Hepatic gene expression variations in response to high-fat diet-induced impaired glucose tolerance using RNAseq analysis in collaborative cross mouse population

H. J. Abu-Toamih Atamni[1] · G. Kontogianni[2] · I. Binenbaum[2,3] · R. Mott[4] · H. Himmelbauer[5,6,7] · H. Lehrach[8] · A. Chatziioannou[2,9] · Fuad A. Iraqi[1]

Abstract

Hepatic gene expression is known to differ between healthy and type 2 diabetes conditions. Identifying these variations will provide better knowledge to the development of gene-targeted therapies. The aim of this study is to assess diet-induced hepatic gene expression of susceptible versus resistant CC lines to T2D development. Next-generation RNA-sequencing was performed for 84 livers of diabetic and non-diabetic mice of 41 different CC lines (both sexes) following 12 weeks on high-fat diet (42% fat). Data analysis revealed significant variations of hepatic gene expression in diabetic versus non-diabetic mice with significant sex effect, where 601 genes were differentially expressed (DE) in overall population (males and females), 718 genes in female mice, and 599 genes in male mice. Top prioritized DE candidate genes were *Lepr*, *Ins*2, *Mb*, *Ckm*, *Mrap*2, and *Ckmt*2 for the overall population; for females-only group were *Hdc*, *Serpina*12, *Socs*1, *Socs*2, and *Mb*, while for males-only group were *Serpine*1, *Mb*, *Ren*1, *Slc4a*1, and *Atp2a*1. Data analysis for sex differences revealed 193 DE genes in health (Top: *Lepr*, *Cav*1, *Socs*2, *Abcg*2, and *Col5a*3), and 389 genes DE between diabetic females versus males (Top: *Lepr*, *Clps*, *Ins*2, *Cav*1, and *Mrap*2). Furthermore, integrating gene expression results with previously published QTL, we identified significant variants mapped at chromosomes at positions 36–49 Mb, 62–71 Mb, and 79–99 Mb, on chromosomes 9, 11, and 12, respectively. Our findings emphasize the complexity of T2D development and that significantly controlled by host complex genetic factors. As well, we demonstrate the significant sex differences between males and females during health and increasing to extent levels during disease/diabetes. Altogether, opening the venue for further studies targets the discovery of effective sex-specific and personalized preventions and therapies.

Introduction

According to the World Health Organization report (WHO 2014), a total of 38 million deaths occurred worldwide in the year 2012 due to non-communicable diseases (NCD), while Type 2 diabetes (T2D) was the fourth major leading cause of

✉ Fuad A. Iraqi
fuadi@tauex.tau.ac.il

1 Department of Clinical Microbiology and Immunology, Sackler Faculty of Medicine, Tel-Aviv University, Ramat Aviv, 69978 Tel Aviv, Israel

2 Institute of Biology, Medicinal Chemistry & Biotechnology, National Hellenic Research Foundation, Athens, Greece

3 Department of Biology, University of Patras, Patras, Greece

4 Department of Genetics, University College of London, London, UK

5 Centre for Genomic Regulation (CRG), Barcelona, Spain

6 Universitat Pompeu Fabra (UPF), Barcelona, Spain

7 University of Natural Resources and Life Sciences (BOKU), Vienna, Austria

8 Department of Vertebrate Genomics, Max Planck Institute for Molecular Genetics, Berlin, Germany

9 e-NIOS Applications PC, 17671 Kallithea, Greece

death (1.5 million deaths per year) (World Health Organization 2014). Diabetes prevalence is continuously increasing worldwide, and the annual number of NCD deaths is increasing steadily (unlike infectious diseases), while diabetes is predicted to be the seventh cause of death by the year 2030 even if the prevalence of obesity remains fixed (Mathers and Loncar 2006).

One of the most noticeable components of T2D is the significant variations in disease pathogenesis between ethnicities, which is attributed to complex genetic components and environmental factors (Samsom et al. 2016). Despite the fact that environmental factors of sedentary lifestyle, over-consumption of fat/carbohydrate enriched diets, socioeconomic status, and air pollution increase the risk for T2D, it is evident that the genetic background plays a critical role in determining the host response towards these environmental conditions (susceptibility versus resistance). Thence, studies of T2D etiology indicate various trends of development and progress worldwide, suggesting different prevention and treatment strategies based on ethnicity (Kaul and Ali 2016).

In addition to the environmental and genetic predisposition risk factors, the sex-related differences play major role in the pathogenesis of T2D (Legato et al. 2006). Sexually dimorphic gene expression varies at different stages of life during development (Yang et al. 2006), and under different environmental challenges, such as nutrition, as reported in the current study. In spite of multiplied expenses and time demands when conducting a research project with both sexes, sex differences must be considered in the advanced studies of translational research and personalized medicine, in order to increase the understanding of physiological and metabolic sex-specific pathways, which will enhance the potential effects of T2D therapies.

Studies on human population ethnic subgroups showed a significant sex effect in T2D onset and progress, which differ between males and females within and between the different ethnic groups (Gale and Gillespie 2001). The genetic basis for T2D development, affected by diet and sex interactions, can be addressed through human GWAS studies. However, using animal models at the earlier stages of the search promises accurate results, since, with such models, researchers can control the environmental conditions better than in human studies. Despite the complexity of T2D and sex effects, recently, NIH has changed its policies and is demanding the use of both sexes in diseases and treatment research unless sex-specific inclusion is unwarranted (Clayton and Collins 2014).

Dissecting the genetic factors controlling host susceptibility to T2D development requires controlled and standardized investigations of the interaction of environment and host genetic background. Such investigations are complex and almost impossible in humans due to multiple constraints, particularly the impossibility of controlling the environmental conditions of the study. Consequently, mouse models were developed for the study of human health and diseases, while the study of complex diseases required a mouse model exhibiting wide genetic diversity, to enable high-resolution mapping of genomic regions. To achieve this, new, genetically highly diverse, recombinant inbred lines of mouse population were established, namely the collaborative cross (CC), created from full reciprocal mating of eight divergent strains of mice: A/J, C57BL/6 J, 129S1/SvImJ, NOD/LtJ, NZO/HiLtJ, CAST/Ei, PWK/PhJ, and WSB/EiJ, generating a unique, genetic reference population (GRP), comprising of a set of approximately 75 recombinant, inbred lines (RIL), dubbed CC lines. In genetically defined CC lines, chromosomal regions responsible for the genetic variance of complex traits can be mapped as quantitative trait loci (QTL) in experimental populations available for precise study under defined environmental conditions. Additionally, to the QTL mapping approach, next-generation cDNA sequencing (mRNA-Seq) is a powerful approach for identifying gene expression variations between health and disease conditions. Once QTL and gene expression results identified, the genetic analysis can be extended successfully to humans, as in previous studies using the CC mouse model (Shusterman et al. 2017; Nashef et al. 2018).

In our previous studies, we have assessed the CC mice for T2D development due to high-fat diet challenge and found that CC mice have a wide range of T2D phenotypic variations development, which is believed due to its genetic diversity, and subsequently, we mapped QTL associated with the disease (Abu-Toamih Atamni et al. 2016a, b; 2017, 2018; Nashef et al. 2017). Herein, we present the results of hepatic gene expression in health (non-diabetic) versus disease (diabetic) in the overall population (regardless of sex), females and males separately, and comparison between males and females. Eventually, we present lists of the top prioritized candidate genes that were differentially expressed (DE) in hepatic tissues of mice with impaired versus non-impaired glucose tolerance.

Materials and methods

Ethical statement

All experiments in this study were approved by the Animal Use and Care Committee at TAU (approved experiment number M-12-025and M-14-007).

Breeding

At this time, about 100 CC lines are under development at the small animal facility of Tel-Aviv University (TAU) between generations of G27 to G64 of inbreeding by full-sib

mating (Iraqi et al. 2012). Due to genetic variations between the CC lines, breeding rate, number, and sex of litters in each cycle might vary. Therefore, the CC lines were assessed based on litters' availability, while making our best efforts to scan as many as possible of the CC lines with representation of both sexes.

Study cohort

Study cohort consisted of 84 mouse liver samples, generated from 41 different CC lines (N of mice between; 1 to 5 per line), of which 20 CC lines present in both sexes. For the females-only analysis, we used 31 out of the 41 lines, and 30 out of the 41 lines for males- only analysis, with N of mice between 1 and 3 per line in each analysis group. The CC line selection was randomly based on breeding availability. Mice were weaned at age of 3 weeks old and housed separately by sex with standard rodents' chow diet (TD.2018SC, Teklad Global, Harlan, Inc., Madison, WI, USA), containing % Kcal from fat 18%, protein 24%, and carbohydrates 58%, and water ad libitum.

Dietary challenge

The experimental mice started at age of 8 weeks old, and diet switched to purified HFD (42% fat), in which 21% of fat source is obtained from milk fat, for a period of 12 weeks with free access to food and water. The HFD, TD. 88137 (Harlan Teklad, Madison, WI, USA) contains 42% of calories from fat and 34.1% from carbohydrate, primarily sucrose.

Intraperitoneal glucose tolerance test (IPGTT)

The glucose tolerance test measures the glucose tolerance ability of the body, by monitoring the clearance rate of an intraperitoneal injected glucose load from the body. It is used to detect disturbances in glucose metabolism that can be linked to diabetes. At the end of 12 weeks on HFD challenge, mice were fasted for 6 h (6:00–12:00 a.m.); with free access to water. The fasting blood glucose levels were determined at time 0, prior to the intra-peritoneal (IP) injection of glucose solution (2.5 g glucose per kg mouse). Subsequently, blood glucose levels were measured at different time points during the following 3 h (time 0, 15, 30, 60, 120, and 180 min after glucose injection). Thereafter, mice were returned back to cages with high-fat food and free access to water for overnight recovery from the IPGTT procedure.

Tissue collecting

Following IPGTT, mice were returned to the same diet (HFD, 42% fat) as before the IPGTT assay for overnight recovery, in order to bring mice to their original T2D status conditions. On the following day, mice were killed by cervical dislocation (CD) protocol, and livers were immediately collected to liquid nitrogen and thereafter stored in a freezer (− 80 ℃) for RNA extraction.

Rationale behind samples selection for gene expression

For the purpose of comprehensively and rapidly defining the gene expression profile during T2D and metabolic syndrome development in the different CC lines, in response to the HFD, next-generation sequencing technologies were used (Voelkerding et al. 2009; Ansorge 2009; Buckingham 2010). Samples for the RNA-sequencing were selected based on diabetic status, where mice that reached the end time point of IPGTT with glucose levels < 400 mg/dL were considered "non-diabetic," while glucose levels > 400 mg/dL were considered "diabetic."

RNA extractions

Total RNA was isolated using a Qiagen kit (Cat.No.73404) as recommended by the manufacturer, and quality of the RNA samples was assessed using the BioAnalyzer 2100 (Agilent). The RNA Integrity Number (RIN) software was used to estimate the integrity of the total RNA sample by calculating the 18S to 28S rRNA ratios. Samples with RIN above 7.0 passed the quality control test.

mRNA-Seq libraries

Samples were prepared using the TruSeq Stranded mRNA library preparation kit (Illumina), including polyA+ RNA selection, RNA fragmentation, cDNA synthesis, and ligation of barcoded Illumina adapters. Indexed libraries were pooled and sequenced on the Illumina HiSeq 2000 and 2500 sequencers with Illumina v3 sequencing chemistry. Paired-end sequencing was performed by reading 50 bases at each end of a fragment. Overall, each sample consisted of 24–37.5 million sequenced cDNA fragments with an average of 31.5 million fragments per sample.

RNA-Seq quantification and analysis

The protocol used in this study has been previously published (Kontogianni et al. 2017) and here was adjusted for the mouse samples, and reads were aligned to the mouse genome and transcriptome (version: GRCm.38). With the aim of comparing the level of gene expression in different states, we utilized Cuffdiff (Trapnell et al. 2012). Cuffdiff allows the tracking of significant changes in transcript expression and splicing, between different states. It can

identify differentially expressed genes and transcripts, as well as genes that are being differentially regulated at the transcriptional and post-transcriptional level. In this context, we focused on the identification of differential expression. Various tests were performed between the samples, namely, all non-diabetic versus all diabetic (ANvsAD), female non-diabetic versus female diabetic (FNvsFD), male non-diabetic versus male diabetic (MNvsMD), female non-diabetic versus male non-diabetic (FNvsMN), and female diabetic versus male diabetic (FDvsMD). The differentially expressed (DE) genes were selected by setting a *P* value threshold of 0.05 and fold change of $|\log_2 FC| \geq 1$. Primary outcomes will be the gene expression levels of each gene in the hepatic tissue samples, and secondary outcomes will be the comparison of gene/s expression variations between the different experimental groups. DE genes were based on p-value and logFC, simultaneously. The reason for that is that the use of solely classical FDR cut-off thresholds filters down the initial lists, excluding causal genes with potential weak effect while keeping several genes with small fold changes between the groups, which run the risk of being artifacts of the normalization process. We chose to include to our lists genes with higher fold change and perform prioritized pathway analysis, exploiting the enrichment scores of gene sets mapped to distinct ontological terms afterwards, to find the significant genes. In this way, genes that simultaneously have a significant p-value and adequate fold change score, but also participate in biological processes with statistically corrected, significant enrichment score, are utilized. Further analysis was performed for the selection of "top prioritized" genes that were selected based on their significance level and enrichment analysis. These genes were presented in the heatmap of each analysis and defined as the most critical genes to ontological clusters on the Gene Ontology Biological process.

Functional analysis

In order to perform functional pathway analysis of the DE genes, we used the BioInfoMiner platform. This platform performs statistical and network analysis exploiting several biological hierarchical vocabularies for functional annotation (like GO Ashburner et al. 2000 and MGI Mammalian Phenotype Smith and Eppig 2009), aiming to detect and rank significantly altered biological processes and the respective driver genes linking these processes. BioInfoMiner corrects for potential semantic inconsistencies on the selected vocabulary's scheme, by linking the annotation of each gene with the ancestors of every direct correlated ontological term, consequently restoring the sound structure of an ontological tree. The BioInfoMiner platform is available online at the website https://bioinfominer.com. Gene regulatory networks were obtained from Consensus Path DB (Kamburov et al.

2012) using differentially expressed genes in normal versus diabetic mice (p-value < 0.05 and |logFC| > 1), including genes from all the performed tests.

Results

Glucose tolerance ability phenotypic variations

Figure 1 shows the glucose clearance ability of the CC lines after 12 weeks of dietary challenge, in response to a glucose solution load and 180 min IPGTT. For this analysis, we focused on glucose levels at the end time point of the IPGTT (Time 180′), where levels below 400 mg/dL were defined as a "success" (i.e.. non-diabetic) and over 400 mg/dL defined as "impaired' glucose tolerance. In Fig. 1, Bar Charts A, B, and C represent trait values for overall population (regardless of the sex), females, and males, respectively. Data analysis showed a significant (*P*-value < 0.01) variation between the CC lines for the actual trait values of blood glucose levels (mg/dL) measured at time point 180 min of the IPGTT. For the overall population, 8 out of the 41 lines showed impaired glucose tolerance, while for females only 5 out of 31, and 11 out of 30 lines for males only.

Hepatic gene expression

Study cohort consisted of 84 mouse liver samples, generated from 41 different CC lines, of which 20 CC lines with presentation of both sexes and collected at the end of 12 weeks of HFD challenge (20 weeks old). Following the approved IACUC Committee guidelines, animals were symptomatically healthy, i.e., did not show extraordinary health such as low locomotion, major weight loss (10% between weekly weights and over 20% of its weight), apathy, physical instability, difficulty in breathing, high levels of basal blood glucose (> 400 mg/dL), or behavior symptoms such as high aggressiveness/isolation. Gene expression data analysis was performed to recognize candidate genes that were DE in diabetic versus non-diabetic conditions, once for overall population (regardless of sex) and once for females-only and males-only groups. Moreover, data analysis for sex differences was performed to determine the differences in hepatic gene expression between females and males in health versus disease—i.e., "diabetic" males versus "diabetic" females, and "non-diabetic" females versus "non-diabetic" males. Summary Table 1 shows the total number of DE genes for each analysis group and the total number of over/under-expressed genes in each. Furthermore, below, we present the results as summarized in heatmaps and lists of the most significant candidate genes. Full details of the top prioritized genes, including gene rank, gene symbol/definition, FC, and significance level are available in the Online Appendix.

Fig. 1 Glucose levels (mg/dL) of 41 CC lines (**a**) and separately for 31 female lines (**b**) and 30 male lines (**c**), after 12 weeks on high-fat diet (HFD, 42% fat) measured at final time point (time 180 min) of IPGTT. *Y*-axis represents blood glucose levels (mg/dL). *X*-axis represents the different CC lines (IL and #). No. of animals: all (**a**) 84 mice, females (**b**) 42 mice, males (**c**) 42 mice
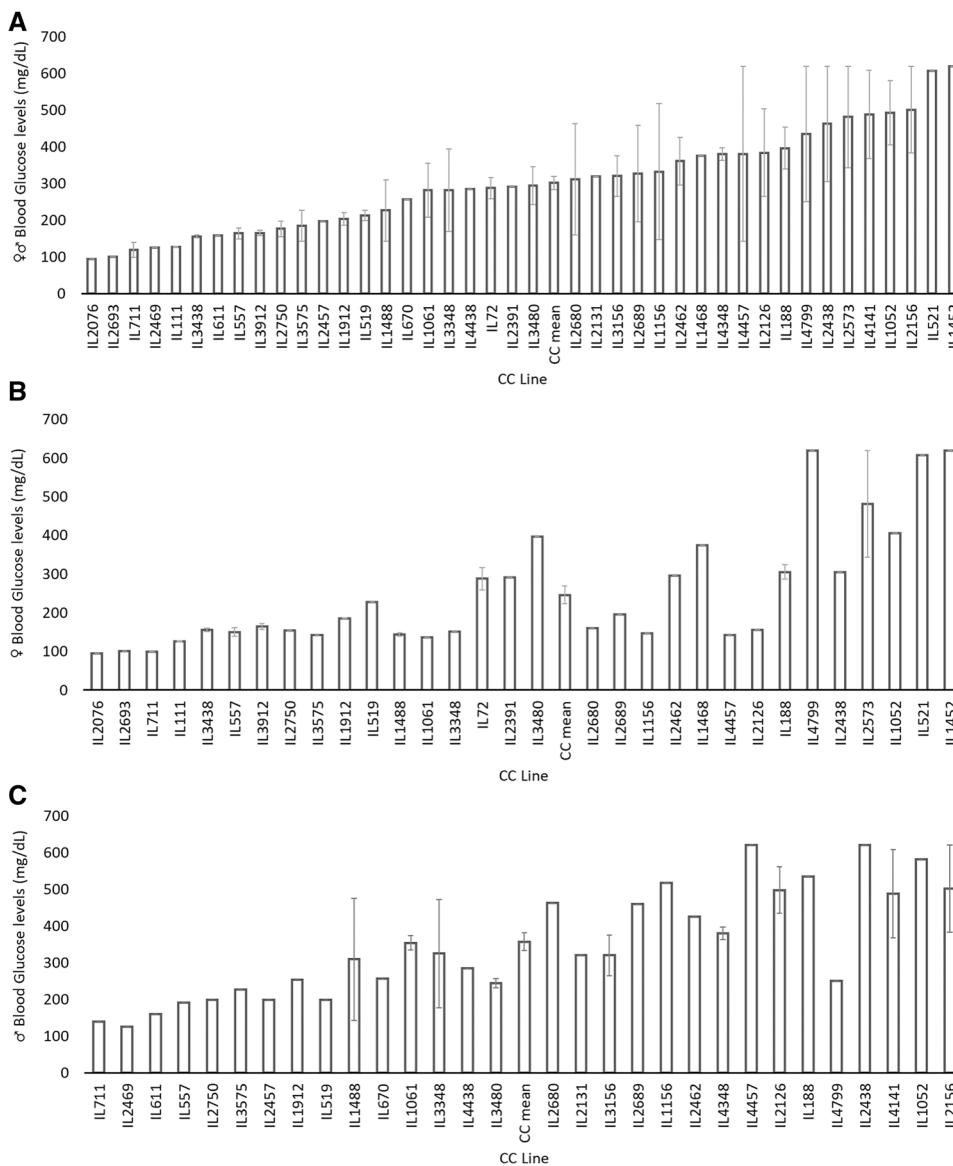
Table 1 Summary table of the total number of DE genes in each analysis group and the number of over-/under-expressed gene within each group

Analysis	Group	Symbol	Total genes	Over-expressed	Under-expressed
Non-diabetic versus diabetic	All	AN versus AD	601	449	152
	Females	FN versus FD	718	529	189
	Males	MN versus MD	599	430	169
Males versus females	Non-diabetic	MN versus FN	734	547	187
	Diabetic	MD versus FD	874	590	284

Moreover, Table 5 in the Online Appendix shows the differential expression of few positive control genes expressed from the X/Y chromosomes, where indeed Y Chr genes as expected show higher expression in males than in females. These results confirm the validity of the approach and obtained results.

Overall population diabetic versus non-diabetic mice

Data analysis revealed 601 genes to be DE (*P*-value < 0.05) between the overall population (both sexes) diabetic to non-diabetic hepatic tissues, of which 99 genes also had |log$_2$

fold change| > 1. In Fig. 2a, we present the heatmap of the top prioritized genes (fpkm values) for this test. The heatmap includes 28 genes defined as the most critical genes to ontological clusters derived from the Gene Ontology Biological process corpus, based on their significance level and enrichment analysis. Furthermore, we performed in-depth analysis of the significant DE genes with absolute log FC ratio higher than one (*P*-value < 0.05 and |logFC| > 1) using the MGI Mammalian Phenotype terms. Analysis results indicated 13 significant mammalian phenotype (MP) clusters/terms, which are highly associated with T2D and MetS phenotypes, directly or as side effect of health complications. Each MP term presents a different known phenotype of T2D and/or MetS, which is significantly enriched in the DE-expressed genes. Five of the foremost DE genes ranked on top of the highly prioritized genes include the *Lepr*, *Ins*2, *Mb*, *Ckm*, and *Ckmt*2 genes (Table 2). Our findings report under-expression of *Lepr* in hepatic tissue of diabetic versus non-diabetic mice (FC = − 1.062, *P*-value = 1.5×10^{-3}), which is significantly involved in five clusters of MP terms: MP:0001556- increased circulating HDL cholesterol level, MP:0004896- abnormal endometrium morphology, MP:0001426—polydipsia, MP:0005217—abnormal pancreatic beta-cell morphology, and MP:0002625—heart left ventricle hypertrophy. Contrarily, expression levels of *Ins2* (MGI:96573) were significantly (FC = 1.013, *P*-value = 0.034) elevated in diabetic versus non-diabetic mice. *Ins*2 gene over-expression was significantly involved in three clusters of MP terms: MP:0002625—heart left ventricle hypertrophy, MP:0005217—abnormal pancreatic beta-cell morphology, and MP:0001426—polydipsia. Additionally, our results report significant under-expression of *Mb* (FC = − 3.70, *P*-value = 5×10^{-5}) in diabetic compared to non-diabetic mice. The MP terms for *Mb* gene were muscle fatigue (MP:0003646) and heart left ventricle hypertrophy (MP:0002625). Two additional co-expressed genes in the *Mb* gene network are *Ckm* (MGI:88413) and *Ckmt*2 (MGI:1923972) genes, which were significantly under-expressed (*Ckm*; FC = − 3.18, *P*-value = 5×10^{-5}/*Ckmt*2; FC = − 1.56, *P*-value = 5×10^{-5}) in diabetic mice. *Ckmt*2 expression alterations in diabetic versus non-diabetic liver tissues were significantly involved in the MP terms for heart left ventricle hypertrophy (MP:0002625) and polydipsia (MP:0001426), which are major phenotypes of T2D etiology.
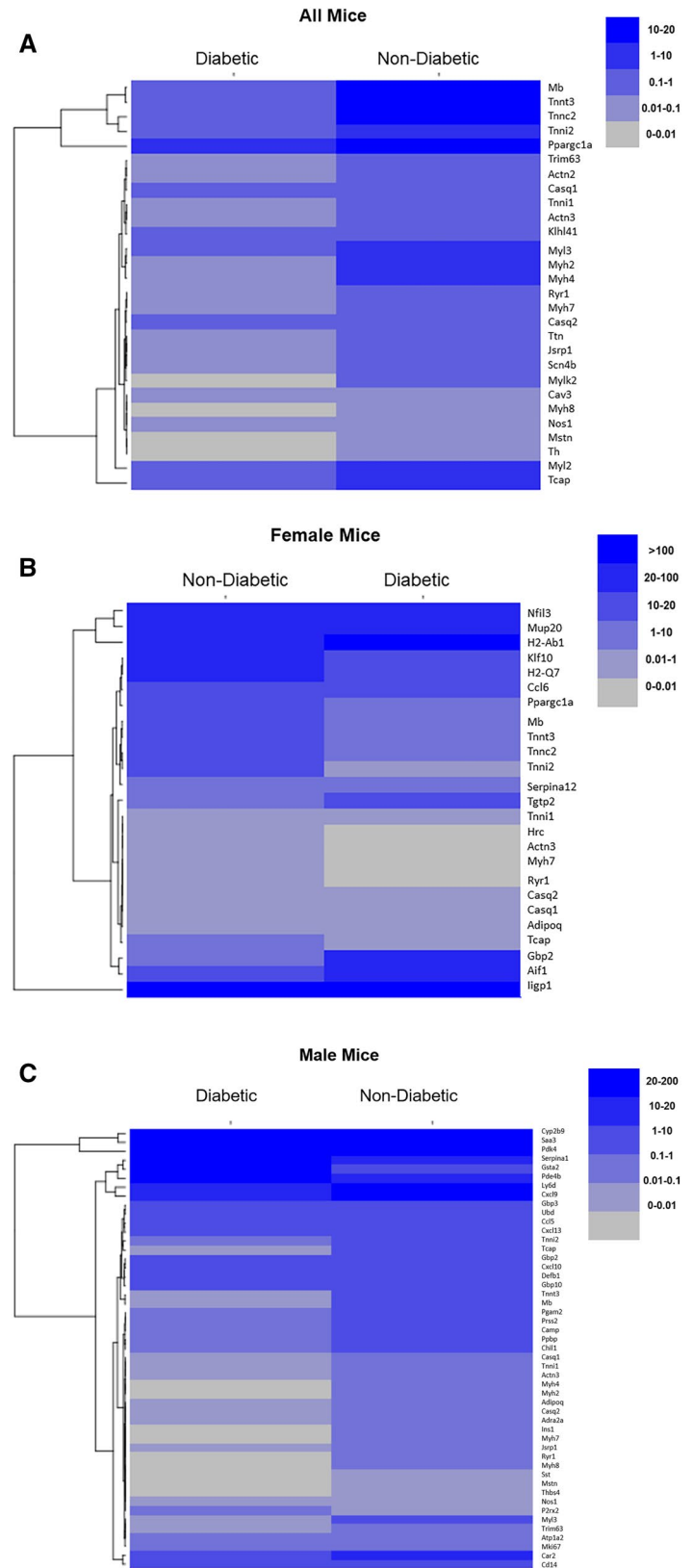
Diabetic versus non-diabetic females

In total, 718 genes were DE (*P*-value < 0.05) between diabetic versus non-diabetic females, of which 178 genes also had |log$_2$ fold change| > 1. A heatmap of the top prioritized genes of the data analysis (fpkm values) is presented in Fig. 2b, suggesting 25 candidate linker genes to be DE.

MGI browser for Mammalian Phenotypes terms revealed 15 significant mammalian phenotype (MP) clusters/terms, including terms of homeostasis/metabolism phenotype (MP:0005376), containing the MP terms of abnormal energy expenditure (MP:0005450), decreased triglyceride level (MP:0005318), and decreased liver triglyceride level (MP:0009356). Moreover, MP terms of T2D and MetS side effects/health complications, for instance MP terms of Abnormal body weight (MP:001259), Muscle fatigue (MP:0003646), Cardiovascular system; Heart left ventricle hypertrophy (MP:0002625), Increased angiogenesis (MP:0005601) and several MP terms of the Immune system. Within the top prioritized genes, the five foremost DE genes were *Hdc*, *Serpina*12, *Socs*1, *Socs*2, and *Mb*. Our results indicate a significant (*P*-value = 3.6×10^{-3}) under-expression (FC = − 1.0019) of the *Hdc* levels in diabetic liver tissues, related to MP terms of abnormal substantia nigra morphology (MP:0000836), heart left ventricle hypertrophy (MP:0002625), increased epididymal fat pad weight (MP:0009288), and abnormal body weight (MP:0001259). Furthermore, a significant under-expression of the *Vaspin* levels observed in diabetic mice (FC = − 1.16, *P*-value = 3.0×10^{-4}) involved in MP phenotypic terms of heart left ventricle hypertrophy (MP:0002625), increased epididymal fat pad weight (MP:0009288), and abnormal body weight (MP:0001259). As for the *Socs1* gene, our data report a significant (*P*-value = 4.5×10^{-3}) over-expression (FC = 1.0060) in diabetic versus non-diabetic livers, to be significantly related to the phenotypic terms of heart left ventricle hypertrophy (MP:0002625), increased circulating interleukin-12b level (MP:0008647), and abnormal body weight (MP:0001259). Our data report a significant (*P*-value = 4×10^{-4}) under-expression (FC = − 1.26) of *Socs*2 levels in diabetic mice significantly related to the same phenotypic terms as of *Socs*1. Furthermore, a significant (*P*-value = 5×10^{-5}) under-expression (FC = − 3.46) of *Mb* gene in diabetic, significantly related to the MP terms of muscle fatigue (MP:0003646), increased angiogenesis (MP:0005601) and heart left ventricle hypertrophy (MP:0002625).

Diabetic versus non-diabetic males

In total, 599 genes were found to be DE (*P*-value < 0.05) between the diabetic versus non-diabetic males, of which 120 genes also had |log$_2$ fold change| > 1. A heatmap is presented in Fig. 2c, suggesting 49 candidate genes to be DE between diabetic and non-diabetic liver tissue. Furthermore, data analysis revealed 11 significant mammalian phenotype (MP) clusters/terms related to T2D and MetS health complications, including terms of different body systems/tissues and terms of increased susceptibility to bacterial infection (MP:0002412) of the immune system. MP term of the

Fig. 2 Heatmaps of hepatic gene expression of the top prioritized genes for the CC mice diabetic versus non-diabetic, in response to 12 weeks of HFD (42% Fat) challenge, for **a** overall population (regardless of sex), **b** females-only group, **c** males-only group. Genes level of expression is expressed in fpkm values (fragments per kilobase of exon per million fragments mapped) presented across diabetic versus non-diabetic samples represented by colors scale. Right *Y*-axis represents associated genes, which includes the top prioritized differentially expressed genes in comparison between diabetic to non-diabetic mice. *X*-axis represents the diabetogenic condition of mice

Table 2 List of top prioritized candidate genes (hub genes) based on BioInfoMiner analysis using Gene Ontology biological process terms, significantly DE in diabetic versus non-diabetic hepatic tissues of overall, females-only, and males-only CC mice population, in response to 12 weeks HFD (42% Fat) challenge, regardless of sex

Group	Gene	Definition	Locus	Fold Change	P-value
Overall	*Lepr*	Leptin receptor	4:101717403-101815352	− 1.06	$1.50e^{-03}$
	*Ins*2	Insulin Π	7:142678655-142699510	1.01	0.034
	Mb	Myoglobin	15:77015488-77057107	− 3.70	$5.00e^{-05}$
	Ckm	Creatine kinase, muscle	7:19411093-19421583	− 3.18	$5.00e^{-05}$
	*Ckmt*2	Creatine kinase, mitochondrial2	13:91853386-91876885	− 1.56	$5.00e^{-05}$
Females only	*Hdc*	Histidine decarboxylase	2:126593666-126619299	− 1.00	$3.60e^{-03}$
	*Serpina*12	Serine (or cysteine) peptidase inhibitor, clade A (alpha-1 antiproteinase, antitrypsin), member 12	12:104028768-104044443	− 1.16	$3.00e^{-04}$
	*Socs*1	Suppressor of cytokine signaling 1	16:10783807-10785536	1.01	$4.50e^{-03}$
	*Socs*2	Suppressor of cytokine signaling 2	10:95385361-95417180	− 1.27	$4.00e^{-04}$
	Mb	Myoglobin	15:77015488-77057107	− 3.46	$5.00e^{-05}$
Males only	*Serpine*1	Serine (or cysteine) peptidase inhibitor, clade E, member 1	5:137061503-137072268	1.15	$5.00e^{-05}$
	Mb	Myoglobin	15:77015488-77057107	− 3.21	0.042
	*Ren*1	Renin 1 structural	1:133350509-133360325	− 1.11	$2.00e^{-04}$
	*Slc*4*a*1	Solute carrier family 4 (anion exchanger), member 1	11:102348823-102366203	− 1.21	$1.00e^{-04}$
	*Atp*2*a*1	ATPase, Ca^{++} transporting, cardiac muscle, fast twitch 1	7:126428758-126463108	− 2.48	$5.85e^{-03}$

For each gene are presented: *Group*—Overall population/females-only or males-only; *Gene Symbol*—short name of the gene; *Definition*—full name of the gene; *Locus*—genomic location of the gene; *Fold Change*– log 2 fold change ratio of gene expression levels between diabetic to non-diabetic conditions; *P-value*—significance level of the fold change level

muscle tissue (abnormal muscle physiology (MP:0002106)) and muscle fatigue (MP:0003646), cardiovascular system (increased angiogenesis (MP:0005601) and cardiac fibrosis (MP:0003141)), renal/urinary system (abnormal urine homeostasis (MP:0009643)) and kidney medulla atrophy (MP:0011422), homeostasis (dehydration (MP:0001429)), increased circulating glycerol level (MP:0003443) and acidosis (MP:0003031). The top ranked five candidate genes were *Serpine*1 *(PAI-1)*, *Mb, Ren*1*, Slc*4*a*1*, and Atp*2*a*1. Data analysis revealed that *PAI*-1 is ranked the first on the candidate genes list, with significantly elevated levels (over-expression; FC = 1.15, *P*-value = 5×10^{-5}) in diabetic, related to four MP terms of increased susceptibility to bacterial infection (MP:0002412), kidney medulla atrophy (MP:0011422), cardiac fibrosis (MP:0003141), and increased angiogenesis (MP:0005601). The hepatic *Mb* gene was significantly under-expressed in diabetic mice (FC = − 3.21, *P*-value = 0.04) and was involved in MP terms of muscle fatigue (MP:0 003646), increased angiogenesis (MP:0005601), and cardiac fibrosis (MP:0003141). Contrariwise, *Ren*1 gene was under-expressed (FC = − 1.11, *P*-value = 2×10^{-4}) in hepatic tissues of diabetic mice and was significantly associated with MP terms of kidney medulla atrophy (MP:0011422) and cardiac fibrosis (MP:0003141). As for the *Slc*4*a*1 gene, a significant under-expression (FC = − 1.21, *P*-value = 1×10^{-4}) was reported in diabetic liver tissues, which is significantly related to MP terms of kidney medulla atrophy (MP:0011422) and

cardiac fibrosis (MP:0003141). Eventually, the *Atp*2*a*1 gene was under-expressed (FC = − 2.48, *P*-value = 5.85×10^{-3}) in diabetic mice tissues and significantly involved in MP terms of muscle fatigue (MP:0003646) and cardiac fibrosis (MP:0003141).

Venn diagrams in Figs. 3a and b present the overlap between DE genes detected by the three tests (all, males-only, and females-only) for *p*-value < 0.05 DE genes (A) and for *p*-value < 0.05 & |log$_2$FC| > 1 DE genes (B).

Females versus males in diabetic and non-diabetic conditions

In this part of the study, we focus on baseline sex differences between females and males during healthy and diabetic conditions, while sharing the same environmental conditions. All the tests performed here are for sex differences, comparing females and males of the same condition (diabetic only or non-diabetic only), by merge of all the alignments for the female mice and all the alignments for the male mice, and then compared those by running the Cuffdiff analysis.

Non-diabetic females versus males

For the non-diabetic mice, 193 genes were DE (|log2 fold change| > 1) in males versus females' hepatic tissues (*P*-value < 0.05). Figure 4a presents a heatmap of the hepatic gene expression (fpkm values) of 36 top prioritized

A

B
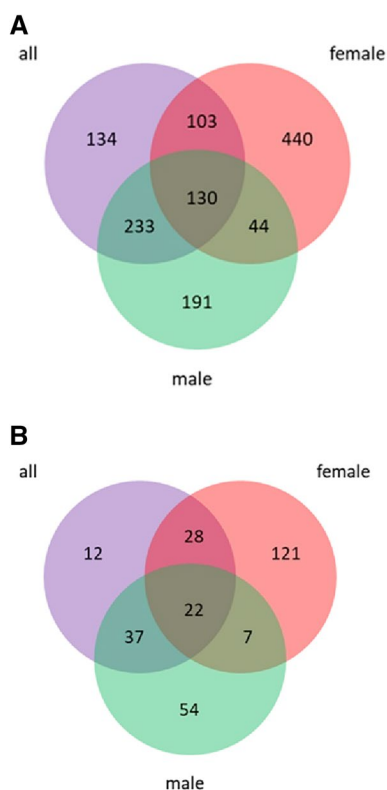
Fig. 3 Venn diagrams presenting the overlaps between DE genes of the three tests of All, females-only and males-only groups for diabetic versus non-diabetic comparison. **a** Venn diagram for p-value < 0.05 DE genes and **b** Venn diagram for p-value < 0.05 & |log2FC| > 1 DE genes

candidate genes. MGI browser for Mammalian phenotypes, reported 14 mammalian phenotype (MP) clusters/terms significantly related (directly/indirectly) to T2D health complications, of which 7 branches were of the abnormal homeostasis term (MP:0001764), including abnormal energy expenditure (MP:0005450), terms of abnormal lipid homeostasis (MP:0002118), and abnormal glucose homeostasis (MP:0002078). As well terms of abnormal adipose tissue morphology (MP:0000003) and abnormal cell physiology (MP:0005621), containing the term of increased pancreatic islet cell apoptosis (MP:0014121). For each of the DE top prioritized genes (|logFC| > 1) between non-diabetic females and males ,the FC—fold change level and P-value are reported, and the direction of FC is under/over-expressed in males relatively to females. Top five prioritized genes were, as ranked from top to bottom, the *Lepr, Cav*1, *Socs*2, *Abcg*2, *and Col*5*a*3. In this context of sex-specific differences, our data represent baseline (non-diabetic) significant under-expression (FC = − 1.55; P-value < 0.05) of *Lepr* gene in liver tissues of males versus females to be significantly associated with MP terms of increased pancreatic islet cell apoptosis (MP:0014121), abnormal

respiratory mechanics (MP:0002314), enhanced lipolysis (MP:0008034), and abnormal respiratory electron transport chain (MP:0010955). And, conversely, over-expression (FC = 1.41) of the hepatic *Cav*1 gene in non-diabetic males compared to females is significantly linked to MP terms of abnormal respiratory mechanics (MP:0002314), enhanced lipolysis (MP:0008034), increased bone trabecula number (MP:0010868), and abnormal respiratory electron transport chain (MP:0010955). Third in rank, the hepatic *Socs*2 gene was under-expressed (FC = − 1.19) in males versus females, suggesting a baseline sex-specific expression variations. *Socs*2 expression variations were significantly associated with MP terms of abnormal respiratory mechanics (MP:0002314), enhanced lipolysis (MP:0008034), and increased bone trabecula number (MP:0010868). Next, fourth in rank, the hepatic *Abcg*2 (FC = 1.30) gene over-expression in males versus females was associated with MP terms of abnormal respiratory electron transport chain (MP:0010955) and enhanced lipolysis (MP:0008034). The *Col*5*a*3 gene, ranked fifth from the top, was over-expressed (FC = 1.54) in males versus females and significantly involved in MP terms of increased pancreatic islet cell apoptosis (MP:0014121) and enhanced lipolysis (MP:0008034).

Diabetic females versus males

For the diabetic mice, 389 genes were DE (|log2 fold change| > 1) between diabetic females versus diabetic males (P-value < 0.05). Figure 4b presents a heatmap of the top prioritized genes, suggesting 68 DE top prioritized candidate genes. MGI browser for Mammalian phenotypes reported 23 mammalian phenotype (MP) clusters/terms to be significantly associated with the candidate genes, while majority of the MP terms related directly or indirectly to T2D health complications. The larger hierarchical MP term family belonged to the immune system phenotype (MP:0005387), followed by families of abnormal lipid homeostasis (MP:0002118), endocrine/exocrine gland phenotype (MP:0005379) containing two MP terms of abnormal pancreatic islets, growth/size/body region phenotype (MP:0005378), and increased food intake (MP:0011939). The top five prioritized genes, as ranked from top to bottom, were the *Lepr, Clps, Ins*2, *Cav*1, *and Mrap*2 *gene*. The *Lepr* gene, ranked on top of the candidate genes lists, is markedly maintained in the groups of "*All mice,*" "*Non-diabetic females versus males,*" and in the current list. Here, *Lepr* levels were under-expressed (FC = − 1.41) in males versus females. While, beside the MP term of increased pancreatic islet cell apoptosis (MP:0014121), *Lepr* sexual dimorphism in diabetic condition exists, suggesting MP terms of decreased susceptibility to endotoxin shock (MP:0008734), decreased circulating triglyceride level (MP:0002644), increased circulating HDL cholesterol level

Fig. 4 Heatmaps of hepatic gene expression for the sex differences top prioritized genes in response to 12 weeks of HFD (42% fat) challenge, at **a** baseline condition (non-diabetic) and **b** disease (diabetes). Genes level of expression are expressed in fpkm values (fragments per kilobase of exon per million fragments mapped) presented across males versus females hepatic tissue represented by colors scale. Right Y-axis represents associated genes, which includes the top prioritized differentially expressed genes in comparison between males and females. X-axis represents the mice sex
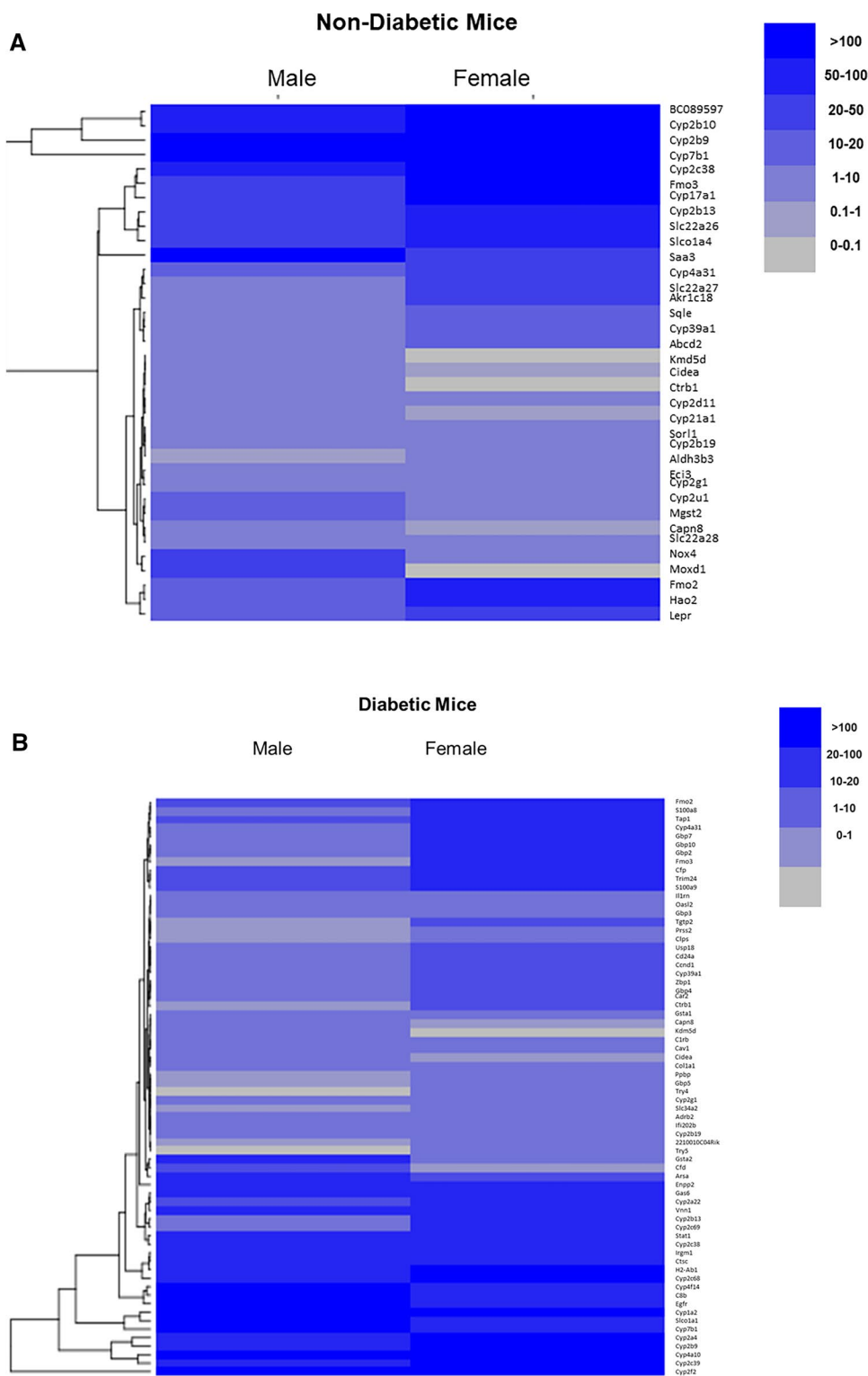
(MP:0001556), and increased compensatory feeding amount (MP:0011958). Second in rank, the *Clps gene,* under-expressed (FC = −3.02) in diabetic males versus females, is associated with decreased susceptibility to endotoxin shock (MP:0008734), sparse hair (MP:0000416), decreased circulating triglyceride level (MP:0002644), and increased compensatory feeding amount (MP:0011958). Following, third in rank, our findings report hepatic *Ins*2 under-expression (FC = −2.53) in diabetic males compared to females, to be related to MP terms of increased pancreatic islet cell

apoptosis (MP:0014121), increased vascular permeability (MP:0003070), decreased circulating triglyceride level (MP:0002644), and increased compensatory feeding amount (MP:0011958). Next, fourth in rank is the *Cav*1 over-expression in males versus females (FC = 1.43), similar to the result in the previous section of "non-diabetic females versus males," but with different MP terms here. The suggested MP terms for *Cav*1 over-expression, here, were decreased susceptibility to endotoxin shock (MP:0008734), increased vascular permeability (MP:0003070), abnormal body water content (MP:0000676), and increased compensatory feeding amount (MP:0011958). Fifth in rank, the hepatic *Mrap*2 gene over-expression (FC = 1.54) in males versus females is associated with MP terms of decreased susceptibility to endotoxin shock (MP:0008734), increased circulating HDL cholesterol level (as in "All mice"; MP:0001556), and increased compensatory feeding amount (MP:0011958).
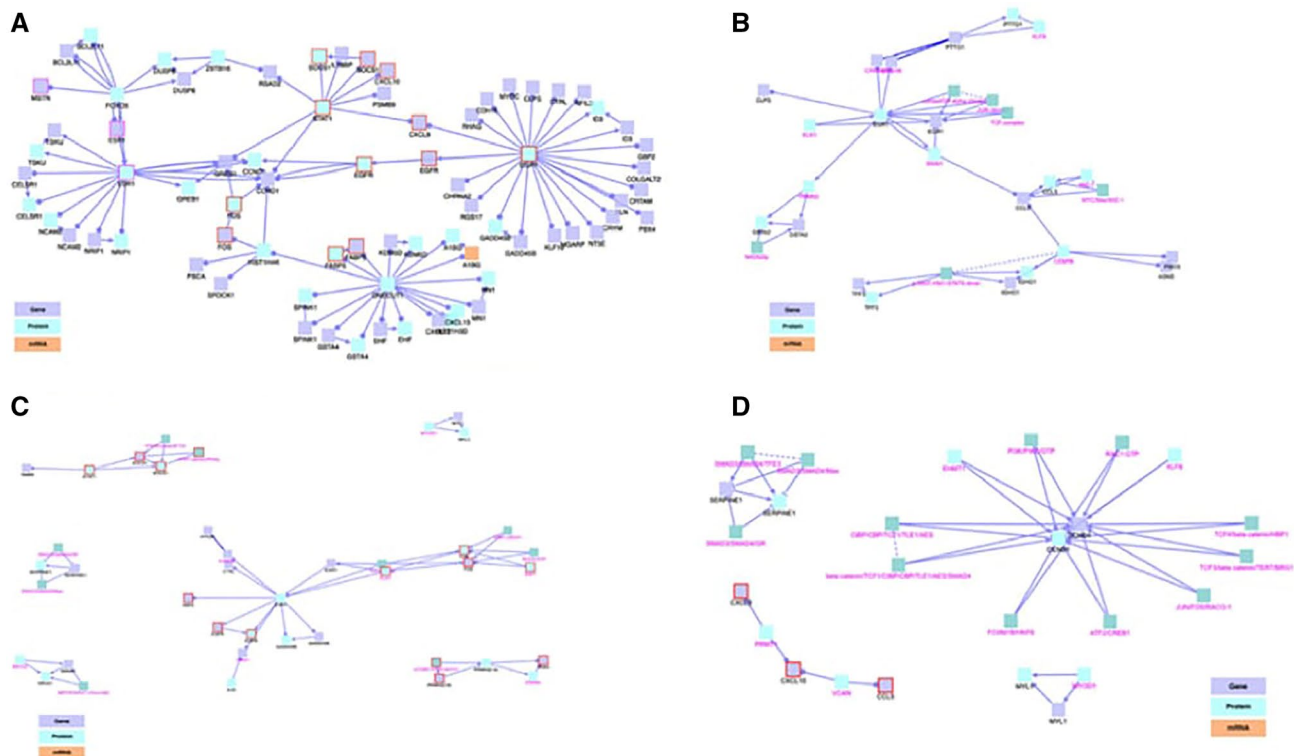
Gene regulatory networks

A gene regulatory network was obtained from the DE genes in normal vs. diabetic mice (*p*-value < 0.05), including genes from the three group of all mice, females-only, and males-only as shown in Figs. 5a–d. The *Onecut1* gene that

plays a prominent role in this network, also known as *HNF6* (hepatocyte nuclear factor 6), encodes a member of the Cut homeobox family pf transcription factors. Its expression is enriched in the liver, where it antagonizes glucocorticoid-stimulated gene transcription. The *Onecut1* gene is potentially involved in several biological processes, including glucose metabolism and cell cycle regulation. Additionally, among the significant genes, we have identified the ones with key, prominent, and regulatory role in the network that potentially plays a major part in the manifestation of T2D.

As for the results from the comparison between males and females, from the generated networks presented in Fig. 6a and b, it is obvious that different key players seem to be involved in T2D manifestation in the two sexes, with *Ppargc*1*a* and *Ccnd*1 acting as key regulators in female and male mice, respectively. Both genes have been previously linked to obesity-associated T2D (Shen and Zhu 2018).

## Gene expression data integration with published QTL data

Integration of the gene expression data with previously reported QTLs of glucose tolerance female-specific QTLs in collaborative cross mice in regions:

Fig. 5 Gene regulatory network obtained from Consensus Path DB (Kamburov et al. 2012) from differentially expressed genes in normal versus diabetic mice (*p* value < 0.05 & |logFC| > 1), **a** including genes from ANvsAD, FNvsFD and MNvsMD. **b** Gene regulatory network from ANvsAD. **c** Gene regulatory network of Female group, from FNvsFD. **d** Gene regulatory network of Male group, from MNvsMD. Highlighted in red are genes prioritized by BioInfoMiner

Fig. 6 Gene regulatory network obtained from Consensus Path DB from differentially expressed genes in **a** diabetic group, female versus male mice (*p* value < 0.05 and |logFC| > 1), **b** non-diabetic group, female versus male mice (*p* value < 0.05 and |logFC| > 1). Highlighted in red are genes prioritized by BioInfoMiner
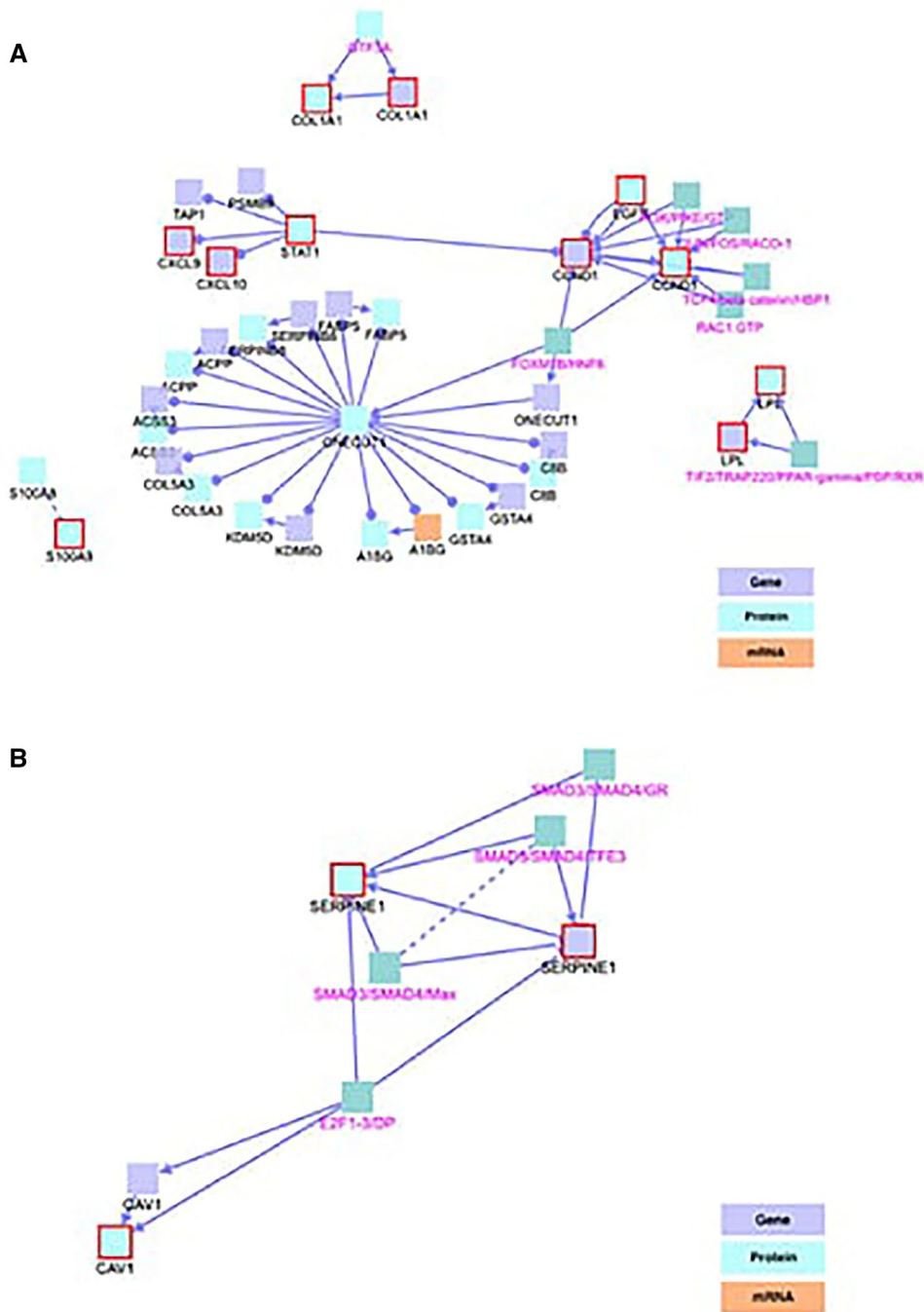
8:32–34.5 Mb (Abu-Toamih Atmni et al. 2017) and liver fattiness in regions 4:35–37 Mb, 17:8–11 Mb, and 18:64–66 Mb (Abu-Toamih Atamni et al. 2016a) revealed no significant variations. Variations were identified in the region 8:32–34.5 Mb, including 2 genes (*Gm6877, 5930422O12Rik*), but these were not significant (female group). Also, in chromosomal region 18:64–66 Mb, for *Alpk2* gene, it was not significant (All mice, female groups); for *Grp* gene, it was not significant (male group). In region 4:35–37 Mb, for *3110043O21Rik* gene, it was

not significant (male group). Additional glucose tolerance sites tested 9:36–49 Mb, 11:62–71 Mb, and 12:79–99 Mb (Clee and Attie 2007). For 9:36–49 Mb, variations were identified in *Gm26249* (logFC = − 2.09, pv = 0.0135, FDR = 0.86) for all mice group and (logFC = − 2.00, pv = 0.004, FDR = 0.9) male-only group, this group also had variations in *Sorl1* with logFC = 1.18, pv = 0.0072, FDR = 0.6. *Scn4b, Crtam, Tmprss, and Zbtb19* were not significant. For 11:62–71 Mb, variations were identified in *Myh1* (logFC = − 1.226, pv = 0.00005, FDR = 0.012)

for all mice group, (logFC = − 1.26, pv = 0.00005, FDR = 0.013) for females, not significant for male group but with logFC = − 0.85, pv = 0.0003, FDR = 0.06. In male's group, *Adcy1* was significantly altered with logFC = − 1.97, pv = 0.00005, FDR = 0.013. *Myh2, Myh4, Myh8, Myh13, 2310065F04Rik, and Dhrs7c* were not significant. For 12:79–99 Mb, *Acot3* gene had (logFC = − 1.44, pv = 0.0007, FDR = 0.1) in all mice group, (logFC = − 2.05, pv = 0.00015, FDR = 0.036) in male group, and *Fos* (logFC = 1.26, pv = 0.0003, FDR = 0.056) in female group. *Acot2, Plekhd1, Acot5, Syndig1 l, and Kcnk10* were not significantly altered.

Discussion

The collaborative cross (CC) is a next-generation mouse genetic reference population (GRP), developed by appreciable efforts of the complex trait community (CTC). The main aim behind the development of the CC multi-parental mouse model was to multiply the genomic mapping resolution of complex traits and complex diseases' etiology of relevance to human health. Herein, we demonstrate for the first time the utilization of this diverse genetic reference population for the study of diet-induced T2D health complications at the level of hepatic gene expression. Next-generation RNA-sequencing was performed to assess the hepatic gene expression profile in health (non-diabetic CC lines) versus disease (diabetic CC lines). Moreover, hepatic gene expression profile of non-diabetic females versus males was assessed in order to track candidate genes responsible for sex differences at baseline state (health), and thereafter, a similar analysis was performed for females versus males in the disease state (diabetic CC lines).

Under-expression of *Lepr* in diabetic condition

Under-expression of *Lepr* in diabetic condition may indicate liver resistance to *Leptin* hormone, which is an abnormal condition leading to impaired regulation of triglyceride metabolism, impaired hepatic glucose homeostasis, and impaired hepatic insulin sensitivity (Do et al. 2016). Leptin function is complex and known to play central regulatory roles in a wide range of systems including regulation of energy homeostasis, satiety (food intake), lipid, and glucose metabolism, alongside the activation of the immune system. Leptin hormone is a well-known candidate in the study of diabetes, obesity, and metabolic syndrome features. Our study's results give support and evidence for the complex roles of leptin receptor far beyond mediating satiety

and suggest it as major player of the hepatic tissue in the pathogenesis of T2D (Joost and Schürmann 2014).

Hepatic compensatory mode under stress

Our findings present an interesting phenomenon of extra-pancreatic expression of *Ins*2, i.e., insulin expression in the liver tissue with significantly higher levels in diabetic mice. A similar phenomenon was previously described in hyperglycemic high-fat diet-fed mice by Kojima et al. (2004) and also Chen et al. (2010), which is possibly one of the compensatory mechanisms of the liver, in response to impaired glucose tolerance and insulin resistance. Consistent with this, the pancreatic *Reg*1 gene expression levels were significantly elevated in the liver tissue of diabetic mice in our study, manifesting hepatic cells expressing pancreatic factors (heterotopic phenotype). A particular situation of hepatic to pancreatic switch occurs nearby injured liver microenvironment or in a compensatory mode under stress conditions (Parikh et al. 2012). Data mining for *Ins*2 expression levels in all the groups revealed that only females express *Ins*2 in the liver. Focusing on the females-only group, *Ins*2 expression during health was very slight and closer to zero, supporting the aforementioned hypothesis of compensatory mechanisms of the liver during diabetes, which are apparently specific for females. And indeed, referring to the literature, we realize that normally mice and rats express two non-allelic insulin genes, *Ins*1 and *Ins*2, located on different chromosomes (19 and 7, respectively), mediating different metabolic pathways in female and males (Templeman et al. 2016). It is important to mention that peripheral interactions of leptin–Insulin pathways are known to play a central role in energy homeostasis and therefore are investigated as therapeutic agents for diabetes.

DE genes of hypoxia/ischemia and tissue injury phenotypes

At the level of hypoxia/ischemia and tissue injury phenotypes during diabetes, and with regard to the T2D co-morbidity with NAFLD, three candidate co-expressed genes were suggested in our study; *Mb (Myoglobin), Ckm,* and *Ckm*2, as significantly under-expressed in diabetic hepatic tissues. Study of *Myoglobin* deficiency in $myo^{-/-}$ mice showed adaptation of the mice to *Mb* deficiency via compensatory mechanisms, and yet, the results are highly variable between mouse strains suggesting a complex regulatory gene network of *Mb* to be controlled by the genetic host background (Schlieper et al. 2004). Myoglobin expression in the liver was suggested to play a protective role against oxidative stress effects by enhancing ATP levels and oxidization ratio (Nitta et al. 2003). Therefore, our findings may indicate

decreased oxidization processes leading to increased risk for liver injury.

DE genes of the inflammatory system

Another major system involved in the basis of T2D and metabolic syndrome pathogenesis is the inflammatory system, which is proposed to be a mediator between insulin resistance and the various metabolic interruptions. One of the candidate genes for females group was the *Hdc* gene (*histidine decarboxylase)* of the histaminergic system, which is involved in multiple body systems and tissues including the immune system and metabolism homeostasis. Indeed, the Fülöp et al. (2003) study, using homozygous *Hdc* knock-out mice, has reported multi-systemic abnormalities, including obesity, impaired glucose tolerance, and hyperleptinemia, while in our study the *Hdc* was under-expressed in diabetic females. Based on studies and on our findings, we may recommend the *Hdc gene* investigation as potential treatment for diabetes health complications (Fülöp et al. 2003). Additional candidate genes of the inflammatory system of the females involve two members of the SOCS (suppressors of cytokine signaling) proteins family genes, *Socs1* (over-expressed) *and Socs2* (under-expressed) in diabetic hepatic tissues. *Socs*1 is an essential negative regulator for a wide range of pathways and systems including homeostasis/metabolism and the immune system. Over-expression of *Socs*1 in the liver is stimulated by increased levels of the inflammatory factors, IL-6 and IFNγ, i.e., mediated via inflammatory factors, implies the strong linkage between inflammation and diabetes (Lebrun and Van Obberghen 2008). Studies of mouse models for *Socs1* deficiency have reported multi-system abnormalities, including inflammation of the liver and pancreas, as well as hepatomegaly and hepatic steatosis (Zhang et al. 2001). Furthermore, constitutive expression of *Socs1* in the liver induces diabetes, metabolic syndrome features, and hepatic steatosis (Ueki et al. 2004). As for the *Socs2* gene, which plays an essential role in hepatic regeneration alongside its role as negative regulator of growth (Masuzaki et al. 2016), hepatic expression levels were reduced in diabetic high-fat-fed females' tissues in our study. Previously, it was shown that high-fat-fed *SOCS2⁻/⁻* mice had higher susceptibility (vs. wild type) to diet in terms of metabolic syndrome features, with elevated levels of inflammatory cytokines (Zadjali et al. 2012). Thence, targeting *Socs*1 suppression and/or *Socs*2 induction of expression in females will be an interesting strategy to investigate for the treatment of diabetes. As for the males' inflammatory system, two central candidate genes are proposed, the *Serpine*1 and *Slc4a1* genes, over- and under-expressed, respectively, in diabetic hepatic tissues of males. The *Serpine*1 named also *PAI*-1 plays a key role as inhibitor of fibrinolysis and important roles in inflammation and wound healing. Over-expression of *PAI*-1

has been related to diabetes and metabolic syndrome features, while hepatic over-expression of *PAI*-1 has been co-associated with NAFLD severity/type and insulin resistance (Jin et al. 2018), which implies the potential targeting of hepatic *PAI-1* as a possible marker for diabetic/metabolic syndrome progress. Interestingly, in this context, a study of human hepatic gene expression in health versus diabetes has reported down-regulation of *plasminogen* in diabetic subjects (Takamura et al. 2004). A recent study of Chan et al. (2015) reported significant linkage between *Slc4a1* variants and increased risk for diabetes.

Hepatic sexual dimorphism in health and disease conditions

The most dominant MP terms within the ontological tree of sex differences were associated with the immune system. Sexual dimorphism in health and disease is highly relevant for the pharmacogenetics and pharmacokinetics of drugs for the treatment of T2D, having a direct impact on the host responsiveness to the treatment. Our study results reveal broad presentation of the cytochromes P450 (*CYP*2) superfamily proteins (enzymes) which are known to metabolize numerous endogenous and exogenous chemicals/xenobiotic/steroids/fatty acids, including drugs metabolism. This means that females to males' differences are seriously considerable for the treatment of diabetes, in this case, and any other disease in general (based on the naïve group). Genes of the *Cyp*2 subfamily: *Cyb2b9*, *Cyb2g1*, and *Cyb2b13*, which are responsible for drug and steroid metabolism, were under-expressed in males versus females during health, and much more reduced during diabetes. This was in agreement with our previous study (Nachshon et al. 2016). Indeed, the sexual dimorphism of hepatic CYPs expression is a well-known phenomenon, which is suggested to be regulated by a complex regulatory network and need to be further investigated (Wiwi et al. 2004). A notable change is the additional reduction in *Cyp2b13* in health versus disease. Based on a recent study using the *Cyp2b9/10/13-null* mouse model, *CypB* hepatic expression is lower in males, which might be mediated by the *constitutive androstane receptor* (CAR) (Kumar et al. 2017). An additional gene with greater FC in diabetic condition is the *Cyp7b1* gene, 7-alpha hydroxylase of steroid nucleus, playing a role in bile acid biosynthesis. This gene was over-expressed in females versus males in health state and even more during diabetes, a similar sexual dimorphism was recently described by Xie et al. in STZ-HFD-induced NASH-HCC mouse model with regard to the study of the cross-talk of the bile acids and microbiota in the gut–liver axis. Over-expression of *Cyp7b1* gene may result in accumulation of cytotoxic bile acids in the liver and eventually increased risk for hepatic cancer in males (Xie et al. 2017). Moreover, another important sexual dimorphic gene

of xenobiotic, lipid, and carbohydrate metabolism roles is the *Fmo*3 gene, member of the *Flavin-containing monooxygenases* (FMOs) enzymes family. *Fmo*3 levels were consistently higher in females than males for both conditions of health and diabetic, with significantly reduced levels in diabetic mice. Similar findings were previously described in several mouse strains (CD-1, Swiss-Webster, C57BL/6, and DBA/2 mice) (Falls et al. 1995). In our study, elevated *Fm*o3 hepatic expression was related to protective effects against CVD and atherosclerosis, in addition to improved lifespan (aging) in mice with higher levels of hepatic *Fmo3* and *Fmo*4 (Rossner et al. 2017). These findings indicate higher risk to a greater distinct in males for T2D co-morbidities with possible impairment of lifespan. As previously and successfully implemented in our lab in translation from mouse to human (Shusterman et al. 2017; Nashef et al. 2018), we will use similar approaches for the translation of our finding here to human. Future studies will be focused on assessing gene expression findings using in vitro tissue culture analysis, which will minimize the use of live animals.

Compliance with ethical standards

Conflict of interests There are no competing financial interests in relation to the work described by all authors, including e/NIOS Company.

References

Abu-Toamih Atamni HJ, Botzman M, Mott R, Gat-Viks I, Iraqi FA (2016a) Mapping liver fat female-dependent quantitative trait loci in collaborative cross mice. Mamm Genome 27(11–12):565–573

Abu-Toamih Atamni HJ, Mott R, Soller M, Iraqi FA (2016b) High-fat-diet induced development of increased fasting glucose levels and impaired response to intraperitoneal glucose challenge in the collaborative cross mouse genetic reference population. BMC Genet 17:10

Abu-Toamih Atamni HJ, Ziner Y, Mott R, Wolf L, Iraqi FA (2017) Glucose tolerance female-specific QTL mapped in collaborative cross mice. Mamm Genome 28(1–2):20–30

Abu-Toamih Atamni HJ, Botzman M, Mott R, Gat-Viks I, Iraqi FA (2018) Mapping novel genetic loci associated with female liver weight variations using collaborative cross mice. Anim Model Exp Med. 1(3):212–220

Ansorge WJ (2009) Next-generation DNA sequencing techniques. N Biotechnol 25(4):195–203

Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT et al (2000) Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat Genet 25:25–29

Chan KH, Chacko SA, Song Y, Cho M, Eaton CB, Wu WC (2015) Genetic variations in magnesium-related ion channels may affect diabetes risk among African American and Hispanic American women. J Nutr 145(3):418–424

Chen X, Larson CS, West J, Zhang X, Kaufman DB (2010) In vivo detection of extrapancreatic insulin gene expression in diabetic mice by bioluminescence imaging. PLoS ONE 5(2):e9397

Clayton JA, Collins FS (2014) Policy: NIH to balance sex in cell and animal studies. Nature 509:282–283

Clee SM, Attie AD (2007) The genetic landscape of type 2 diabetes in mice. Endocr Rev 28(1):48–83 **Review**

Do OH, Gunton JE, Gaisano HY, Thorn P (2016) Changes in beta cell function occur in prediabetes and early disease in the Lepr (db) mouse model of diabetes. Diabetologia 59(6):1222–1230

Falls JG, Blake BL, Cao Y, Levi PE, Hodgson E (1995) Gender differences in hepatic expression of flavin-containing monooxygenase isoforms (FMO1, FMO3, and FMO5) in mice. J Biochem Toxicol 10(3):171–177

Fülöp AK, Földes A, Buzás E, Hegyi K, Miklós IH, Romics L (2003) Hyperleptinemia, visceral adiposity, and decreased glucose tolerance in mice with a targeted disruption of the histidine decarboxylase gene. Endocrinology 144(10):4306–4314

Gale EA, Gillespie KM (2001) Diabetes and gender. Diabetologia 44(1):3–15 **Review**

Iraqi FA, Mahajne M, Salaymah A et al (2012) The Genome architecture of the collaborative cross mouse genetic reference population. Genetics 190(2):389–402

Jin R, Krasinskas A, Le NA, Konomi JV, Holzberg J, Romero R (2018) Association between plasminogen activator inhibitor-1 and severity of liver injury and cardiovascular risk in children with non-alcoholic fatty liver disease. Pediatr Obes 13(1):23–29

Joost HG, Schürmann A (2014) The genetic basis of obesity-associated type 2 diabetes (diabesity) in polygenic mouse models. Mamm Genome 25(9–10):401–412

Kamburov A, Stelzl U, Lehrach H, Herwig R (2012) The consensus path DB interaction database: 2013 update. Nucleic Acids Res 41:D793–D800

Kaul N, Ali S (2016) Genes, genetics, and environment in type 2 diabetes: implication in personalized medicine. DNA Cell Biol 35(1):1–12 **Review**

Kojima H, Fujimiya M, Matsumura K, Nakahara T, Hara M, Chan L (2004) Extrapancreatic insulin-producing cells in multiple organs in diabetes. Proc Natl Acad Sci USA 101(8):2458–2463

Kontogianni, G., Papadodima, O., Mitrakas, A., et al. (2017) An RNA-seq analysis from non-small cell lung cancer biopsies suggests an important role for aberrant alternative splicing in its pathophysiology. Health and Technology, 1-8

Kumar R, Mota LC, Litoff EJ et al (2017) Compensatory changes in CYP expression in three different toxicology mouse models: CAR-null, Cyp3a-null, and Cyp2b9/10/13-null mice. PLoS ONE 12(3):e0174355

Lebrun P, Van Obberghen E (2008) SOCS proteins causing trouble in insulin action. Acta Physiol (Oxf) 192(1):29–36 **Review**

Legato MJ, Gelzer A, Goland R, Ebner SA, Rajan S, Villagra V (2006) Gender-specific care of the patient with diabetes: review and recommendations. Gend Med 3(2):131–158 **Review**

Masuzaki R, Zhao S, Valerius MT, Tsugawa D, Oya Y, Ray KC (2016) SOCS2 Balances metabolic and restorative requirements during liver regeneration. J Biol Chem 291(7):3346–3358

Mathers CD, Loncar D (2006) Projections of global mortality and burden of disease from 2002 to 2030. PLoS Med 3(11):e442

Nachshon A, Abu-Toamih Atamni HJ, Steuerman Y et al (2016) Dissecting the effect of genetic variation on the hepatic expression of drug disposition genes across the collaborative cross mouse strains. Front Genet 7:172

Nashef A, Abu-Toamih Atamni HJ, Buchnik Y et al (2017) Collaborative cross mouse population for studying alveolar bone changes and impaired glucose tolerance comorbidity after high-fat diet consumption. J Periodontol 88(9):e150–e158

Nashef A, Qabaja R, Salaymeh Y et al (2018) Integration of murine and human studies for mapping periodontitis susceptibility. J Dent Res 97(5):537–546

Ng SB, Buckingham KJ, Lee C et al (2010) Exome sequencing identifies the cause of a Mendelian disorder. Nat Genet 42(1):30–35

Nitta T, Xundi X, Hatano E et al (2003) Myoglobin gene expression attenuates hepatic ischemia reperfusion injury. J Surg Res 110(2):322–331

Parikh A, Stephan AF, Tzanakakis ES (2012) Regenerating proteins and their expression, regulation and signaling. Biomol Concepts 3(1):57–70

Rossner R, Kaeberlein M, Leiser SF (2017) Flavin-containing monooxygenases in aging and disease: emerging roles for ancient enzymes. J Biol Chem 292(27):11138–11146 **Review**

Samsom M, Trivedi T, Orekoya O, Vyas S (2016) Understanding the importance of gene and environment in the etiology and prevention of type 2 diabetes mellitus in high-risk populations. Oral Health Case Rep 2(1):112

Schlieper G, Kim JH, Molojavyi A et al (2004) Adaptation of the myoglobin knockout mouse to hypoxic stress. Am J Physiol Regul Integr Comp Physiol 286(4):R786–R792

Shen J, Zhu B (2018) Integrated analysis of the gene expression profile and DNA methylation profile of obese patients with type 2 diabetes. Mol Med Rep 17:7636–7644

Shusterman A, Munz M, Richter G et al (2017) The PF4/PPBP/CXCL5 gene cluster is associated with periodontitis. J Dent Res 96(8):945–952

Smith CL, Eppig JT (2009) The mammalian phenotype ontology: enabling robust annotation and comparative analysis. Wiley Interdiscip Rev 1:390–399

Takamura T, Sakurai M, Ota T, Ando H, Honda M, Kaneko S (2004) Genes for systemic vascular complications are differentially expressed in the livers of type 2 diabetic patients. Diabetologia 47(4):638–647

Templeman NM, Mehran AE, Johnson JD (2016) Hyper-variability in circulating insulin, high fat feeding outcomes, and effects of reducing Ins2 dosage in male Ins1-null mice in a specific pathogen-free facility. PLoS ONE 11(4):e0153280

Trapnell C, Roberts A, Goff L, Pertea G, Kim D, Kelley DR (2012) Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nat Protoc 7(3):562–578

Ueki K, Kondo T, Tseng YH, Kahn CR (2004) Central role of suppressors of cytokine signaling proteins in hepatic steatosis, insulin resistance, and the metabolic syndrome in the mouse. Proc Natl Acad Sci USA 101(28):10422–10427

Voelkerding KV, Dames SA, Durtschi JD (2009) Next-generation sequencing: from basic research to diagnostics. Clin Chem 55(4):641–658 **Review**

Wiwi CA, Gupte M, Waxman DJ (2004) Sexually dimorphic P450 gene expression in liver-specific hepatocyte nuclear factor 4alpha-deficient mice. Mol Endocrinol 18(8):1975–1987

World Health Organization (2014) Global status report on noncommunicable diseases 2014. WHO. http://www.who.int/nmh/publications/ncd-status-report-2014/en/. Accessed 25 Nov 2015

Xie G, Wang X, Zhao A et al (2017) Sex-dependent effects on gut microbiota regulate hepatic carcinogenic outcomes. Sci Rep 7:45232

Yang X, Schadt EE, Wang S, Wang H, Arnold AP, Ingram-Drake L (2006) Tissue-specific expression and regulation of sexually dimorphic genes in mice. Genome Res 16(8):995–1004

Zadjali F, Santana-Farre R, Vesterlund M, Carow B, Mirecki-Garrido M, Hernandez-Hernandez I (2012) SOCS2 deletion protects against hepatic steatosis but worsens insulin resistance in high-fat-diet-fed mice. FASEB J 26(8):3282–3291

Zhang JG, Metcalf D, Rakar S, Asimakis M, Greenhalgh CJ, Willson TA (2001) The SOCS box of suppressor of cytokine signaling-1 is important for inhibition of cytokine action in vivo. Proc Natl Acad Sci USA 98(23):13261–13265