegaMuga (77,800 markers) SNP arrays to confirm their genotype status (Iraqi et al. 2012). Although the CC lines are not yet completely inbred, we have shown previously that using genotypes from a single representative from each line is sufficient for QTL mapping purposes (Durrant et al. 2011; Aylor et al. 2011; Philip et al. 2011; Vered et al. 2014).

QTL analysis

The genome of each CC line is a mosaic of the inbred founders, which we reconstructed using a hidden Markov Model (HMM) HAPPY software (Mott et al. 2000) across the genotypes to compute probabilities of descent from the founders (setting the generation parameter to 7). In CC line *k* at SNP interval (locus) *L*, the HMM probability of descent from founder strain *s* is denoted by $P_{Lk(s)}$. The presence of a QTL at locus *L* is tested using a linear regression framework, in which the residual deviance from the mean probability of death Y_k for an individual from line *k* is:

$$\ln \frac{\pi(y_k)}{1-\pi(y_k)} = \mu + \sum_s P_{Lk}(s)\beta_s$$

Where the overall mean (incorporating any effects of batch and sex) is μ, and β_s is the effect of founder haplotype *s* at locus *L*.

The presence of a QTL was assessed through an ANOVA test by comparing the fit of the model with that of a simpler model in which the $\beta_s = 0$ (the null hypothesis). Age was used as co-factor. Significance is reported as the minus $\log_{10}$ *P* value , as computed by the anova function in R. Genome-wide significance was estimated by permutation, where the CC line labels were permuted between the phenotypes. The median probability of death across replicates within each CC line was used in the QTL analysis. QTL effect sizes were estimated as the proportion of the log likelihood explained by the locus effects at the QTL.

The confidence interval of the QTL was defined based on permutations, using similar approach as presented in our previous studies (Durrant et al. 2011; Vered et al. 2014) to take into account local patters of linkage disequilibrium. Permutation based FDR was calculated as follows: for a given *P* value threshold, the peaks above this threshold in the permutated datasets are considered as 'false positives' (FPs), and the peaks above this threshold in the real dataset are considered as 'positives' (P). Using these values, the permutation-based false discovery rate (FDR) was defined as FP/P.

Results

Liver Fat content by DEXA scan

We assessed percentage of liver fat using DEXA-scanner of females (Fig. 1A) and males (Fig. 1B) from 24 and 37 CC lines, respectively, after maintaining them on standard rodents' diet since weaning day at age of 3 weeks old until the age of 20 weeks old. The one-way ANOVA for females data shows that the 24 female CC lines differed significantly (*P* value<0.05) in their tendency to liver fat accumulation. Among the females of 24 CC lines the mean ratio of percentage of liver fat is 54.35 ±2.01. Percentage of liver fat ranged between maximum value of 100% fat in line IL1513, and minimum level of 20.26% in line IL1488.

The One Way ANOVA for males data shows that the 37 CC lines differed, significantly (P<0.05) as well, with their tendency to liver fat accumulation. Among the males of 37 CC inbred lines, the mean ratio of percentage of liver fat is 38.20 (±2.01). While the maximum level of percentage of liver fat was 72.26% Fat (IL4141) and the minimum level was 16.60% (IL2141), 2.68 times under the CC population mean.

Knowing that the CC lines have different genetic makeup, these results show the strong role of genetic background in determining the percentage of liver fat accumulation at standard environmental conditions. Notably, these results were obtained without western high fat dietary challenge, indicating that the original genetic predisposition indeed underlies the liver fat accumulation pattern.

Broad-sense heritability and Coefficients of genetic variation

As part of assessing the genetic effects underlying the observed inter-individual phenotypic variation, both the heritability (H^2) and the genetic coefficient of variation (CV_G) metrics were calculated (see Methods). We found that the heritability value is 0.32 and the genetic coefficient of variation is 0.28 for liver percentage fat among females from the 24 CC lines, while for the males from the 37 CC lines heritability value is 0.22 and the genetic coefficient of variation is 0.29.

QTL mapping and founder effect

We used the HAPPY software and QTL mapping method as previously described in a study of host susceptibility to Aspergillosis (Durrant et al. 2011) and Klebsiella (Vered et al. 2014) in CC lines, testing the percentage of liver fat as the quantitative trait.Initially, permutation tests for identifying and setting the genome wide significant thresholds were performed for the data. The threshold is presented in Figure 2 and is found to be minus log *P value*=5.5, with corresponding permutation-based FDR = 0.17 (Methods).

Data analysis for females' cohort has shown two significant QTLs on chromosomes 17 and 18. The QTL in Chr17 reached minus log *P* value of 5.618 at SNP JAX00431012, with genomic intervals between 8-11Mb (3 Mb). The QTL in Chr18 reached minus log *P* value of 5.58 at SNP JAX00462266, with genomic interval between 64-66Mb (2 Mb). A third putative QTL was just slightly under the significance threshold (minus log *P* value=5.41). This peak was located at SNP JAX00548268 and mapped to chromosome 4: 35-37Mb (2 Mb). These resolutions of mapping are therefore high considering that the mapping population consisted of only 24 CC lines. We designated the three QTLs with names *Flal*1- *Flal*3 referring to ***F****atty* ***L****iver* ***A****ccumulation* ***L****ocus* 1-3 for the QTLs on chromosomes 17, 18 and 4 (respectively).

Finally, we evaluated the effect of each founder haplotype on the percentage of females liver fat Estimation was conducted as deviation relative to the WSB/EiJ parental strain, which was arbitrarily assigned the baseline zero effect. Results of this analysis are presented for the two QTLs *Flal1* and *Flal2* (Fig. 3A and B, respectively).The two loci showed complex pattern of haplotype effects of the founders, with the wild-derived strains (mainly PWK) playing a role, although other strains also contributed to the overall QTL effect.

Conversely, data analysis for males cohort did not reveal any significant QTL (Figure 1 supplement), indicating that the differences of the male trait of liver fat accumulations under standard chow diet, is probably less associated with genetic background contribution.

Candidate genes underlying the mapped QTLs

Following the QTLs mapping, the mouse genome database (http://www.informatics.jax.org/) was queried for suggested candidate genes within each of the mapped QTLs.

The *Flal1* region harbors the chemokine (C-C motif) receptor 6 (*Ccr6*), a gene that was previously identified as a liver-specific receptor and an activation-regulated chemokine (LARC), namely a C chemokine that is mainly expressed in the liver. LARC with CCR6 receptor may play roles in inflammatory and immunological responses, and also in the normal lymphocyte trafficking and microenvironmental homing that are essential for development and maintenance of various lymphoid tissues (Baba et al. 1997).

The second gene is the T cell activation Rho GTPase activating protein (*Tagap*), which is a member of the Rho GTPase-activator protein superfamily. Tagap is involved in T-cell activation, where Tagap releases GTP from GTP-bound Rho and acts as molecular switch in cells. Therefore, polymorphisms in the Tagap gene have a key role in autoimmune diseases, including rheumatoid arthritis, Type 1 diabetes, celiac disease, and multiple sclerosis (Chen et al. 2011; Chatzikyriakidou et al. 2013; Eyre 2010)

Furthermore, several of previously mapped QTLs - associated with inflammation, obesity and diabetes - were also identified in the same interval. This includes the obesity QTL 19 (Obq19) (Ishimori 2004) and the insulin-dependent diabetes susceptibility 23 (Idd23) (Deruytter 2004).

In the case of *Flal2,* our analysis revealed several candidate genes, including *Nedd41* (the neural precursor cell expressed, developmentally down-regulated gene 4-like)- a gene involved in insulin regulation pathways, and Akt (protein kinase B) and Sgk (serum- and glucocorticoid-dependent kinase), two components of the insulin signaling pathway (Lee et al. 2007).

A number of previously mapped QTLs associated with inflammation and diseases related to NAFLD were also mapped to this region, including obesity 4 (Obsty4) (Cheverud et al. 2004) T helper cell response (Thcr) (Zhang et al. 2003) and insulin dependent diabetes susceptibility 21.2 (Idd21.2) (Hollis-Moffatt et al. 2005).

Discussion and Conclusions

In this study, we have used Collaborative Cross mice, both females and males, in order to identify QTLs and suggest candidate genes contributing to fat accumulation in the liver. The identification of common or sex-specific genetic resistant factors to this disease will help to understand the observed wide range of liver fat accumulation among individuals, and hopefully will suggest new prevention approaches as well.

Our findings confirm the validity and accuracy of the application of dual X-ray absorptiometry (DEXA) in a mouse model for studying liver fat accumulation tendency under naïve standard conditions. Furthermore, this is the first report to present results of liver fat accumulation in the newly developed CC mouse resource population.

Our cohort in this study consisted of 24 CC lines for females and 37 CC lines for males, imbalance due to nature selective and breeding differences among the CC lines. Despite a small number of CC lines (24 and 37 CC lines for females and males, respectively), data analyses showed significant differences between CC lines for both females and males, while only for females CC lines revealed QTLs that contribute to the liver tendency to accumulate less/more fat under naïve environmental conditions. Surprisingly, even with assessing of larger cohort male than female CC lines, our analysis could not identify QTL associated with the trait. This is strong evidence that the host genetic background has less effect on phenotypic variations of fat liver accumulations.

Nevertheless, the highly significant differences in trait values among lines and trait heritability indicate that line genotype is an important determinant of the differences in values of liver fat content. As shown in our results, heritability value for liver fat accumulations was higher in females (0.32) than males (0.22). These results enabled the mapping of the most significant female-dependent QTLs to a presidential high-resolution intervals, 3 and 2 Mb for Chromosomes 17 and 18 QTL, respectively, and identification of host genes that may control fat liver accumulation and inflammation pathways. These two significant QTLs were designated as *Flal*1- *Flal*2 referring to ***F**atty **L**iver **A**ccumulation **L**ocus* 1-2. The third QTL mapped on chromosome 4 distal region was named ***F**atty **L**iver **A**ccumulation **L**ocus* 3.

The high heritability (0.32) of the female tested trait has confirmed that this assessed phenotype is strongly controlled by genetic factors, which enabled us to map these QTLs with the limited number of CC lines. Furthermore, the genetic coefficient of variation of 0.28 has also confirmed the high genetic variations within the CC lines, which will allow identifying novel genetic variants that underlie the phenotype. It is believed that the high genetic variations in the CC lines was introduced from the three wild-derived strains, CAST/Ei (*M. m. castaneus)*, PWK/PhJ (*M. m. musculus)*, and WSB/EiJ (*M. m. domesticus)* (Churchill et al. 2004), but not segregating among classical strains descended from *M.m.domesticus* (most classical strains differ from the reference C57BL/6J at about 4 million SNPs, PWK and CAST each differ at about 17 million SNPs, and WSB at 6 million (Keane et al. 2011; Beck et al. 2000)) . Over 35 million SNPs segregate between the CC founders (Keane et al. 2011). Wild mice are constantly under environmental influence, and efficiency of their innate mechanisms of defense is under strong selective pressure.

Liver is the central metabolic organ. In addition to its role in maintaining plasma glucose, the liver also processes, synthesizes, and secretes lipids, namely triglyceride (TG) and cholesterol. Under pathological conditions (e.g. obesity), the liver is also responsible for storing of excess lipids giving rise to NAFLD. Lipid processing in the liver and NAFLD play important roles in metabolic syndrome diseases and diabetes. We suggest through forward genetic screening in the mouse several novel genes or alleles that are associated with those processes. Furthermore, we found several previously mapped QTLs within our *Flal* QTLs interval that are involved in body weight, growth and susceptibility to diabetes. These results are in agreement with the well established relationships between fatty liver accumulation, obesity, and a variety of inflammatory diseases.

This report and others (Durrant et al. 2011; Aylor et al. 2011; Philip et al. 2011; Vered et al. 2014) demonstrate the utility of CC lines in the analysis of complex traits in mouse models of human disease. Our results underline the importance of the contribution of wild-derived alleles to the CC, where *Flal*1 and *Flal*2 are mainly driven by wild strain. The wild-derived founders may have different immune response mechanisms compared with the classical mouse strains. If so, we expect to identify novel response mechanisms to the variation of fat liver accumulation. This has two

consequences. First, we can identify more genes using populations in which these variants segregate than in classical populations. Second, only sequence variants segregating in the CC founders that follow the same strain distribution pattern can be causal for the QTL. Therefore having a complete catalogue of sequence variants in the founders is of great utility. The combination of high-density genotypes in the CC and the genome sequences of the CC founders yield an approximate reconstruction of the sequence of each CC line as a mosaic of fragments of the founders' genomes. While this is currently limited to those regions of the genome that can be assembled from short-read sequence data, as sequencing technologies improve, we expect to generate a complete catalogue of variation in the CC model.

Overall, dissection of the complex genetics underling liver fat accumulation ability was the goal of our study. Our results support our principal idea of mapping and identify QTLs involved in liver fat accumulation pathways by using small number of CC lines. These results illustrates the importance of a mouse model sources such as CC population for identifying genetic factors affecting host susceptibility or resistibility towards future development of NAFLD disease under high fat dietary environments. Once the genetic basis of liver fat accumulation pathways is understood, such information may be of preventive and therapeutic value for the NAFLD disease development. In future, we will study the gene expression variations using RNAseq approach in different CC mouse lines, which differently express the trait of % fat in the liver, and subsequently identify these genes. Expression levels differences will be co-localized with phenotypic QTL given in our study enabling the tracking of positional candidate genes with higher resolution. We will combining the current QTL findings with the RNAseq results for better identifying candidate genes, which later on can be assessed and confirmed its association with the disease using specific gene knockout approach. As far, we know the reported QTLs in our study are new and were not linked, previously reported in any GWAS studies to liver diseases.

In the very near future, over a hundred inbred and genotyped CC lines will be available to the research community. Using more lines will drastically improve the resolution of mapping and statistical power. Nonetheless, the current study shows that even a modest number of lines are useful, if there is sufficient replication within each line.

Competing interests

The authors declare no potential conflict of interest with respect to financial or Non-financial competing interests, the authorship and/or publication of this article e.g. pharmaceutical stock ownership, consultancy.

Authors' contribution

HJA participated in the design of the study, carried out the mice assessment and participated in data analysis. MB participated in the data analysis and edits the first draft through the final submitted version. IGV participated in the data analysis and edits the first draft through the final submitted version. FI participated in designing the study and in developing the first draft through the final submitted version of the manuscript. All authors read and approved the final manuscript.

Acknowledgments

The present work is part of a Ph.D. thesis by Hanifa J. Atamni theses. The authors declare no competing financial interests or other associations that might pose a conflict of interest (e.g., pharmaceutical stock ownership, consultancy). This work was supported by the Hendrech and EiranGotwert Fund for studying diabetes, Wellcome Trust grants 085906/Z/08/Z, 075491/Z/04, Wellcome Trust core funding grant 090532/Z/09/Z, and core funding by Tel-Aviv University. This work was supported by the Israeli Centers of Research Excellence (I-CORE): Center No. 41/11 (MB, I.G-V) and a fellowship from the Edmond J. Safra Center for Bioinformatics at Tel-Aviv University (MB).

References

Adams LA, Lymp JF, St Sauver J, Sanderson SO, Lindor KD, Feldstein A, et al (2005) The natural history of nonalcoholic fatty liver disease: a population-based cohort study. Gastroenterology; 129:113-121.

Aylor DL, Valdar W, Foulds-Mathes W, Buus RJ, Verdugo RA, Baric RS, et al (2011) Genetic analysis of complex traits in the emerging Collaborative Cross. Genome Res; 21:1213-1222.

Baba M, Imai T, Nishimura M, Kakizaki M, Takagi S, Hieshima K, et al (1997) Identification of CCR6, the specific receptor for a novel lymphocyte-directed CC chemokine LARC. J Biol Chem ;272:14893-14898.

Beck JA, Lloyd S, Hafezparast M, Lennon-Pierce M, Eppig JT, Festing MF, et al (2000) Genealogies of mouse inbred strains. Nat Genet; 24:23-25.

Browning JD, Szczepaniak LS, Dobbins R, Nuremberg P, Horton JD, Cohen JC, et al (2004) Prevalence of hepatic steatosis in an urban population in the United States: impact of ethnicity. Hepatology; 40:1387-1395.

Caldwell SH, Oelsner DH, Iezzoni JC, Hespenheide EE, Battle EH, Driscoll CJ (1999) Cryptogenic cirrhosis: clinical characterization and risk factors for underlying disease. Hepatology; 29:664-669.

Charlton M (2004) Nonalcoholic fatty liver disease: a review of current understanding and future impact. Clin Gastroenterol Hepatol ;2:1048-1058.

Charlton MR, Burns JM, Pedersen RA, Watt KD, Heimbach JK, Dierkhising RA (2011) Frequency and outcomes of liver transplantation for nonalcoholic steatohepatitis in the United States. Gastroenterology; 141:1249-1253.

Chalasani N, Guo X, Loomba R, Goodarzi MO, Haritunians T, Kwon S, et al (2010) Nonalcoholic Steatohepatitis Clinical Research Network. Genome-wide association study identifies variants associated with histologic features of nonalcoholic Fatty liver disease. Gastroenterology; 139:1567-1576.

Chatzikyriakidou A, Voulgari PV, Lambropoulos A, Georgiou I, Drosos AA (2013) Validation of the TAGAP rs212389 polymorphism in rheumatoid arthritis susceptibility. Joint Bone Spine;80: 543-544.

Chen R, Stahl EA, Kurreeman FA, Gregersen PK, Siminovitch KA, Worthington J, et al (2011) Fine mapping the TAGAP risk locus in rheumatoid arthritis. Genes Immun; 12:314-318.

Cheverud JM, Ehrich TH, Hrbek T, Kenney JP, Pletscher LS, Semenkovich CF (2004) Quantitative trait loci for obesity- and diabetes-related traits and their dietary responses to high-fat feeding in LGXSM recombinant inbred mouse strains. Diabetes; 53:3328-3336.

Churchill GA, Airey DC, Allayee H, Angel JM, Attie AD, Beatty J, et al (2004) The Collaborative Cross, a community resource for the genetic analysis of complex traits. Nat Genet; 36:1133-1137.

Cusi K (2012) Role of obesity and lipotoxicity in the development of nonalcoholic steatohepatitis: pathophysiology and clinical implications. Gastroenterology; 142:711-725.

Deruytter N, Boulard O, Garchon HJ (2004) Mapping non-class II H2-linked loci for type 1 diabetes in nonobese diabetic mice. Diabetes ;53:3323-3327.

Durrant C, Tayem H, Yalcin B, Cleak J, Goodstadt L, de Villena FP, et al (2011) Collaborative Cross mice and their power to map host susceptibility to Aspergillus fumigatus infection. Genome Res; 21:1239-1248.

Ekstedt M, Franzén LE, Mathiesen UL, Thorelius L, Holmqvist M, Bodemar G, et al (2006) Long-term follow-up of patients with NAFLD and elevated liver enzymes. Hepatology; 44:865-873.

Eyre S, Hinks A, Bowes J, Flynn E, Martin P, Wilson AG, et al (2010) Overlapping genetic susceptibility variants between three autoimmune disorders: rheumatoid arthritis, type 1 diabetes and coeliac disease. Arthritis Res Ther;12:R175.

Hollis-Moffatt JE, Hook SM, Merriman TR (2005) Colocalization of mouse autoimmune diabetes loci Idd21.1 and Idd21.2 with IDDM6 (human) and Iddm3 (rat). Diabetes; 54:2820-2825.

Iraqi FA, Churchill G, Mott R (2008) The Collaborative Cross, developing a resource for mammalian systems genetics: a status report of the Wellcome Trust cohort. Mamm Genome;19:379-381.

Iraqi FA, Athamni H, Dorman A, Salymah Y, Tomlinson I, Nashif A, et al (2014) Heritability and coefficient of genetic variation analyses of phenotypic traits provide strong basis for high-resolution QTL mapping in the Collaborative Cross mouse genetic reference population. Mamm Genome; 25:109-119.

Iraqi FA, Mahajne M, Salaymeh A, Sandovsky H, Tayem H, Vered K, et al (2012) The genome architecture of the Collaborative Cross mouse genetic reference population. Genetics; 190:389-401.

Ishimori N, Li R, Kelmenson PM, Korstanje R, Walsh KA, Churchill GA, et al (2004) Quantitative trait loci that determine plasma lipids and obesity in C57BL/6J and 129S1/SvImJ inbred mice. J Lipid Res; 45:1624-1632.

Keane TM, Goodstadt L, Danecek P, White MA, Wong K, Yalcin B, et al (2011) Mouse genomic variation and its effect on phenotypes and gene regulation. Nature;477:289-294.

Larter CZ, Chitturi S, Heydet D, Farrell GC (2010) A fresh look at NASH pathogenesis. Part 1: the metabolic movers. J Gastroenterol Hepatol;25:672-690.

Lee IH, Dinudom A, Sanchez-Perez A, Kumar S, Cook DI (2007) Akt mediates the effect of insulin on epithelial sodium channels by inhibiting Nedd4-2. J Biol Chem; 282:29866-29873.

Loomba R, Abraham M, Unalp A, Wilson L, Lavine J, Doo E, et al (2012) Nonalcoholic Steatohepatitis Clinical Research Network. Association between diabetes, family history of diabetes, and risk of nonalcoholic steatohepatitis and fibrosis. Hepatology; 56:943-951.

Masuoka HC, Chalasani N (2013) Nonalcoholic fatty liver disease: an emerging threat
to obese and diabetic individuals. Ann N Y Acad Sci; 1281:106-122.

Mott R, Talbot CJ, Turri MG, Collins AC, Flint J (2000) A method for fine mapping quantitative trait loci in outbred animal stocks. Proc Natl Acad Sci USA; 97:12649-12654.

Philip VM, Sokoloff G, Ackert-Bicknell CL, Striz M, Branstetter L, Beckmann MA, et al (2011) Genetic analysis in the Collaborative Cross breeding population. Genome Res; 21:1223-1238.

Ratziu V, Poynard T (2006) Assessing the outcome of nonalcoholic steatohepatitis? It's time to get serious. Hepatology; 44:802-805.

Romeo S, Kozlitina J, Xing C, Pertsemlidis A, Cox D, Pennacchio LA, et al (2008) Genetic variation in PNPLA3 confers susceptibility to nonalcoholic fatty liver disease. Nat Genet; 40:1461-1465.

Targher G, Bertolini L, Padovani R, Rodella S, Zoppini G, Zenari L, et al (2006) Relations between carotid artery wall thickness and liver histology in subjects with nonalcoholic fatty liver disease. Diabetes Care; 29:1325-1330.

Valenti L, Al-Serri A, Daly AK, Galmozzi E, Rametta R, Dongiovanni P, et al (2010) Homozygosity for the patatin-like phospholipase-3/adiponutrin I148M polymorphism influences liver fibrosis in patients with nonalcoholic fatty liver disease. Hepatology; 51:1209-1217.

Vardi P, Alkalai KS, Sprecher E, Koren I, Zadik Z, Sabbah M, et al (2007) Components of the metabolic syndrome (MTS), hyperinsulinemia, and insulin resistance in obese Israeli children and adolescents. Diabetes & Metabolic Syndrome: Clinical Research & Reviews; 1: 97-103.

Vered K, Durrant C, Mott R, Iraqi FA (2014) Susceptibility to Klebsiella pneumonaie infection in collaborative cross mice is a complex trait controlled by at least three loci acting at different time points. BMC Genomics; 15:865.

Vernon G, Baranova A, Younossi ZM (2011) Systematic review: the epidemiology and natural history of non-alcoholic fatty liver disease and non-alcoholic steatohepatitis in adults. Aliment Pharmacol Ther; 34:274-285.

Welsh CE, Miller DR, Manly KF, Wang J, McMillan L, Morahan G, et al (2012) Status and access to the Collaborative Cross population. Mamm Genome;23:706-712.

Yang H, Ding Y, Hutchins LN, Szatkiewicz J, Bell TA, Paigen BJ, et al (2009) A customized and versatile high-density genotyping array for the mouse. Nat Methods; 6:663-666.

Zhang F, Liang Z, Matsuki N, Van Kaer L, Joyce S, Wakeland EK, et al (2003) A murine locus on chromosome 18 controls NKT cell homeostasis and Th cell differentiation. J Immunol; 171:4613-4620.

Figure legends:

Fig. 1. Means of percentage of liver fat (±SEM) by DEXA scan analysis among different CC lines at age of 20 weeks old on standard rodents diet. (A) Means of percentage of liver fat (±SEM) for female CC lines. (B) Means of percentage of liver fat (±SEM) for Male CC lines. *x* axis – line number, representing the different CC lines. *y* axis – represents the percentage of liver fat values as monitored by DEXA. Significant variation was found between the different CC lines at $P<0.05$.

Fig. 2. Genome scans for significant QTLs associated with percentage of liver fat across female CC lines. Two significant QTLs were mapped to chromosomes 17 and 18 with female data. A less significant QTL was also mapped to chromosome 4. Experiment-wide thresholds of significance at 95% levels are minus log *P* value=5.5 with corresponding permutation-based false discovery rate (FDR) = 0.17 (horizontal line).

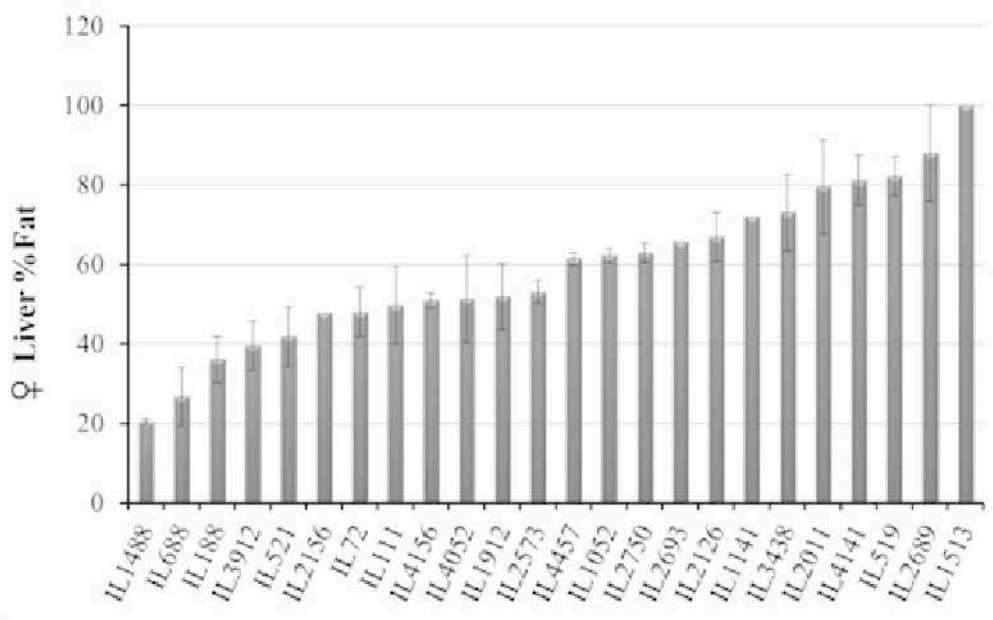
Fig. 3. Estimated haplotype effects at QTL for the percentage of liver fat on chromosomes 17 and 18. Plots A, B show the estimated haplotype effects on chromosome 17 and 18, respectively (*y* axis) across founder strains (*x* axis). Effects are shown as deviations relative to WSB/EiJ, which was arbitrarily assigned trait effect = 0.

Supplement figure 1. Genome scans for significant QTLs associated with percentage of liver fat across male CC lines. Experiment-wide thresholds of significance at 95% levels are minus log *P* value=5.5 with corresponding permutation-based false discovery rate (FDR) = 0.17 (horizontal line).

Supplement figure 2. Microscopic apperance and hematoxylin and Eosin staining of mouse livers at diferent stages with fat liver accumopations, at 6, 8, 12 and 20 weeks. Picture was taken from http://www.psychogenics.com/ .

Figure 1. Means of percentage of liver fat by DEXA scan analysis among different CC lines.

A



B

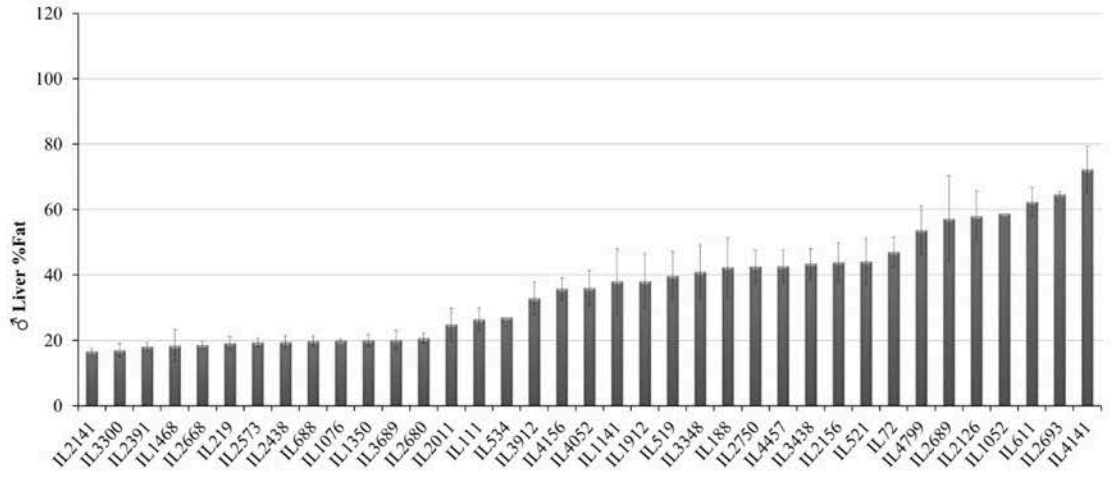


Figure 2. Genome scans for significant QTLs associated with percentage of liver fat across female CC lines

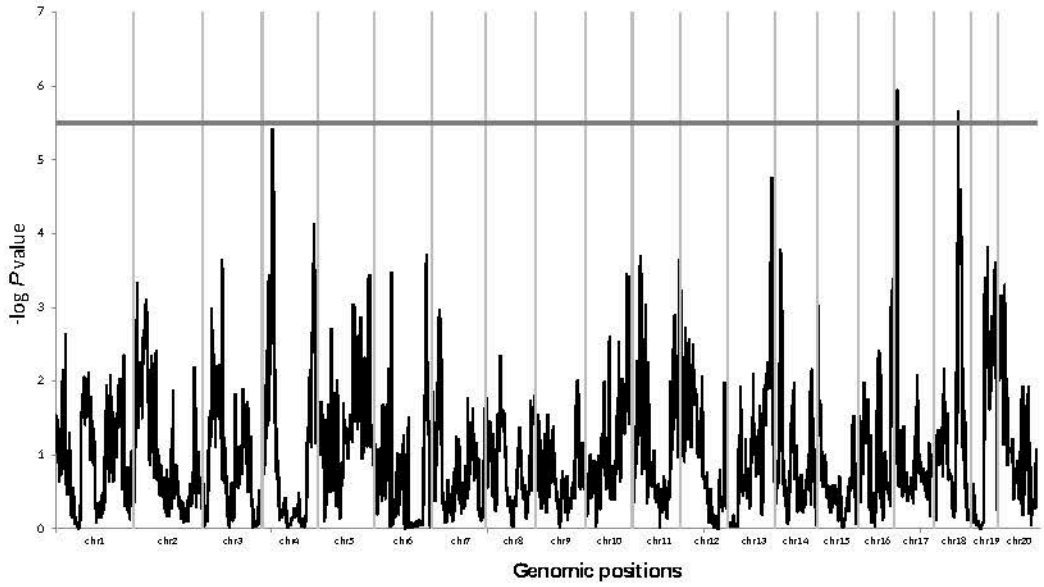
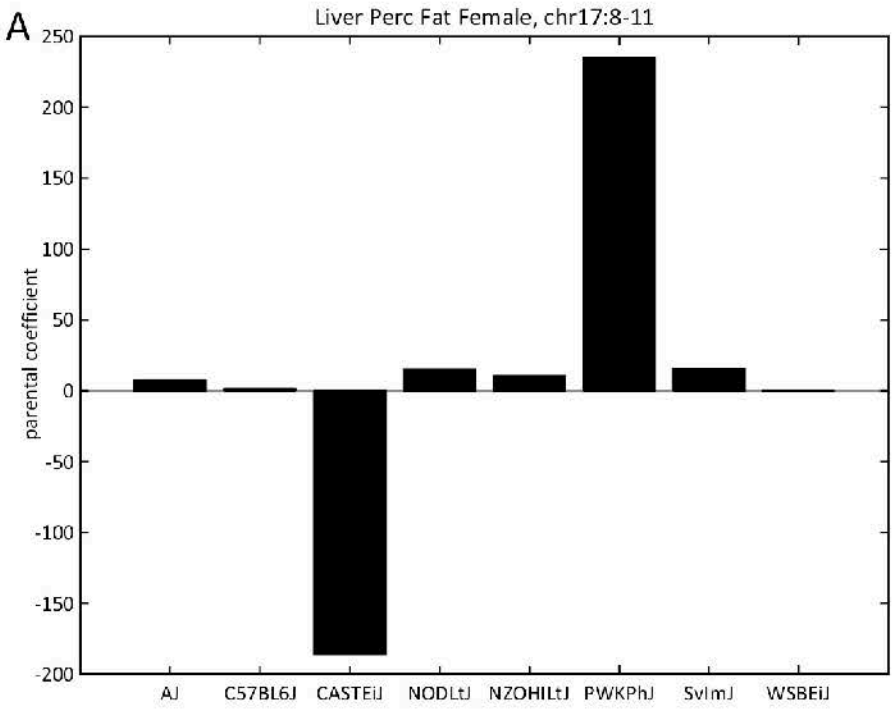
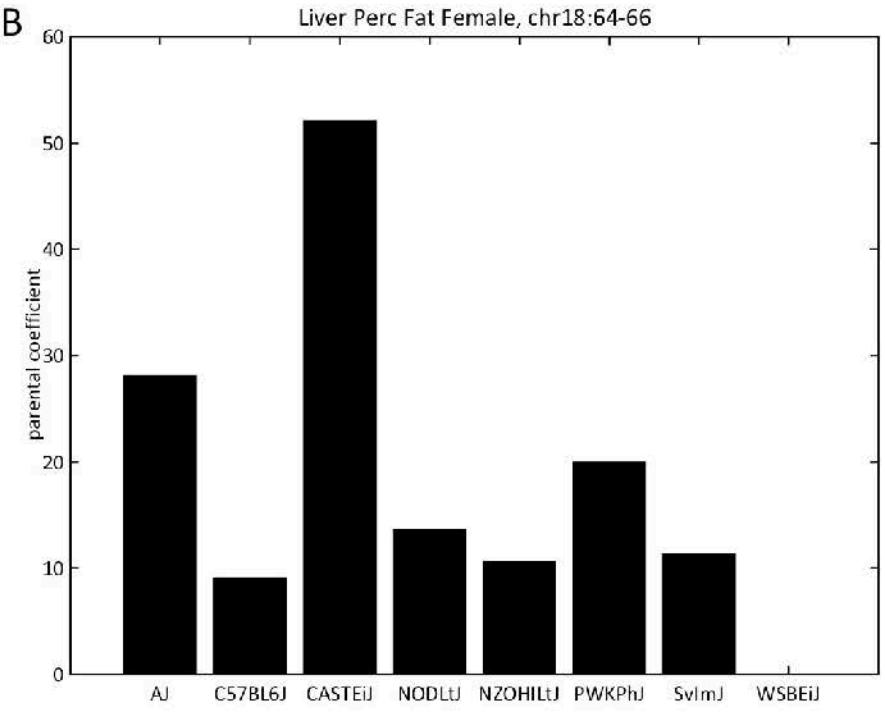


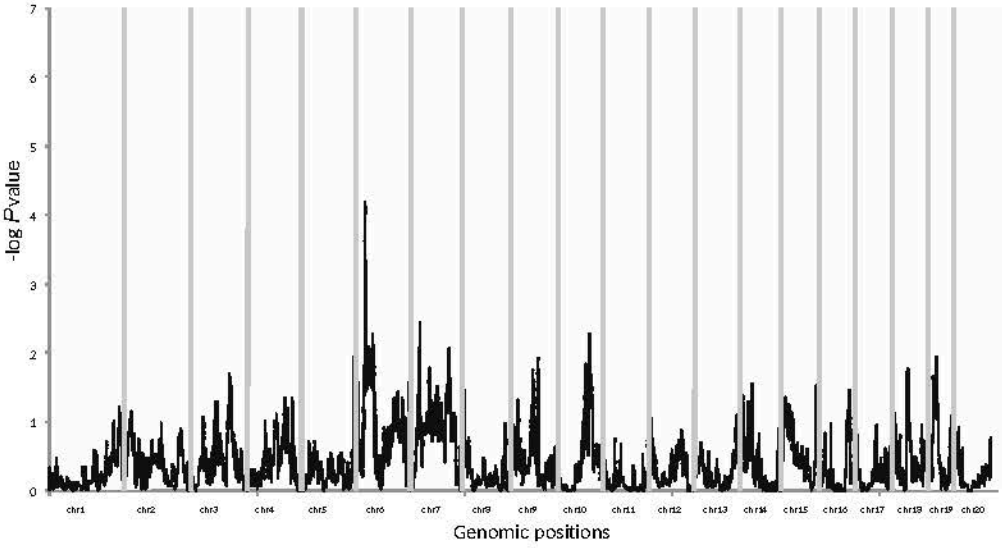
Figure 3. Estimated haplotype effects at QTL for the percentage of liver fat on chromosomes 17 (A) and chromosome 18 (B).



B


Liver Perc Fat Female, chr18:64-66
parental coefficient
60
50
40
30
20
10
0
AJ
C57BL6J
CASTEiJ
NODLtJ
NZOHILtJ
PWKPhJ
SvImJ
WSBEiJ

Supplement figure 1. Genome scans for significant QTLs associated with percentage of liver fat across male CC lines.



Supplement figure 2. Microscopic apperance and hematoxylin and Eosin staining of mouse livers at diferent stages with fat liver accumopations, at 6, 8, 12 and 20 weeks. Picture was taken from http://www.psychogenics.com/ .