Genomics Data 2 (2014) 351–353

Contents lists available at ScienceDirect

Genomics Data

journal homepage: http://www.journals.elsevier.com/genomics-data/

Data in Brief

Gene expression in rat models for inter-generational transmission of islet dysfunction and obesity

Ruby C.Y. Lin [a,b], Sheau-Fang Ng [c,d], Margaret J. Morris [d,*]

[a] Asbestos Diseases Research Institute, NSW, 2139 Sydney, Australia
[b] School of Biotechnology and Biomolecular Sciences, University of New South Wales, NSW 2052, Australia
[c] Department of Paediatrics, Faculty of Medicine, Universiti Kebangsaan Malaysia Medical Center, Universiti Kebangsaan Malaysia, 56000 Kuala Lumpur, Malaysia
[d] Department of Pharmacology, School of Medical Sciences, University New South Wales, NSW 2052, Australia

ARTICLE INFO

Article history:
Received 8 September 2014
Accepted 29 September 2014
Available online 16 October 2014

Keywords:
Intergenerational transmission
Islet dysfunction
Paternal transmission
High fat diet
Gene signatures

ABSTRACT

Paternal high fat diet (HFD) consumption triggers unique gene signatures, consistent with premature aging and chronic degenerative disorders, in both white adipose tissue (RpWAT) and pancreatic islets of daughters. In addition to published data in Nature, 2010, 467, 963–966 (GSE: 19877, islet) and FASEB J 2014, 28, 1830–1841 (GSE: 33551, RpWAT), we describe here additional details on systems-based approaches and analysis to develop our observations. Our data provides a resource for exploring the complex molecular mechanisms that underlie intergenerational transmission of obesity.

Specifications	
Organism/cell line/tissue	Sprague–Dawley rat
Sex	Female
Sequencer or array type	Affymetrix Rat GeneChip Gene 1.0 ST arrays (Affymetrix)
Data format	Quantile Normalized Log2
Experimental factors	Islet and white adipose tissue (RpWAT) harvested from F1 daughters from F0 fathers fed on high fat diet (HFD) vs F1 daughters from F0 fathers fed on normal diet
Experimental features	Genome-wide expression analysis comparing islet and RpWAT in F1 daughters from F0 fathers fed on either HFD or normal diet.
Consent	N/A
Sample source location	Sydney, Australia

Direct link to deposited data

Effects of dietary obesity in fathers on gene expression of white adipose tissue in the female offspring, http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE33551.

Effects of dietary obesity in fathers on gene expression of islets in the female offspring, http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE19877.

Experimental design, materials and methods

Male Sprague–Dawley rats were fed high fat diet (HFD) for 11 weeks prior to mating with females consuming control diet [1]. Females consumed regular chow prior to mating and throughout gestation and lactation. Offspring were weaned at 21 days onto control diet, and were killed at 14 weeks of age after an overnight fast. Here female offspring from 7 controls and 6 HFD fathers are included; one animal per litter was used in the analysis. This protocol was approved by the Animal Care and Ethics Committee, University of New South Wales, Australia.

The offspring of obese fathers showed impaired glucose tolerance at both 6 and 12 weeks of age, and reduced total islet area as a result of lost of large-sized islets. Islets were isolated and retroperitoneal white adipose tissue (RpWAT) dissected at 14 weeks of age; total mRNA was extracted using a miRNeasy Mini kit (Qiagen) according to the manufacturer's protocol. RNA concentration and purity were assessed spectrophotometrically (Shimadzu BioSpec-nano; Kyoto, Japan) and RNA with a RIN (RNA integrity number) (Agilent) >7.5 were selected for transcriptomics using Affymetrix Rat GeneChip® Gene 1.0 ST arrays (Affymetrix).

Microarray study design

Islet and RpWAT tissues

100 ng of total RNA from islet was labeled and hybridized onto Affymetrix Rat GeneChip® Gene 1.0 ST arrays (Affymetrix; n = 7

* Corresponding author.

http://dx.doi.org/10.1016/j.gdata.2014.09.013

(control) and n = 6 (HFD)) according to the manufacturer's instructions (Ramaciotti Centre for Genomics, UNSW, Australia). Affymetrix data were processed using the standard approach described in the Affymetrix I GeneChip Expression Analysis Technical Manual; 2006 [2]. This dataset is deposited on Gene Expression Omnibus (www.ncbi.nlm.nih.gov/geo), accession number: GSE19877. RpWAT tissue was processed as described above and the GEO accession is http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE33551.

Robust Multiarray Averaging (RMA) was used for background correction; this adjusts for probe intensities for a number of properties such as fragment length, GC content and sequence allele position [3]. Next RMA normalized data [4], n = 13 was subjected to one-way ANOVA to examine differential gene expression between HFD and control (Partek Genome Suite v6.5). Partek has summarization algorithms to compile the data of a probe set to a transcript, yielding a single number that represents a central tendency for that probe set. In this instance, we used median polish summarization, which is a commonly used summarization algorithm for Affymetrix probesets for arrays of this size (detailed in Bioconductor; https://stat.ethz.ch/pipermail/bioconductor/2003-September/002498.html).

ANOVA, analysis of variance, was used to test for differences in means of gene expression between HFD and Control animals. The only assumptions here are that the HFD and Control data are normally distributed and that the variance is approximately equal between the groups (homogeneity of variance). Thus ANOVA is very powerful and robust to detect differential gene expression between these 2 groups. Configuring the ANOVA in Partek allowed an adjustment for factors such as batch differences between the arrays, systematic technical errors, tissue type and treatment (HFD vs Control) thus allowing specification of main effects, i.e., diet. This also takes into account of Random vs Fixed effects, i.e., one factor is diet (HFD vs Control) and another factor is subject, which is the rat selected for the experiment. So in this case, Diet is a fixed factor, since levels of Control and HFD represent all conditions of interest. Subject, however is a random effect, since the rats are only a random sample of all the levels of that factor. In addition, Least Squares Mean (LS Mean) is calculated as the linear combination (sum) of the estimated means from a linear model e.g. ANOVA, regression. In an unbalanced experiment, i.e., HFD and Controls, n = 13, the LS means are preferred because they reflect the model being fitted to the data. We log transformed data prior to statistical analysis in order to transform a multiplicative effect into an additive effect.

False discovery rate

When ANOVA is used to test for a difference between three or more groups, a *post hoc* analysis is performed to see which groups are different. In this instance, we only looked at one factor, i.e., Diet (HFD vs Control), hence we did not carry out False Discovery Rate analysis. In addition, we interrogated the differentially expressed gene lists in the context of biological processes and systems, thus a *post hoc* adjustment was not applied, as this would bias the gene list to only statistically significant genes at the expense of biologically significant genes.

Systems network analysis

We used a network-based approach to enable a global view of molecular connectivity and their coordinated program in both islet and RpWAT tissues, which complements the conventional 'one variant at a time' method of linear causality models. The importance of drawing protein–protein interaction networks has recently been recognized in cancer systems analysis whereby differentially expressed genes with strong discriminatory power tend to be "passenger" or bystander genes whereas subtle changes of insignificant genes determined by pairwise statistics and threshold (gene–gene effect) can be the "drivers"

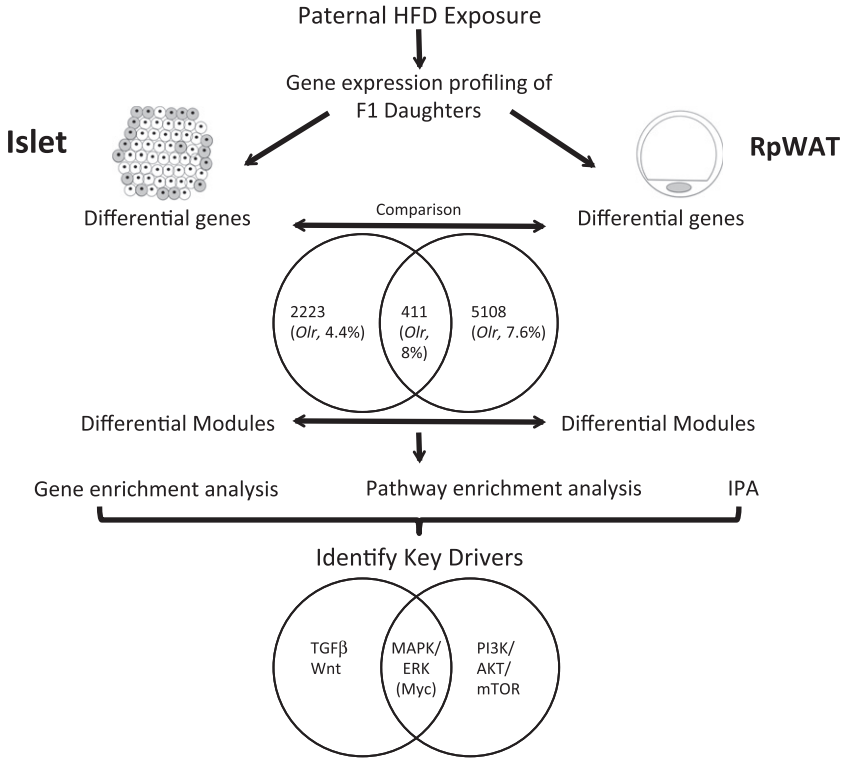

Fig. 1. We used an integrated system biology approach to interrogate the islet and RpWAT transcriptome [1,12] in order to understand the molecular events underlying the impact of paternal HFD consumption. Some of the main findings and key players are illustrated above. Of interest, there was an over-representation of olfactory receptor (*Olr*) genes in both islet and RpWAT transcriptome [12]. There were 411 differentially expressed genes in common between RpWAT and islet, of which *Olr* belonging to the olfactory transduction pathway is the most significantly enriched (P = 0.0003) [12]. Many of the differentially expressed molecules/enriched pathways or networks in both tissues belong to the MAPK/ERK pathway; one of the common molecular networks between islet and RpWAT tissues, i.e., Cell cycle also shares a common hub gene, Myc [12]. Taken together, these findings indicate paternal HFD exposure affects common network topologies in islets and RpWAT, thus implicating crosstalk or developmental changes that persist between these two metabolically related tissues.

of an integrated network response that regulate and sustain the functioning of multiple disease pathways on the proteomic scale [5–8]. It was suggested that genes with differential expression levels deemed insignificant have an effect on the molecular network through protein–protein interactions [9]. Similarly, the magnitude of differentially expressed genes encoding highly connected protein hubs has been found to be small in schizophrenia and Parkinson's disease based on human protein–protein interactome [10]. As a consequence, biologically significant genes may have been filtered out. Thus this practice undermines a biological discovery because small changes to nodes central to a network have the potential, by the nature of their connectivity, to initiate a greater impact than large changes to the peripheral nodes of the network. This phenomenon is likely to be true in complex diseases such as obesity, type 2 diabetes and cardiovascular diseases. More recently, a network-based analysis has been used in the field of developmental programming to understand the mechanisms of inter-generational transmission of disease risk [11]. We therefore, used the gene lists from unadjusted $P < 0.01$ and $P < 0.05$ in Nature [1] and $P < 0.05$ in FASEB J [12] for our systems network analysis in IPA (Ingenuity). False Discovery Rate was applied in the molecular network modeling using IPA. The data processing and discovery process is outlined in Fig. 1.

Acknowledgment

This work was supported by a project grant of the National Health and Medical research council (NHMRC) of Australia (to MJM) and funding from New South Innovations, 09_2385 (University of New South Wales (UNSW)) to MJM and SFN. SFN was supported by the Ministry of Higher Education and Universiti Kebangsaan Malaysia. RCYL was supported by the UNSW Vice-Chancellor Postdoctoral Fellowship. We thank the Ramaciotti Centre for Genomics, UNSW for array processing.

References

[1] S.F. Ng, R.C. Lin, D.R. Laybutt, R. Barres, J.A. Owens, M.J. Morris, Chronic high-fat diet in fathers programs beta-cell dysfunction in female rat offspring. Nature 467 (2010) 963–966.
[2] Affymetrix, Affymetrix GeneChip® Expression Analysis Technical Manual. 2006.
[3] B.M. Bolstad, R.A. Irizarry, M. Astrand, T.P. Speed, A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. Bioinformatics 19 (2003) 185–193.
[4] R.A. Irizarry, B. Hobbs, F. Collin, Y.D. Beazer-Barclay, K.J. Antonellis, U. Scherf, T.P. Speed, Exploration, normalization, and summaries of high density oligonucleotide array probe level data. Biostatistics 4 (2003) 249–264.
[5] A.L. Barabasi, N. Gulbahce, J. Loscalzo, Network medicine: a network-based approach to human disease. Nat. Rev. Genet. 12 (2011) 56–68.
[6] J.L. Chen, J. Li, W.M. Stadler, Y.A. Lussier, Protein-network modeling of prostate cancer gene signatures reveals essential pathways in disease recurrence. J. Am. Med. Inform. Assoc. 18 (2011) 392–402.
[7] C. Lefebvre, P. Rajbhandari, M.J. Alvarez, P. Bandaru, W.K. Lim, M. Sato, K. Wang, P. Sumazin, M. Kustagi, B.C. Bisikirska, K. Basso, P. Beltrao, N. Krogan, J. Gautier, R. Dalla-Favera, A. Califano, A human B-cell interactome identifies MYB and FOXM1 as master regulators of proliferation in germinal centers. Mol. Syst. Biol. 6 (2010) 377.
[8] C.C. Wu, K. Kannan, S. Lin, L. Yen, A. Milosavljevic, Identification of cancer fusion drivers using network fusion centrality. Bioinformatics 29 (2013) 1174–1181 (Oxford, England).
[9] F. Azuaje, Y. Devaux, D.R. Wagner, Coordinated modular functionality and prognostic potential of a heart failure biomarker-driven interaction network. BMC Syst. Biol. 4 (2010) 60.
[10] J.C. Mar, N.A. Matigian, A. Mackay-Sim, G.D. Mellick, C.M. Sue, P.A. Silburn, J.J. McGrath, J. Quackenbush, C.A. Wells, Variance of gene expression identifies altered network constraints in neurological disease. PLoS Genet. 7 (2011) e1002207.
[11] M.K. Skinner, M. Manikkam, M.M. Haque, B. Zhang, M.I. Savenkova, Epigenetic transgenerational inheritance of somatic transcriptomes and epigenetic control regions. Genome Biol. 13 (2012) R91.
[12] S.F. Ng, R.C. Lin, C.A. Maloney, N.A. Youngson, J.A. Owens, M.J. Morris, Paternal high-fat diet consumption induces common changes in the transcriptomes of retroperitoneal adipose and pancreatic islet tissues in female rat offspring. FASEB J. 28 (2014) 1830–1841.