Identification of Genetic Variants Regulating the Abundance of Clinically Relevant Plasma Proteins using the Diversity Outbred Mouse Model

Stéphanie Philtjens
Indiana University
Dominic J. Acri
Indiana University
Byungwook Kim
Indiana University
Hyewon Kim
Mayo Clinic
Jungsu Kim (jk123@iu.edu)
Indiana University

Research Article

Keywords: Diversity Outbred , protein quantitative trait locus , Giga Mouse Universal Genotyping Array , Olink , Mouse Exploratory Panel , plasma , genetic modifier , Interleukin 17A , Regulatory Factor X1 , Aryl Hydrocarbon Receptor

Posted Date: February 8th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-171899/v1

Abstract

Background: Levels of plasma proteins are under control of environmental and genetic factors. To use plasma proteins in biomarker studies, we need to understand how genetic modifiers influence their abundance. Although there has been expression quantitative trait loci (eQTL) studies on a few limited numbers of proteins, the effect of genetic variants on the levels of multiple plasma proteins still warrants more systematic investigation.

Results: To identify genetic modifiers that influence the levels of clinically relevant plasma proteins, we performed protein quantitative trait locus (pQTL) mapping on the 92 proteins present in the Olink Mouse Exploratory Panel using the Diversity Outbred (DO) mouse population. We identified 12 significant pQTL that were located in *cis* and 6 that were in *trans*. Among them, we discovered that the presence of coding variants in the gene encoding for the Aryl Hydrocarbon Receptor (*Ahr*) had a significant effect on its abundance in plasma. Most interestingly, we identified variants in the Regulatory Factor X1 (*Rfx1*) gene that influence the abundance of the IL-17A protein in plasma.

Conclusion: Our study reports an innovative pipeline for the identification of genetic modifiers that may be targeted for drug development.

Background

Proteins expressed in blood plasma are diverse and their levels are dependent on environmental factors and genetic background [1]. Genes that influence the expression of other genes are called modifiers and can be detected using quantitative trait loci (QTL) mapping [2, 3]. Although protein QTLs (pQTLs) have been detected in the plasma of humans and mice, they mainly detected *cis*-acting pQTLs, not *trans*-acting pQTLs, due to a lack of statistical power, a low genetic diversity, or low-throughput protein level screening platforms [1, 4–9].

To overcome the lack of genetic diversity in mouse models, the Diversity Outbred (DO) mouse model was established through a multiparent paradigm from eight founder strains, comprising five inbred and three wild-derived strains [10, 11]. The combination of eight founder strains results in a much greater level of genetic diversity than existing recombinant inbred lines and the genetic variants are more uniformly distributed across the genome than in other genetic reference populations [12]. Each DO mouse is a genetically unique individual with a high level of allelic heterozygosity, providing precision for mapping QTL with relatively small sample sizes compared to human mapping studies. For example, QTL mapping in these multiparent populations resulted in the identification of genetic modifiers for the viral response [13, 14], kidney disease [15], atherosclerosis [16] and heart size [17].

We measured the abundance of clinically relevant plasma proteins in 140 DO mice using the Olink Mouse Exploratory Panel. This proximity by extension assay (PEA) measures 92 proteins in just 1 µL of plasma [9, 18]. In addition, each DO mouse was genotyped using the third generation Mouse Universal Genotyping Assay (MUGA), the GigaMUGA [19]. To identify new modifier genes that influence the level of

plasma proteins, we performed pQTL mapping (Fig. 1a). To the best of our knowledge, this is the first discovery study to report *trans*-acting pQTLs in the plasma of DO mice.

Results

Plasma pQTL mapping identifies 18 significant and 5 suggestive pQTL

To identify modifier genes that affect levels of plasma proteins, we performed pQTL mapping for the analytes measured using the Mouse Exploratory Panel from Olink Proteomics with genetic relatedness as a covariate. Significant pQTL were defined by a genome-wide P-value < 0.05, while suggestive pQTL were defined by 0.05 < P-value < 0.1. We identified a total of 18 significant and 5 suggestive pQTL (Fig. 1b and c, Table 1). Six of the 18 significant pQTL were *trans* pQTL (Fig. 1b, blue lines) and 12 were *cis* pQTL (Fig. 1b, red lines), while four of the suggestive pQTL were *trans* pQTL (Fig. 1c, blue lines) and one was a *cis* pQTL (Fig. 1c, red lines). A list of all identified pQTL can be found in Table 1 and a graphical presentation of the QTL support intervals can be found in Fig. S1.

Table 1
pQTL for plasma proteins in DO mice

Protein	Chromosome with pQTL peak	Peak LOD[1]	Position (Interval, cM)	*Cis*/*trans*	Significant ($p < 0.05$)[2]
ADAM23	1	10.2	28.7 (28.4–28.8)	*cis*	Yes
ADAM23	5	7.5	27.8 (27.4–28.9)	*trans*	No
AHR	12	12.8	14.5 (14.2–16.2)	*cis*	Yes
CA13	3	22.9	1.9 (1.8–2.3)	*cis*	Yes
CNTN1	15	21.7	40.8 (40.7–40.8)	*cis*	Yes
CNTN4	6	37.4	45 (44.9–45)	*cis*	Yes
ENO2	6	21.4	55.5 (55.5–55.6)	*cis*	Yes
FAS	15	8.3	44.8 (44.5–45.5)	*trans*	Yes
IL17A	8	8.3	35.5 (35.4–36.3)	*trans*	Yes
IL23R	6	35.8	28.1 (27-28.1)	*cis*	Yes
MATN2	14	8.7	16.7 (16.3–16.8)	*trans*	Yes
NADK	16	7.4	34 (29.7–34.9)	*trans*	No
NOTCH3	17	7.5	15.9 (14.6–16.4)	*cis*	No
RGMA	8	14.1	18.8 (18.8–18.8)	*trans*	No
RGMA	14	16.1	35 (35–35)	*trans*	Yes
SEZ6L2	7	29.6	60 (60-60.2)	*cis*	Yes
TNFRSF11B	1	8.4	60.3 (60.2–61)	*trans*	Yes
TPP1	7	8.8	45.8 (45.5–47)	*cis*	Yes

[1]LOD = logarithm of the odds; [2] "Yes" indicates a $p < 0.05$ while "No" corresponds to $0.1 < p < 0.05$.

Protein	Chromosome with pQTL peak	Peak LOD[1]	Position (Interval, cM)	*Cis*/*trans*	Significant ($p < 0.05$)[2]
TPP1	10	7.7	37.7 (37.4–39.3)	*trans*	No
VEGFD	X	12.1	68.7 (68.3–68.8)	*cis*	Yes
VSIG2	7	7.4	8 (7.8–27)	*trans*	No
VSIG2	9	10.1	15.5 (15.5–17.3)	*cis*	Yes
WFIKKN2	11	10	51.5 (50.6–51.9)	*cis*	Yes

[1]LOD = logarithm of the odds; [2] "Yes" indicates a $p < 0.05$ while "No" corresponds to $0.1 < p < 0.05$.

Variants in Ahr influence its plasma protein abundance

We identified a significant *cis* pQTL for the aryl hydrocarbon receptor (AHR) protein with a logarithm of the odds (LOD) score of 12.8 and a peak located at 33.3 Mb on chromosome 12 (Fig. 1b and 2a). Fine mapping of this pQTL resulted in the identification of the missense variant p.K432R in *Ahr*, where the presence of the minor G allele significantly increased the protein abundance of AHR (P-value = 3.41×10^{-7}, Fig. 2b). Protein sequence alignment showed conservation of the lysine at position 432 between mouse and human (Fig. S2a). In addition, we identified a variant located in the 5' untranslated region (UTR) of *Ahr*, where a significant decrease in AHR protein levels was observed with the presence of the minor A allele of c.-100G > A (P-value = 6.3×10^{-9}, Fig. 2b).

Rfx1 is a genetic modifier for IL-17A in plasma

One of the significant *trans* pQTL we identified was for the interleukin 17A (IL-17A) protein with a LOD score of 8.3 and a peak at 84.8 Mb on chromosome 8 (Fig. 3a). Fine mapping of this region resulted in a locus of approximately 3 Mb in size (Fig. 3b) [20, 21]. To identify the gene that could explain the variation in IL-17A protein levels, we investigated the 163 genes located in a 4 Mb interval around the peak. The "Shortest Path" algorithm from the MetaCore™ "Build Network" function was used to connect the candidate modifier genes with IL-17A. A direct interaction was identified between regulatory factor X1 (RFX1) and IL-17A (Fig. 3c, green highlight). Therefore, we hypothesize that *Rfx1* is a genetic modifier for IL-17A abundance in plasma. We identified two coding variants, p.F612S and p.A724, (Fig. 3d) in *Rfx1*, as well as one in the 3' UTR (c.*935G > A, Fig. 3d). Although the presence of one minor C allele of p.F612S did not affect IL-17A protein levels, we did observe a significant decrease in IL-17A protein levels in DO mice carrying two minor C alleles (P-value = 0.0036, Fig. 3d) compared to one. For the silent variant,

p.A724, we did not observe any significant effect on the protein abundance of IL-17A (Fig. 3d), as might be expected. Because p.F612 is conserved between mouse and human (Fig. S2b), this amino acid residue might regulate the function of RFX1 protein. Furthermore, we also observed a significant decrease in IL-17A protein abundance in heterozygous carriers of c.*935G > A compared to homozygous wild type carriers (P-value = 0.0049, Fig. 3d), while no significant difference was observed in homozygous carriers of the minor A allele. The c.*935G > A variant was also conserved between human and mouse (Fig. S2c). Because microRNAs are known to bind 3' UTR and downregulate gene expression, we searched for potential microRNAs that might bind around this SNP. However, this variant is not located in a known miRNA binding site according to PicTar and miRDB database [22, 23].

Discussion

Although plasma proteins are the preferred source of disease biomarkers, our understanding of how genetic variants affects the levels of these proteins is still limited [7, 18]. Furthermore, their abundance is also highly dependent on environmental factors, making it hard to study the genetic factors that influence their levels in humans. In this study, we measured the levels of 92 proteins in plasma obtained from 140 DO mice. The DO mouse population was chosen because of the advantage of eliminating the effects of environment while profiting from a genetic diversity similar to what is observed in humans. Additionally, genetic modifiers that have been discovered via QTL mapping in the DO mouse population have uncovered genetic mechanisms that have been successfully translated to humans [24, 25].

The Olink Mouse Exploratory Panel was used to measure 92 clinically relevant proteins in just 1 μL of plasma [9, 18]. Of the 92 proteins in the assay, 76 were reliably detected in more than 10% of the samples (Table S1) and were used for pQTL mapping. We identified a total of 18 significant and 5 suggestive pQTL (Fig. 1b and c). Of the significant pQTL, 12 were located in *cis* while 6 were in *trans*. One of the significant *cis* pQTL was for AHR. We identified a missense (p. K432R) and a 5' UTR (c.-100G > A) variant in *Ahr* that significantly affected AHR protein levels in plasma (Fig. 2b). Although p.K432R is not located in a known protein domain [26] nor is c.-100G > A located in a predicted 5' UTR motif, our data do suggest that coding variants in *Ahr* are responsible for the differences in AHR protein abundance.

We also demonstrated that the protein level of IL-17A in plasma is linked to significant *trans* pQTL on chromosome 8. At this locus, three different variants, one missense (p.F612S), one silent (p.A724) and one located in the 3' UTR (c.*935G > A, Fig. 3d) in *Rfx1* were identified and we showed that the genotypes of both the missense and 3' UTR variant significantly changed the protein abundance of IL-17A. Therefore, our data suggest that *Rfx1* is a genetic modifier for IL-17A plasma levels. This hypothesis is further supported by a recent article by Zhao and colleagues [27]. They showed that the levels of IL-17A mRNA and protein significantly increased after knock-down of RFX1 in $CD4^+$ T cells, while a decrease in IL-17A level was observed after overexpressing RFX1. Furthermore, they demonstrated that RFX1 directly affects IL-17A expression by binding one of its two binding sites upstream of the transcription start site

of the *IL17A* gene. RFX1 also played an important role in methylation and acetylation of the promoter region of *IL17A* [27]. The results from Zhao et al. [27] provide functional evidence directly linking RFX1 and IL-17A. This proof-of-concept case study clearly demonstrates the high potential of our approach to the identification of genetic modifiers for many other proteins.

Conclusions

We identified 18 significant pQTLs for the discovery of novel genetic modifiers that control levels of plasma proteins. Of note, we found new coding variants in *Ahr* that alter AHR protein abundance in plasma. We also report both coding and non-coding variants in the *Rfx1* gene that alter IL-17A plasma protein abundance, replicating a known connection between RFX1 and IL-17A in mouse and human [28]. Additionally, we show that the genetic diversity of the DO mouse model makes it possible to identify *trans* pQTL that can easily be translated to human, making this mouse model suitable for translational discovery studies. Importantly, we demonstrated that fewer than 150 DO mice are sufficient to identify statistically significant pQTL compared to the larger number of DO mice used in previous studies [16, 17, 25, 29–32]. The methodology detailed in this study can be used to unravel the complex mechanisms between genetic loci and protein abundance.

Methods

Ethics statement

All experiments and methods were performed in accordance with relevant guidelines and regulations. All animal experiments were performed in accordance with the approved animal protocol and guidelines established by the Institutional Animal Car Committee at Mayo Clinic, Jacksonville, FL (#A00003398-00). This study was carried out in compliance with the ARRIVE guidelines.

Mice

Male DO mice (n = 140, Stock No. 009376) were obtained in two batches from The Jackson Laboratory at 4-weeks of age and at generation 28 (G28 litter 1 and G28 litter 2) of outcrossing. The mice were housed under standard laboratory conditions on 12-h light:dark cycles in a specific pathogen-free environment at Mayo Clinic Jacksonville, Florida. At 13-weeks of age, the mice were anesthetized using ketamine (90 mg/kg, *i.p.*) and xylazine (10 mg/kg, *i.p.*). Around 200 µL of whole blood was collected into an EDTA coated tube (BrainTree Scientific Inc, Cat. No. MV-CB300-16444-BX) from the left ventricle and centrifuged at 2,000 x *g* for 15 minutes to separate out plasma. After collection of the blood, the mice were sacrificed by transcardial perfusion with ice-cold phosphate buffered saline.

Genotyping

Tail samples of all DO mice (n = 140) were collected and sent to GeneSeek (Neogen) for genotyping on the GigaMUGA. The GigaMUGA contains 143,259 genetic markers that were specifically designed for

genetic mapping in the DO mouse population [19, 33]. Genotype quality was assessed using the R package argyle. All variants were used for further analysis [34].

Olink Mouse Exploratory Panel

Plasma samples of all 140 DO mice were randomly distributed across 96-well plates and shipped to Olink Proteomics (Olink Proteomics, Uppsala, Sweden). The Olink technology allows the measurement of up to 92 proteins in 1 µL of plasma using the proximity extension assay (PEA). Distinct polyclonal oligonucleotide-labeled antibodies were added to the samples and bind their target proteins. Proteins were quantified by real-time quantitative PCR of the amplified oligonucleotide tags [35]. Primary data acquisition and quality control analysis were performed by Olink Proteomics that produced the Normalized Protein eXpression (NPX), an arbitrary unit used by Olink in Log2 scale [9, 35]. 83% (76/92) of the detected proteins were detected in more than 10% of the samples and were used for pQTL mapping (Table S1 and File S1).

pQTL mapping

pQTL mapping was performed using the R package R/qtl2 [36]. Founder haplotype probabilities were predicted using a Hidden Markov Model adapted for multi-parent populations and the protein abundance was regressed on these founder haplotype probabilities. To account for genetic similarity between mice, the kinship matrix was determined based on the leave-one-chromosome-out method. Genome scans were performed, and a random effect was included to account for kinship. Significance thresholds were determined using 1,000 permutations and the mapping statistic is the logarithm of the odds (LOD) score. pQTL significance intervals were defined by the 95% Bayesian credible interval. Significance threshold of a pQTL was set at $P < 0.05$, and suggestive at $0.05 < P < 0.1$. Genes within 2 Mb of the top SNPs, both upstream and downstream, were considered as candidate modifiers for further study.

Motif and miRNA binding site discovery, sequence alignment and network analysis

To identify whether the 5' UTR variant in Ahr (c.-100G > A) was located in an UTR motif, the online tools MEME Suite 5.1.1 [37] (http://meme-suite.org/index.html, Accessed on 05/26/2020) and Regma 2.0 [38] (http://regma2.mbc.nctu.edu.tw/index.html, Accessed on 05/26/2020) were used. To identify whether the 3' UTR variant in Rfx1 (c.*935G > A) was located in a miRNA binding site, the online tools PicTar [23] (https://pictar.mdc-berlin.de/, Accessed on 05/26/2020) and miRDB [22] (http://mirdb.org/, Accessed on 05/26/2020) were used. The Basic Local Alignment Search Tool (BLAST) was used to compare mouse and human sequences at the level of the transcript and protein. Coding variants were numbered relative to the translation initiation codon in the Ahr transcript (RefSeq NM_013464.4) or Rfx1 transcript (mouse: RefSeq NM_009055.4; human: RefSeq NM_002918). Amino acid numbering is according to AHR (human: GenPept accession number NP_001612.1; mouse: GenPept accession number NP_038492.1) or RFX1 (human: GenPept accession number NP_002909.4; mouse: GenPept accession number NP_033081.3). An overview of all identified variants and their position can be found in the Table S2. Network analysis was performed using the "Build Network" function in MetaCore™ (Clarivate Analytics, Accessed on

03/25/2020). The “Shortest path” algorithm was used to connect the genes located in the associated locus (“from”) with IL17A (“to”). The maximum number of steps allowed between the candidate modifier genes and IL17A was two, the minimum number of steps allowed using the “Shortest path”.

Statistical analysis

One-way ANOVA and Tukey multiple pairwise-comparison analyses were performed using R (R version 3.5.2).

Abbreviations

AHR: Aryl Hydrocarbon Receptor

CD4: Cluster of Differentiation 4

DO: Diversity Outbred

eQTL: Expression Quantitative Trait Locus

IL-17A: Interleukin 17a

LOD: Logarithm of Difference

MUGA: Mouse Universal Genotyping Array

pQTL: Protein Quantitative Trait Locus

PEA: Proximity Extension Assay

RFX1: Regulatory Factor X1

NPX: Normalized Protein eXpression

UTR: Untranslated Region

Declarations

Ethics approval and consent to participate

All animal experiments were performed in accordance with the approved animal protocol and guidelines established by the Institutional Animal Car Committee at Mayo Clinic, Jacksonville, FL (#A00003398-00).

Consent for publication

Not applicable

Availability of data and materials

The Olink Proteomics data are included as supplementary data (File S1), while the genotype data of the DO mice are available from the corresponding author.

Competing interests

The authors declare that they have no competing interests.

Funding

The research was supported by Eli Lilly-Stark Neuroscience fellowship (SP), the Indiana Clinical and Translational Sciences Institute, funded in part by grant #UL1TR002529 from the National Institutes of Health, National Center for Advancing Translational Sciences (SP), and the Paul and Carole Stark Fellowship (DJA). JK laboratory was supported by the Strategic Research Initiative (Indiana University), Precision Health Initiative (Indiana University), NIH R01AG054102, R01AG053500, R01AG053242, and R21AG050804.

Authors' contributions

SP conceived the design of the studies, participated in the dissection of the mice, performed the pQTL mapping and additional statistical analyses, interpreted the data, drafted the manuscript and created the figures. DJA assisted in pQTL mapping, interpretation of the data, and drafted the manuscript and figures. BK participated in the dissection of the mice, the collection of plasma, and revised the manuscript. HK participated in the dissection of the mice and the collection of plasma. JK conceived and oversighted the project and was involved in the design of the studies, data interpretation, provided funding, contributed to the writing and critically reviewed the manuscript.

Acknowledgements

We thank Matthew Heck for his contribution in dissecting the DO mice. We would also like to thank Dr. Luke Dabin for his critical revision of this manuscript. Illustrations in this paper were created using Seriver Medical Art templates, which are licensed under a Creative Commons Attribution 3.0 Unported License.

References

1. Melzer D, Perry JR, Hernandez D, Corsi AM, Stevens K, Rafferty I, Lauretani F, Murray A, Gibbs JR, Paolisso G *et al*: **A genome-wide association study identifies protein quantitative trait loci (pQTLs)**. *PLoS Genet* 2008, **4**(5):e1000072.

2. Génin E, Feingold J, Clerget-Darpoux F: **Identifying modifier genes of monogenic disease: strategies and difficulties**. *Hum Genet* 2008, **124**(4):357-368.
3. Grisel JE, Crabbe JC: **Quantitative Trait Loci Mapping**. *Alcohol Health Res World* 1995, **19**(3):220-227.
4. Solomon T, Lapek JD, Jr., Jensen SB, Greenwald WW, Hindberg K, Matsui H, Latysheva N, Braekken SK, Gonzalez DJ, Frazer KA *et al*: **Identification of Common and Rare Genetic Variation Associated With Plasma Protein Levels Using Whole-Exome Sequencing and Mass Spectrometry**. *Circ Genom Precis Med* 2018, **11**(12):e002170.
5. Holdt LM, von Delft A, Nicolaou A, Baumann S, Kostrzewa M, Thiery J, Teupser D: **Quantitative Trait Loci Mapping of the Mouse Plasma Proteome (pQTL)**. *Genetics* 2013, **193**(2):601-608.
6. Lourdusamy A, Newhouse S, Lunnon K, Proitsi P, Powell J, Hodges A, Nelson SK, Stewart A, Williams S, Kloszewska I *et al*: **Identification of cis-regulatory variation influencing protein abundance levels in human plasma**. *Human Molecular Genetics* 2012, **21**(16):3719-3726.
7. Sun BB, Maranville JC, Peters JE, Stacey D, Staley JR, Blackshaw J, Burgess S, Jiang T, Paige E, Surendran P *et al*: **Genomic atlas of the human plasma proteome**. *Nature* 2018, **558**(7708):73-79.
8. Yao C, Chen G, Song C, Keefe J, Mendelson M, Huan T, Sun BB, Laser A, Maranville JC, Wu H *et al*: **Genome-wide mapping of plasma protein QTLs identifies putatively causal genes and pathways for cardiovascular disease**. *Nat Commun* 2018, **9**(1):3268.
9. Folkersen L, Fauman E, Sabater-Lleal M, Strawbridge RJ, Franberg M, Sennblad B, Baldassarre D, Veglia F, Humphries SE, Rauramaa R *et al*: **Mapping of 79 loci for 83 plasma protein biomarkers in cardiovascular disease**. *PLoS Genet* 2017, **13**(4):e1006706.
10. Churchill GA, Gatti DM, Munger SC, Svenson KL: **The Diversity Outbred mouse population**. *Mamm Genome* 2012, **23**(9-10):713-718.
11. Chick JM, Munger SC, Simecek P, Huttlin EL, Choi K, Gatti DM, Raghupathy N, Svenson KL, Churchill GA, Gygi SP: **Defining the consequences of genetic variation on a proteome-wide scale**. *Nature* 2016, **534**(7608):500-505.
12. Collaborative Cross C: **The genome architecture of the Collaborative Cross mouse genetic reference population**. *Genetics* 2012, **190**(2):389-401.
13. Martin MD, Sompallae R, Winborn CS, Harty JT, Badovinac VP: **Diverse CD8 T Cell Responses to Viral Infection Revealed by the Collaborative Cross**. *Cell Reports* 2020, **31**(2):107508.
14. Manet C, Simon-Lorière E, Jouvion G, Hardy D, Prot M, Flamand M, Panthier J-J, Sakuntabhai A, Montagutelli X: **Genetic diversity of Collaborative Cross mice controls viral replication, clinical severity and brain pathology induced by Zika virus infection, independently of Oas1b**. *bioRxiv* 2019:677484.
15. Huda MN, VerHague M, Albright J, Smallwood T, Bell TA, Que E, Miller DR, Roshanravan B, Allayee H, Manuel de Villena FP *et al*: **Dissecting the Genetic Architecture of Cystatin C in Diversity Outbred Mice**. *G3 (Bethesda)* 2020, **10**(7):2529-2541.
16. Smallwood TL, Gatti DM, Quizon P, Weinstock GM, Jung KC, Zhao L, Hua K, Pomp D, Bennett BJ: **High-resolution genetic mapping in the diversity outbred mouse population identifies Apobec1 as a**

candidate gene for atherosclerosis. *G3 (Bethesda)* 2014, **4**(12):2353-2363.

17. Shorter JR, Huang W, Beak JY, Hua K, Gatti DM, de Villena FP, Pomp D, Jensen BC: **Quantitative trait mapping in Diversity Outbred mice identifies two genomic regions associated with heart size**. *Mamm Genome* 2018, **29**(1-2):80-89.
18. Assarsson E, Lundberg M, Holmquist G, Björkesten J, Bucht Thorsen S, Ekman D, Eriksson A, Rennel Dickens E, Ohlsson S, Edfeldt G *et al*: **Homogenous 96-Plex PEA Immunoassay Exhibiting High Sensitivity, Specificity, and Excellent Scalability**. *PLOS ONE* 2014, **9**(4):e95192.
19. Morgan AP, Fu CP, Kao CY, Welsh CE, Didion JP, Yadgary L, Hyacinth L, Ferris MT, Bell TA, Miller DR *et al*: **The Mouse Universal Genotyping Array: From Substrains to Subspecies**. *G3-Genes Genom Genet* 2016, **6**(2):263-279.
20. Bryant CD, Ferris MT, De Villena FPM, Damaj MI, Kumar V, Mulligan MK: **Chapter 8 - Reduced Complexity Cross Design for Behavioral Genetics**. In: *Molecular-Genetic and Statistical Techniques for Behavioral and Neural Research.* Edited by Gerlai RT. San Diego: Academic Press; 2018: 165-190.
21. Paigen K: **1 - Genetic Factors in Developmental Regulation**. In: *Physiological Genetics.* Edited by Scandalios JG: Academic Press; 1979: 1-61.
22. Chen Y, Wang X: **miRDB: an online database for prediction of functional microRNA targets**. *Nucleic Acids Res* 2020, **48**(D1):D127-D131.
23. Krek A, Grün D, Poy MN, Wolf R, Rosenberg L, Epstein EJ, MacMenamin P, da Piedade I, Gunsalus KC, Stoffel M *et al*: **Combinatorial microRNA target predictions**. *Nature Genetics* 2005, **37**(5):495-500.
24. Winter JM, Gildea DE, Andreas JP, Gatti DM, Williams KA, Lee M, Hu Y, Zhang S, Program NCS, Mullikin JC *et al*: **Mapping Complex Traits in a Diversity Outbred F1 Mouse Population Identifies Germline Modifiers of Metastasis in Human Prostate Cancer**. *Cell Syst* 2017, **4**(1):31-45 e36.
25. Yang C, Wang Y, Xu W, Liu Z, Zhou S, Zhang M, Cui D: **Genome-wide association study using diversity outcross mice identified candidate genes of pancreatic cancer**. *Genomics* 2018.
26. Fukunaga BN, Probst MR, Reisz-Porszasz S, Hankinson O: **Identification of functional domains of the aryl hydrocarbon receptor**. *J Biol Chem* 1995, **270**(49):29270-29278.
27. !!! INVALID CITATION !!! [23].
28. Zhao M, Tan Y, Peng Q, Huang C, Guo Y, Liang G, Zhu B, Huang Y, Liu A, Wang Z *et al*: **IL-6/STAT3 pathway induced deficiency of RFX1 contributes to Th17-dependent autoimmune diseases via epigenetic regulation**. *Nat Commun* 2018, **9**(1):583.
29. Kemis JH, Linke V, Barrett KL, Boehm FJ, Traeger LL, Keller MP, Rabaglia ME, Schueler KL, Stapleton DS, Gatti DM *et al*: **Genetic determinants of gut microbiota composition and bile acid profiles in mice**. *PLOS Genetics* 2019, **15**(8):e1008073.
30. Yuan JT, Gatti DM, Philip VM, Kasparek S, Kreuzman AM, Mansky B, Sharif K, Taterra D, Taylor WM, Thomas M *et al*: **Genome-wide association for testis weight in the diversity outbred mouse population**. *Mamm Genome* 2018, **29**(5-6):310-324.

31. Recla JM, Bubier JA, Gatti DM, Ryan JL, Long KH, Robledo RF, Glidden NC, Hou G, Churchill GA, Maser RS *et al*: **Genetic mapping in Diversity Outbred mice identifies a Trpa1 variant influencing late-phase formalin response**. *Pain* 2019, **160**(8):1740-1753.
32. Logan RW, Robledo RF, Recla JM, Philip VM, Bubier JA, Jay JJ, Harwood C, Wilcox T, Gatti DM, Bult CJ *et al*: **High-precision genetic mapping of behavioral traits in the diversity outbred mouse population**. *Genes Brain Behav* 2013, **12**(4):424-437.
33. Chesler EJ, Gatti DM, Morgan AP, Strobel M, Trepanier L, Oberbeck D, McWeeney S, Hitzemann R, Ferris M, McMullan R *et al*: **Diversity Outbred Mice at 21: Maintaining Allelic Variation in the Face of Selection**. *G3 (Bethesda)* 2016, **6**(12):3893-3902.
34. Morgan AP: **argyle: An R Package for Analysis of Illumina Genotyping Arrays**. *G3 (Bethesda)* 2015, **6**(2):281-286.
35. Assarsson E, Lundberg M, Holmquist G, Bjorkesten J, Thorsen SB, Ekman D, Eriksson A, Rennel Dickens E, Ohlsson S, Edfeldt G *et al*: **Homogenous 96-plex PEA immunoassay exhibiting high sensitivity, specificity, and excellent scalability**. *PLoS One* 2014, **9**(4):e95192.
36. Broman KW, Gatti DM, Simecek P, Furlotte NA, Prins P, Sen S, Yandell BS, Churchill GA: **R/qtl2: Software for Mapping Quantitative Trait Loci with High-Dimensional Data and Multiparent Populations**. *Genetics* 2019, **211**(2):495-502.
37. Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, Ren J, Li WW, Noble WS: **MEME SUITE: tools for motif discovery and searching**. *Nucleic Acids Res* 2009, **37**(Web Server issue):W202-208.
38. Chang TH, Huang HY, Hsu JB, Weng SL, Horng JT, Huang HD: **An enhanced computational platform for investigating the roles of regulatory RNA and for identifying functional RNA motifs**. *BMC Bioinformatics* 2013, **14 Suppl 2**:S4.

Figures

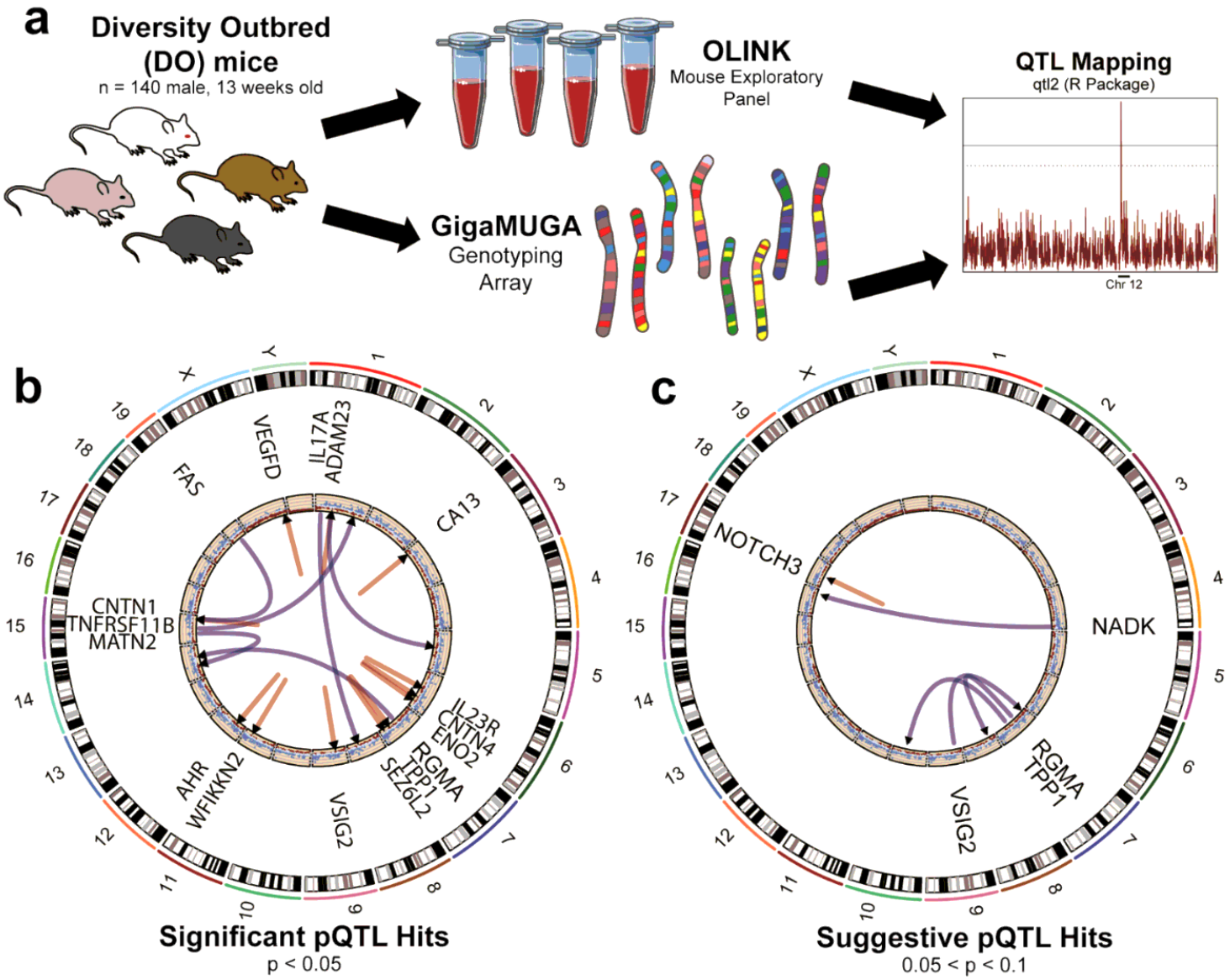

Figure 1

pQTL mapping of plasma proteins reveals 18 significant and 5 suggestive candidate modifier loci. (a) Workflow for pQTL mapping. (b) pQTL mapping of plasma proteins detected by the Olink Mouse Exploratory panel identified 18 significant (P-value < 0.05) pQTL. The first track includes the protein symbols mapped to their respective loci in the mouse genome (mm10). The second track is a scatterplot showing the density of the GigaMUGA genotyping array (# markers/0.5Mb; red: ≤ 20 markers; light blue: ≤ 30 markers; dark blue: ≥ 30 markers; clipped at 100 markers/0.5Mb). The inner track displays trans (blue) and cis (red) pQTL with a black triangle at the associated locus. (c) pQTL mapping of plasma proteins detected 5 suggestive pQTL (0.05 < P-value < 0.1). Visualization is similar to above, with trans (blue) and cis (red) pQTL.

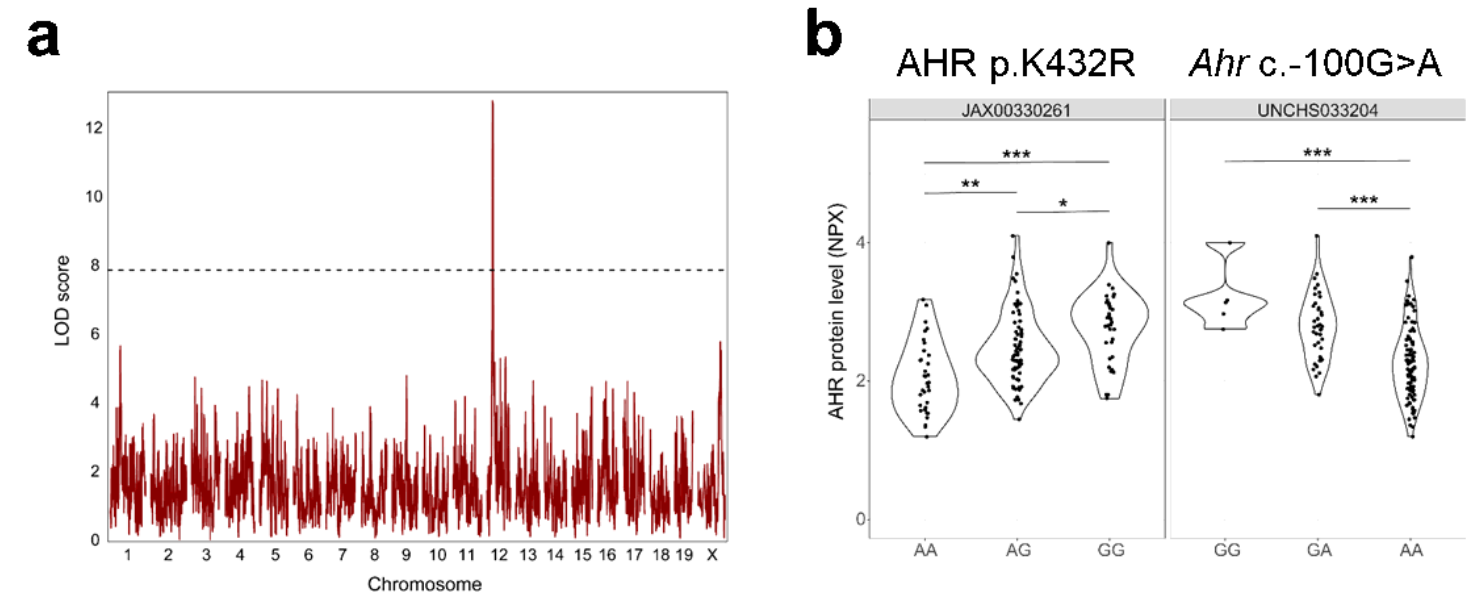

Figure 2

Variants in Ahr have a significant effect on AHR levels in plasma of DO mice. (a) Genome scan of the AHR protein shows a significant peak on chromosome 12. Chromosome position is on the x-axis and LOD score on the y-axis. Dashed line at P-value ≤ 0.05 significance threshold. (b) Violin plots showing the effect of the p.K432R and c.-100G>A genotypes on AHR protein abundance. *** P-value < 0.001; ** P-value < 0.01; * P-value < 0.05

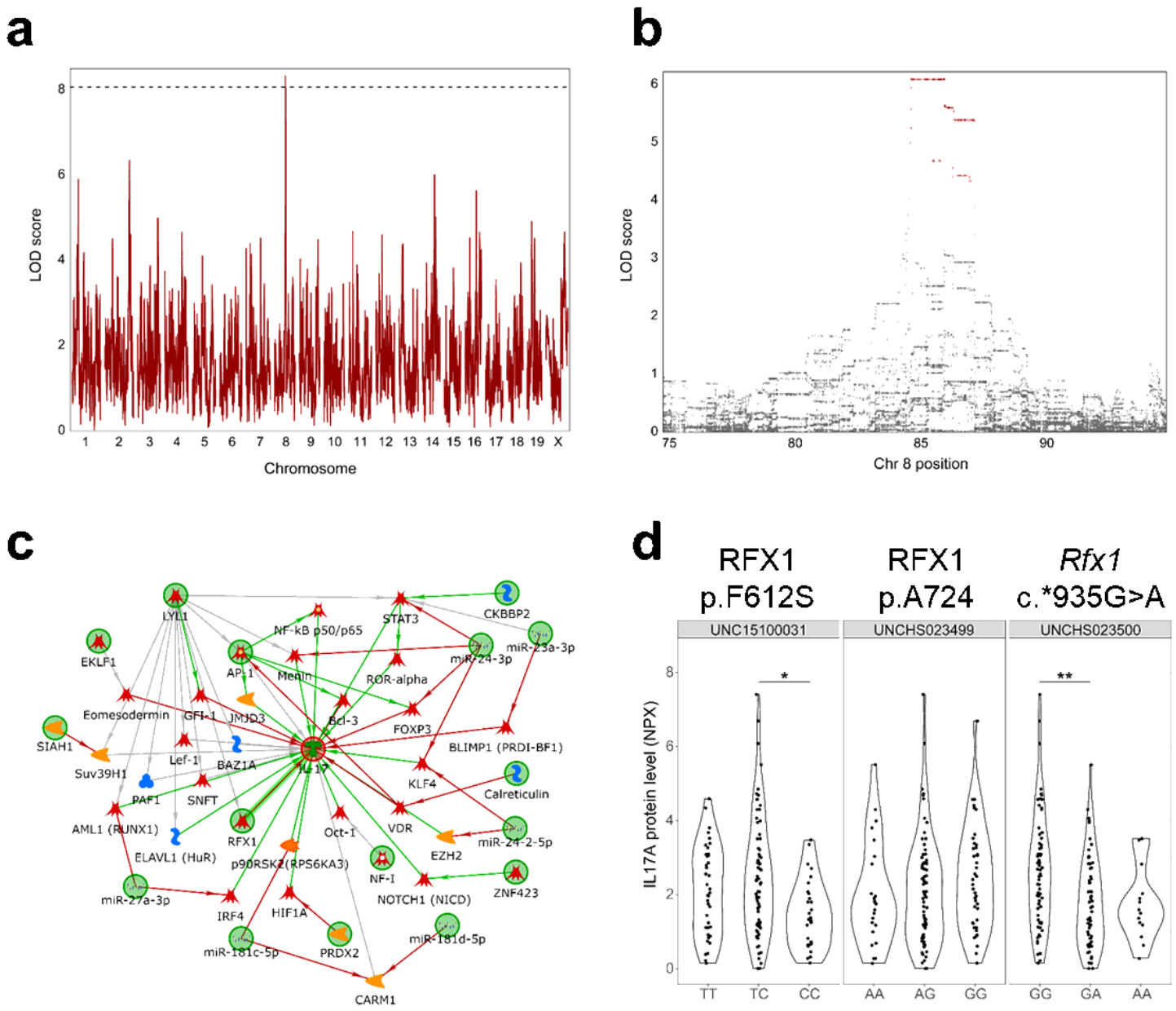

Figure 3

Variants in Rfx1 have a significant effect on IL-17A levels in plasma of DO mice. (a) Genome scan for the IL-17A protein shows a significant peak on chromosome 8. (b) Association mapping of the IL-17A pQTL. Zoom in on chromosome 8 showing the QTL support interval. Magenta marks SNPs with a LOD drop < 2 from the top SNP. (c) MetaCore™ network analysis connecting the genes located in the pQTL locus on chromosome 8 for IL-17A. Green circles indicate either IL-17A or genes that were present in the locus. (d) Violin plots showing the effect of the RFX1 p.F612S, p.A724 and c.*935G>A genotypes on IL-17A protein abundance. ** P-value < 0.01; * P-value < 0.05

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Additionalfile1.xlsx
- OlinkSuppv1120210201SPH.pdf