Transcriptome analysis in blood cells from children reveals potential early biomarkers of metabolic alterations

J Sánchez, C Picó, W Ahrens, R Foraita, A Fraterman, L A Moreno, P Russo, A Siani, A Palou, on behalf of the IDEFICS Consortium

Cite this article as: J Sánchez, C Picó, W Ahrens, R Foraita, A Fraterman, L A Moreno, P Russo, A Siani, A Palou, on behalf of the IDEFICS Consortium, Transcriptome analysis in blood cells from children reveals potential early biomarkers of metabolic alterations, *International Journal of Obesity* accepted article preview 6 June 2017; doi: 10.1038/ijo.2017.132.

Received 18 January 2017; revised 6 April 2017; accepted 23 May 2017; Accepted article preview online 6 June 2017

Transcriptome analysis in blood cells from children reveals potential early biomarkers of metabolic alterations

Juana Sánchez[1], Catalina Picó[1], Wolfgang Ahrens[2,3], Ronja Foraita[2], Arno Fraterman[4], Luis A. Moreno[5], Paola Russo[6], Alfonso Siani[6], Andreu Palou[1] on behalf of the IDEFICS Consortium.

[1]Laboratory of Molecular Biology, Nutrition and Biotechnology (Nutrigenomics), University of the Balearic Islands (UIB) and CIBER Fisiopatología de la Obesidad y Nutrición (CIBEROBN), Palma de Mallorca, Spain.

[2]Leibniz Institute for Prevention Research and Epidemiology - BIPS, Bremen, Germany.

[3]Faculty of Mathematics and Computer Science, University of Bremen (UNIHB), Bremen, Germany.

[4]MVZ Eberhard & Partner Laboratoriumsmedizin, Dortmund, Germany.

[5]GENUD (Growth, Exercise, Nutrition and Development) Research Group, Faculty of Health Sciences, Universidad de Zaragoza, Instituto Agroalimentario de Aragón (IA2), Instituto de Investigación Sanitaria Aragón (IIS Aragón), Centro de Investigación Biomédica en Red de Fisiopatología de la Obesidad y Nutrición (CIBEROBN), Zaragoza, Spain.

[6]Unit of Epidemiology and Population Genetics, Institute of Food Sciences, National Research Council, Avellino, Italy.

Running title: Early biomarkers of metabolic alterations

Corresponding author and person to whom reprint should be addressed: Prof. Catalina Picó. Universitat de les Illes Balears, Biología Molecular, Nutrición y Biotecnología (Nutrigenómica). Edifici Mateu Orfila. Carretera de Valldemossa Km 7.5, 07122-Palma de Mallorca, Spain. Telephone: +34 971173070; Fax: +34 971173426. e-mail: cati.pico@uib.es

Disclosure statement: The authors have nothing to disclose.

Abstract

Objectives. The development of effective strategies to prevent childhood obesity and its comorbidities requires new, reliable early biomarkers. Here, we aimed to identify in peripheral blood cells (PBCs) potential transcript-based biomarkers of unhealthy metabolic profile associated to overweight/obesity in children.

Methods. We performed a whole-genome microarray analysis in blood cells to identify genes differentially expressed between overweight and normal weight children to obtain novel transcript-based biomarkers predictive of metabolic complications.

Results. The most significant enriched pathway of differentially expressed genes was related to oxidative phosphorylation, for which most of genes were down-regulated in overweight versus normal weight children. Other genes were involved in carbohydrate metabolism/glucose homeostasis or in lipid metabolism (e.g. TCF7L2, ADRB3, LIPE, GIPR), revealing plausible mechanisms according to existing biological knowledge. A set of differentially expressed genes was identified to discriminate in overweight children those with high or low triglyceride levels.

Conclusion. Functional microarray analysis has revealed a set of potential blood-cell transcript-based biomarkers that may be a useful approach for early identification of children with higher predisposition to obesity-related metabolic alterations.

Introduction

Obesity is nowadays considered one of the main public health problems, affecting not only adults but also children. WHO estimated that 42 million children under the age of 5 were overweight or obese in 2013 (1). What is even more worrying is the greater probability of an obese child to maintain obesity into adulthood. Moreover, early onset of obesity is associated with an increased incidence of co-morbidities, such as type 2 diabetes, non-alcoholic fatty liver disease, cardiovascular diseases and metabolic syndrome (2, 3). Knowing that obesity is influenced by several genetic, environmental and behavioural factors, and that the success of treatment is limited, prevention of obesity at early ages becomes a major need and should be one of the main focuses of attention. Thus, early diagnosis seems important for management and prevention of childhood obesity. The identification of biomarkers of predisposition to obesity-associated metabolic alterations could aid present/current strategies in decreasing in a more effective manner obesity associated co-morbidities, by targeting the underlying processes.

Transcriptome analysis in peripheral blood cells (PBCs), considering whole peripheral blood or a subpopulation of white blood cells, the so-called peripheral blood mononuclear cells (PBMCs), has been proposed as a research strategy to identify new biomarkers and candidate-genes for a number of diseases based on differentially expressed mRNA profiles (4-7). PBCs have the advantage of being easy to collect, unlike invasive biopsies. Moreover, gene expression in these cells may reflect the responses of internal organs, such as adipose tissue and liver, and thereby have been proposed as a source of biomarkers of health and disease (8-10). More specifically, PBCs have been proposed as an appropriate method for studying the cardiovascular system (6), or as a source of transcript-based biomarkers for very early stages of acute coronary disease (5). In addition, the suitability of blood cells as a potential source for biomarkers of metabolic adaptations to food intake and body weight maintenance has been reported both in animal (8, 9, 11) and human studies (12, 13). Besides, in murine models, expression levels of selected genes in PBCs have been proposed as predictive biomarkers of a healthy or distorted metabolic state due to interventions during the perinatal period (14, 15).

In the present study, a whole genome microarray analysis was performed in PBCs to identify genes differentially expressed between overweight and normal weight children, using a cross-sectional design, in order to derive novel transcript-based biomarkers that could be predictive of metabolic complications in early life. These biomarkers could be helpful for identifying particularly those obese children who are at risk of metabolic complications, hence allowing early interventions in a more effective way.

Materials and Methods

Participants

IDEFICS is a large European multi-centre study on childhood obesity (details in (16)). Using a cross-sectional design, whole-genome microarray analysis has been performed in PBCs from a subgroup of 32 individuals with normal weight (17) and overweight (15), belonging to the Spanish IDEFICS cohort. Children included in the analysis were randomly selected within normal weight and overweight individuals and matched by sex. BMI categories were defined according to Cole (17). The age range was 4.7-8.0 years.

Blood sampling and processing for gene expression analysis

For each participant, a total of 2.5 mL peripheral blood was collected under fasting conditions into PAXgene vacutainer tubes (Qiagen, Izasa-Barcelona, Spain) via antecubital fossa venipuncture. Total RNA was isolated using the PAXgene blood RNA kit according to the manufacturer's instructions (Qiagen) and as previously described (13).

Microarray processing

From each sample, 80 ng of RNA was reverse transcribed to complementary DNA (cDNA) using the Agilent Low Input Quick Amp Labeling kit (Agilent Technologies, Inc., CA, USA) according to the manufacturer's protocol. Then, half of the cDNA sample (10 μl) was used for the linear amplification of RNA and labelling with cyanine-3 (Cy3) or Cy5. Transcription and labelling were carried out at 40 °C for 2 h. The labelled and amplified cRNA samples were purified using Qiagen Rneasy MiniSpin columns (Qiagen, Madrid, Spain). The incorporation of

dyes and cRNA concentration was measured using the NanoDrop ND 1000 spectrophotometer (NanoDrop Techonologies, Ins., Wilmington, DE). Then, 825 ng of cRNA labelled with Cy5 from each sample and 825 ng of Cy3 pool was hybridized on 4x44K G4845A human whole genome Agilent microarrays (Agilent Technologies, Inc., Santa Clara, CA, USA) for 17 h at 65 °C in hybridization chambers in an oven rotating at 10 rpm (Agilent Technologies). After hybridization, the arrays were washed with "GE wash buffer 2" for 1 min at 37 °C, followed by acetonitrile for 10 s at room temperature, and finally with a solution for stabilization and drying for 30 s at room temperature, according to the manufacturer's protocol (Agilent Technologies).

Microarray data analysis

The arrays were scanned with an Agilent Microarray Scanner (Agilent Technologies). Scanned images were examined for visible defects and proper grid alignment. The intensities of the signals from each spot were quantified, and the raw data were extracted using Feature Extraction Software version 10.10.1.1 (Agilent Technologies).

Background correction was performed by subtracting the background intensities from the foreground intensities. Normalisation within and between arrays was conducted using the loess method and the quantile method, respectively. All samples showed correct normalisation and were thus used in further statistical analysis.

The Babelomics 4.3 platform (http://babelomics.bioinfo.cipf.es), a suite of web tools for microarray data analysis (18), was applied for statistical analysis. The patterns of probes matching for the same gene were merged, and hence obtaining the median. Next, differential gene expression between overweight and normal weight children was assessed by the Limma test. Fold changes were the difference in log2 mean values between experimental groups. The threshold of significance for this statistical test was set at $p<0.01$. To control for false positives we selected gene expression differences for a false discovery rate (FDR) (19) of 16%. Biological information of the identified genes was obtained in available databases (NCBI, UniProt, Genecards). In addition, to gain insight into biological mechanisms, pathway analysis was conducted using Metacore™ (GeneGo Inc). Genes with expression levels significantly

different between normal weight and overweight children were plotted in a heat map using R (v 3.1.1, R Development Core Team). Partial least squares discriminant analysis (PLS-DA) was performed using the Excel add-in Multibase (Numerical Dynamics, Japan) to select a more reduced number of gene transcripts to discriminate between overweight and normal weight subjects, with high and low metabolic risk.

Reverse transcription quantitative polymerase chain reaction (RT-qPCR) analysis

To validate the results from the microarray analysis, mRNA expression levels of 10 selected genes were measured by RT-qPCR. These genes included the five genes from oxidative phosphorylation chain that displayed the greatest changes in the array analysis and other selected genes related to food intake control, signalling and carbohydrate metabolism (see Supplemental Table 1 for further details). RT-qPCR was performed as previously described (13). Briefly, 0.25 µg of total RNA was denatured at 65 °C for 10 min and then reverse transcribed to cDNA using MuLV reverse transcriptase at 20ºC for 15 min, 42 °C for 30 min, with a final step of 5 min at 95 °C in an Applied Biosystems 2720 Thermal Cycler. Each PCR was performed from diluted (1/5) cDNA template, forward and reverse primers (1 mM each) and Power SYBER Green PCR Master Mix. Real time PCR was performed using the Applied Biosystems StepOnePlus™ Real-Time PCR Systems with the following profile: 10 min at 95 ºC, followed by a total of 40 two- temperature cycles (15 s at 95 ºC and 1 min at 60 ºC). In order to verify the purity of the products, a melting curve was produced after each run according to the manufacturer's instructions. The threshold cycle (Ct) was calculated using the instrument's software (StepOne Software version 2.2.2), and the relative expression of a target gene to a housekeeping gene was calculated. Tripartite motif containing 27 (TRIM27) was chosen as housekeeping gene because showed equal expression levels across all groups.

Statistical analysis

No blinding was carried out for data analysis. Data are presented as mean and the standard error of the mean (SEM). The statistical analysis of microarray data was conducted as described above. Statistical analysis of the anthropometric and biochemical parameters was performed

using SPSS version 20. The median value of each gene's expression distribution in the overall group was selected as an arbitrary cut-off point to group gene expressions in low and high expression categories. Levene´s test was performed to assess whether the variance is equal between groups. Single comparisons between groups were assessed by Student's t-test. Two-way ANOVA was used to determine differences between BMI and gene expression categories (high vs low). Threshold of significance was defined at $p<0.05$.

Results

Subject characteristics

Characteristics of the subjects included in the study are shown in Table 1. The random sample consists of 32 children from the Spanish cohort of the IDEFICS study, uniformly distributed by weight and sex, and with a range of age between 4.7 and 8.0 years. Compared with normal weight, overweight children exhibited as expected higher body weight, BMI and BMI z-score (20), subcutaneous fat (calculated as the sum of four different skin folds), waist/height ratio and HOMA-IR index, and also presented higher insulinemia and triglyceridemia ($p<0.01$, Student's t-test). No differences were found between normal weight and overweight children concerning total, HDL and LDL cholesterol levels. Mean age of overweight children was significantly higher than that of normal weight children.

Gene expression profiling in PBCs

Of the 45,220 probes tested in the microarray analysis, 29,094 probes remained after background correction, normalization and merging replicated probes and were used for further analysis. In total, 1,716 unique genes were found to be expressed significantly different between normal weight and overweight children ($p<0.01$; Limma test). There were 550 unknown genes, 50 non coding genes, and 39 were pseudogenes and hence were not considered in the following steps. From the remaining 1,077 known coding genes, 505 were down-regulated and 572 up-regulated in overweight compared to normal weight children. The results from the array

concerning the 1,077 genes differently expressed between normal weight and overweight children across individual samples are depicted in a heat map (see Figure 1).

Functional and biological analysis

Using available databases (NCBI, UniProt and Genecards), the 1,077 known genes with significantly different expression levels between normal weight and overweight children were classified into several biological processes according to their function (Supplemental Figure 1). The biological processes showing the highest number of differentially expressed genes were related to transcription/translation machinery (n=223), cell turnover (n=123), signalling (n=121), immune system (n=78), metabolism of proteins (n=56), transport (n=53), metabolism of lipids (n=50), nervous system (n=45), central metabolism (n=44) and cytoskeleton (n=34). Other biological processes with a notable number of differently expressed genes (n=12-30) were related to neural signalling, blood, sensory perception of taste and smell, metabolism of carbohydrates, redox metabolism, epigenetic modification, metabolism of amino acids and protein or lipid glycosylation. The remaining genes were related to the metabolism of nucleotides and amines, metabolism of vitamins and minerals, cell communication, food intake control and vascular homeostasis (with 11 or less genes involved in each process). Other genes related to biological processes such as autophagy, cell-cell adhesion, collagen, keratins, extra cellular matrix, etc. were grouped under 'miscellanea'

The set of 1,077 known genes was further analysed for functional and biological pathways using the MetaCore™ platform (GeneGo Inc). The top ten pathway maps are listed in Table 2. The most significant enriched pathway in our data set was related to mitochondrial oxidative phosphorylation. Notably, most of the genes involved in this pathway were down-regulated in overweight compared to normal weight children. Results regarding the five genes related to oxidative phosphorylation displaying the greatest changes in the microarray analysis (ATP5O, NDUFS5, NDUFB8, COX6C and ATP1F1) were fully validated by RT-qPCR (Supplemental Table 1).

Expression levels of genes involved in the tricarboxylic acid (TCA) cycle were also reassessed, since this process supplies NADH and succinate to oxidative phosphorylation. Notably, key genes in the TCA cycle were generally down-regulated in overweight compared to normal weight children, but particularly regarding succinate handling. This is the case of SUCLG2 (succinate-CoA ligase, GDP-forming, beta subunit), SDHB (succinate dehydrogenase complex, subunit B, integral membrane protein, 15kDa), SDHC (succinate dehydrogenase complex, subunit C, integral membrane protein, 15kDa) and MDH2 (malate dehydrogenase 2, NAD (mitochondrial)). Furthermore, genes encoding for two subunits of the pyruvate dehydrogenase (PDH) multi-enzyme complex, PDHB and PDHX, which catalyse the overall conversion of pyruvate to acetyl coenzyme and provide the primary link between glycolysis and the TCA cycle, were down-regulated in overweight children compared to normal weight children ($p=0.013$ and $p=0.017$, respectively, Limma test,). Up-regulated mRNA expression was also detected in overweight children for PC (pyruvate carboxylase), which catalyses carboxylation of pyruvate to form oxaloacetate.

Besides the above mentioned, other differentially expressed genes between normal weight and overweight children are suggested as relevant regarding obesity-related metabolic alterations, according to existing biological knowledge. Among them are the transcription factor 7-like 2 (TCF7L2) and the gastric inhibitory polypeptide receptor (GIPR) (associated with carbohydrate metabolism and glucose homeostasis) or the adrenergic, beta-3-, receptor (ADRB3) and the lipase, hormone-sensitive (LIPE) (associated with lipid metabolism).

Biomarker selection

To refine intergroup differences between normal weight and overweight children and select the most relevant and discriminatory genes with the greatest potential as biomarkers, data were also analysed with the partial least squares discriminant analysis (PLS-DA). The set of 1,077 genes was included in the analysis. PLS-DA score plot also showed a clear separation between normal weight and overweight children (Figure 2A). We retained those 30 genes with the highest

loading in PC1 and in PC2 (Supplemental Table 2) and carried out a new PLS-DA. We obtained a model with a clear separation between normal weight and overweight individuals (Figure 2B). PC1 and PC2 accounted for 45.7% and 29.9 % of the variance. The ten variables that mostly contributed to the model, all of them included in the PC1, were zinc finger protein 418 (ZNF418), natriuretic peptide A (NPPA), polymerase (RNA) I polypeptide C, 30kDa (POLR1C), adrenergic, beta-3-, receptor (ADRB3), purinergic receptor P2X, ligand-gated ion channel, 2 (P2RX2), dynein, axonemal, light chain 4 (DNAL4), neuronal signalling (cortistatin (CORT), trefoil factor 1 (TFF1), neuronal growth regulator 1 (NEGR1) and Interleukin 15 receptor, alpha (IL15RA) (Figure 2C).

Association studies of gene expression with circulating triglyceride levels

Children were grouped into categories (low/ high) according to selected biochemical parameters, HOMA index, insulin and triglycerides (TG). The median value of each parameter's distribution in the overall group was selected as a cut-off point to define these categories. A new PLS-DA was built (including the 1,077 genes differentially expressed between normal weight and overweight children) to identify genes that differentiate children according not only to their body weight, but also to their metabolic risk. Only the results obtained in PLS-DA for discrimination of children with high and low TG levels within normal weight and overweight children displayed a valid model (Supplemental Figure 2). The PLS-DA showed that within normal weight children the model does not discriminate between those having high or low TG levels, because both groups are located in the quadrant with the higher score for PC1 (Figure 3). However, this model shows better separation within overweight children (Figure 3). Overweight children with low TG levels were mainly located on the right side of PC1, as normal weight children. This suggests that those genes with a high loading score in PC1 may be good candidates as biomarkers to discriminate between overweight children with high or low TG levels. The genes are listed in Supplemental Table 3. It is noteworthy that the list of genes with high loadings in PC1 includes 8 of the 10 genes previously listed in Figure 2C (ZNF418, NPPA, POLR1C, ADRB3, P2RX2, DNAL4, CORT, TFF1) as those that mostly contributed to the

separation between normal weight and overweight children. The discriminatory power of the expression profile of the above-mentioned genes on TG levels within overweight children was further studied by considering each individual gene (Figure 4). For that, normal weight and overweight children were subdivided according to their gene expression levels (low vs high) by using the median value of each gene expression's distribution of the whole sample as a cut-off point. Overweight children with high expression levels of most of the genes (ZNF418, NPPA, POLR1C, ADRB3, DNAL4, CORT and TFF1), exhibited higher TG levels than those with low expression levels ($p<0.05$, Student's t-test), whereas no differences were found in normal weight children. In the case of P2RX2, the trend was similar, but differences were not statistically significant.

Discussion

To identify biomarkers that can be predictive of unhealthy metabolic alterations associated to overweight/obesity in children, we first performed a comparative analysis of the transcriptome profile in PBCs of normal weight and overweight children. Overall 1,077 known coding genes differently expressed between these two groups of children were detected, clearly separated by both PLS-DA and the heat map. Microarray data were confirmed by RT-qPCR performed for a selection of 10 genes.

Pathway enrichment analysis revealed that the most significant enriched pathway was related to mitochondrial oxidative phosphorylation. Most of the genes involved in this pathway were down-regulated in overweight compared to normal weight children. A reduction in the expression levels of genes involved in the oxidative phosphorylation has also been described in the skeletal muscle (21, 22), visceral adipose tissue (23) and in PBMCs (24) from adult patients with type 2 diabetes. It has also been reported that the reduction in the set of genes involved in oxidative phosphorylation in PBMCs seems to be intrinsic to the diabetic state since altered expression was not changed by glycaemic control (24). Notably, overweight children involved in the present study were not diagnosed with type 2 diabetes, although they displayed higher insulin levels and HOMA index values compared to normal weight children with no alteration

in fasting glucose levels. In addition, we also observed a parallel reduction in key genes of the TCA cycle and of genes encoding for two subunits of the pyruvate dehydrogenase complex, which catalyses the conversion of pyruvate to acetyl-CoA and provides the primary link between glycolysis and the TCA cycle. Conversely, pyruvate carboxylase, which catalyses the carboxylation of pyruvate to oxaloacetate, was up-regulated in overweight children, suggesting that pyruvate may be driven to the conversion to oxaloacetate more than to acetyl-CoA. Overall, the expression level of key genes in PBCs pointed to a disturbed oxidative capacity in overweight children, which could reflect impairment in other internal tissues. This may be a hallmark of overweight/obesity, also indicative of a pre-diabetic state, regardless of an alteration in glucose levels.

Moreover, also in reference to type 2 diabetes, differences in the expression levels of transcription factor 7-like 2 (TCF7L2) may be of potential interest. TCF7L2 is a key regulator of (pro)insulin synthesis and processing (25). TCF7L2 harbors common genetic variants (particularly the single-nucleotide polymorphism rs7903146) with the strongest and most widely replicated effect on type 2 diabetes risk already described (26, 27). The risk T-allele of rs7903146 has been associated with impaired glucose-stimulated insulin secretion and processing and increased insulin resistance (28), and carriers of the T-allele have been reported to display increased expression of TCF7L2 in human pancreatic islets (29). The frequency of the risk T-allele has been found to be higher not only in diabetic patients, but also in those who develop macro-vascular complications (30). Moreover, centenarians show the lowest frequency of T-allele and the highest frequency of C-allele, which emerges as a protective longevity variant (30).

Notably, in the present study, TCF7L2 has been identified as one of the genes differently expressed between normal weight and overweight children, being up-regulated in overweight children. These results might suggest a higher presence of the risk T-allele in overweight individuals, although this has not been addressed here since it was not the objective of the study.

In any case, present findings show that greater TCF7L2 expression levels in PBCs, regardless of the polymorphism variant, may by indicative of higher risk of obesity and diabetes in children.

PLS-DA is considered a powerful and simple tool for analysing microarray data that allows to predict the phenotype given the expression levels (31). Here, PLS-DA was successfully applied to select a more reduced number of gene transcripts to discriminate between overweight and normal weight subjects. The 10 genes, whose transcripts have been identified as most relevant to the discrimination between overweight and normal weight children in our PLS-DA model, were related to metabolism (ADRB3), nervous system (NEGR1), neuronal signalling (CORT and P2RX2), signalling (IL15RA), vascular homeostasis (NPPA), transcription/translation machinery (POLR1C and ZNF418), cytoskeleton (DNAL4) and TFF1. mRNA levels of all of them were up-regulated in PBCs of overweight compared to normal weight children. Moreover, using a new PLS-DA model, we defined a set of genes whose expression levels may discriminate between overweight individuals with high or low TG levels. It is worth noting that the latter set of genes includes 8 of the list of 10 genes above mentioned (ZNF418, NPPA, POLR1C, ADRB3, P2RX2, DNAL4, CORT and TFF1). Other genes with a relevant role in lipid metabolism (such as LIPE), or glucose homeostasis (such as GIPR) are also included. Both genes were also up-regulated in overweight children.

Some genes identified as candidate biomarkers in both PLS-DA models seem to be especially remarkable in terms of biological plausibility regarding body weight and metabolic aspects. Particularly, ADRB3 is related to lipid metabolism. The ADRB3 gene encodes for the adrenergic receptor beta 3, which is mainly present in adipose tissue and involved in the regulation of lipolysis and thermogenesis. A single nucleotide polymorphism in the coding region of this gene has been associated with childhood obesity and other features of the metabolic syndrome (32).

In addition, genome-wide association studies have identified sequence variants of loci near to NEGR1 and GIPR associated with BMI (33-35). GIPR encodes a G-protein coupled receptor

for gastric inhibitory polypeptide (GIP). GIP is well known for its insulin tropic action and its secretion is stimulated by both fat and glucose (36). In mice, knock-out or antagonism of GIPR protects from obesity and metabolic abnormalities, such us insulin resistance induced by the consumption of a high-fat diet (37, 38). In humans, GIP has been shown to increase adipose tissue expression and plasma concentration of MCP-1 and seems to promote a crosstalk between macrophages and adipocytes to induce an inflammatory response associated to high-density food intake and obesity (39). Regarding NEGR1, the protein encoded by this gene is a cell adhesion molecule expressed in different areas of CNS. Experimental data with NEGR1-deficient mice showed that the loss of function of NEGR1 has negative consequences on body mass and produces a reduction in food intake, although the mechanisms involved are still unknown (40). In humans, the at-risk allele of the associated variant (rs2568958) has been related to increased expression (per allele copy) of NEGR1 in blood (34), pointing that up-regulation of NEGR1 in human may be associated to increased BMI (40).

CORT may also be of potential interest. This gene encodes for the cortistatin (CORT) neuropeptide, which shares high structural and functional similarities with somatostatin (SST). In addition to their roles in the CNS, both peptides also regulate endocrine secretion. CORT has been shown to bind with high affinity to the ghrelin receptor (GHSR) and play important actions on pancreatic function (41). In islets isolated from C57BL6 mice, CORT action has been reported to reduce the glucose-stimulated insulin secretion (42). Therefore, the presence of higher expression levels of this gene in PBCs from overweight children may also be in agreement with a greater predisposition to type 2 diabetes.

Despite the care taken in selecting the study cohort to have a similar number of normal weight and overweight children within the range of age considered, the study has the limitation that the mean age of the overweight group is significantly higher than that of the normal weight group. Although it is not expected that the age difference is responsible for the differences found between groups, it does lead to limitations in the interpretation of the results that may be considered in further studies. Another fact to take into account is the use of the whole fraction of

blood cells, instead of PBMC sub-populations, because this provides methodological advantages. Therefore, differences in PBMC sub-populations could be masking some changes in gene expression that could be used as early biomarkers. Finally, the cross-sectional nature of the study may also be considered as a limitation, since it may not provide definite information on cause-and-effect relationships. The predictive value of the identified biomarkers needs further validation in follow-up studies.

In summary, functional microarray analysis of PBCs from normal weight and overweight children has revealed early changes associated to overweight, indicative of a reduced capacity of metabolic handling through the TCA cycle and oxidative phosphorylation, which, if generalized to other tissues, may contribute to the energy imbalance and later metabolic disorders. Moreover, other genes related with glucose homeostasis and lipid metabolism revealed other putative mechanisms according to biological knowledge. With a PLS-DA model, a set of potential transcript-based biomarkers has been identified that, once validated in future studies, may be a useful approach for early identification of children with higher predisposition to obesity and its related metabolic alterations, particularly in terms of insulin resistance and hypertriglyceridemia. However, due to the cross-sectional nature of the analysis, the predictive value of the identified biomarkers should be considered with some reservations and requires verification.

Acknowledgments

We are grateful for the support provided by school boards, headmasters, teachers, school staff and communities, and for the effort of all study nurses and data managers. This work was done as part of the IDEFICS study (http://www.idefics.eu) and the I.Family Study (http://www.ifamilystudy.eu/). We gratefully acknowledge the financial support of the European Community within the Sixth RTD Framework Programme Contract No. 016181 (FOOD) for the IDEFICS study and within the Seventh RTD Framework Programme Contract No. 266044 for the I.Family study. This work was also supported by the Instituto de Salud

Carlos III, Centro de Investigación Biomédica en Red Fisiopatología de la Obesidad y Nutrición, CIBERobn. LBNB is a member of the European Research Network of Excellence NuGO (The European Nutrigenomics Organization, EU Contract: nº FP6-506360).

Conflict of interest: The authors declare that they have no conflict of interest

References

1. WHO. Obesity and overweight. 2015 [updated January 2015]; Available from: http://www.who.int/mediacentre/factsheets/fs311/en/.

2. Ahrens W, Moreno LA, Marild S, Molnar D, Siani A, De Henauw S, et al. Metabolic syndrome in young children: definitions and results of the IDEFICS study. Int J Obes. 2014 Sep;38 Suppl 2:S4-14.

3. Choquet H, Meyre D. Genomic insights into early-onset obesity. Genome Med. 2010;2(6):36.

4. Bittel DC, Kibiryeva N, Sell SM, Strong TV, Butler MG. Whole genome microarray analysis of gene expression in Prader-Willi syndrome. Am J Med Genet A. 2007 Mar 1;143A(5):430-42.

5. Silbiger VN, Luchessi AD, Hirata RD, Lima-Neto LG, Cavichioli D, Carracedo A, et al. Novel genes detected by transcriptional profiling from whole-blood cells in patients with early onset of acute coronary syndrome. Clin Chim Acta. 2013 Jun 5;421:184-90.

6. Visvikis-Siest S, Marteau JB, Samara A, Berrahmoune H, Marie B, Pfister M. Peripheral blood mononuclear cells (PBMCs): a possible model for studying cardiovascular biology systems. Clin Chem Lab Med. 2007;45(9):1154-68.

7. Sanchez J, Bonet ML, Keijer J, van Schothorst EM, Molller I, Chetrit C, et al. Blood cells transcriptomics as source of potential biomarkers of articular health improvement: effects of oral intake of a rooster combs extract rich in hyaluronic acid. Genes Nutr. 2014 Sep;9(5):417.

8. Caimari A, Oliver P, Keijer J, Palou A. Peripheral blood mononuclear cells as a model to study the response of energy homeostasis-related genes to acute changes in feeding conditions. Omics. 2010 Apr;14(2):129-41.

9. Oliver P, Reynes B, Caimari A, Palou A. Peripheral blood mononuclear cells: a potential source of homeostatic imbalance markers associated with obesity development. Pflugers Arch. 2013 Apr;465(4):459-68.

10. de Mello VD, Kolehmanien M, Schwab U, Pulkkinen L, Uusitupa M. Gene expression of peripheral blood mononuclear cells as a tool in dietary intervention studies: What do we know so far? Molecular nutrition & food research. 2012 Jul;56(7):1160-72.

11. Caimari A, Oliver P, Rodenburg W, Keijer J, Palou A. Feeding conditions control the expression of genes involved in sterol metabolism in peripheral blood mononuclear cells of normoweight and diet-induced (cafeteria) obese rats. The Journal of nutritional biochemistry. 2010 Nov;21(11):1127-33.

12. Leonardson AS, Zhu J, Chen Y, Wang K, Lamb JR, Reitman M, et al. The effect of food intake on gene expression in human peripheral blood. Hum Mol Genet. 2010 Jan 1;19(1):159-69.

13. Sanchez J, Priego T, Pico C, Ahrens W, De Henauw S, Fraterman A, et al. Blood cells as a source of transcriptional biomarkers of childhood obesity and its related metabolic alterations: results of the IDEFICS study. The Journal of clinical endocrinology and metabolism. 2012 Apr;97(4):E648-52.

14. Konieczna J, Sanchez J, Palou M, Pico C, Palou A. Blood cell transcriptomic-based early biomarkers of adverse programming effects of gestational calorie restriction and their reversibility by leptin supplementation. Sci Rep. 2015;5:9088.

15. Konieczna J, Sanchez J, van Schothorst EM, Torrens JM, Bunschoten A, Palou M, et al. Identification of early transcriptome-based biomarkers related to lipid metabolism in peripheral blood mononuclear cells of rats nutritionally programmed for improved metabolic health. Genes Nutr. 2014 Jan;9(1):366.

16. Ahrens W, Bammann K, Siani A, Buchecker K, De Henauw S, Iacoviello L, et al. The IDEFICS cohort: design, characteristics and participation in the baseline survey. Int J Obes. 2011 Apr;35 Suppl 1:S3-15.

17. Cole TJ, Bellizzi MC, Flegal KM, Dietz WH. Establishing a standard definition for child overweight and obesity worldwide: international survey. Bmj. 2000 May 6;320(7244):1240-3.

18. Medina I, Carbonell J, Pulido L, Madeira SC, Goetz S, Conesa A, et al. Babelomics: an integrative platform for the analysis of transcriptomics, proteomics and genomic data with advanced functional profiling. Nucleic Acids Res. 2010 Jul;38(Web Server issue):W210-3.

19. Benjamini Y, Drai D, Elmer G, Kafkafi N, Golani I. Controlling the false discovery rate in behavior genetics research. Behav Brain Res. 2001 Nov 1;125(1-2):279-84.

20. Cole TJ, Lobstein T. Extended international (IOTF) body mass index cut-offs for thinness, overweight and obesity. Pediatr Obes. 2012 Aug;7(4):284-94.

21. Mootha VK, Lindgren CM, Eriksson KF, Subramanian A, Sihag S, Lehar J, et al. PGC-1alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. Nat Genet. 2003 Jul;34(3):267-73.

22. Patti ME, Butte AJ, Crunkhorn S, Cusi K, Berria R, Kashyap S, et al. Coordinated reduction of genes of oxidative metabolism in humans with insulin resistance and diabetes: Potential role of PGC1 and NRF1. Proceedings of the National Academy of Sciences of the United States of America. 2003 Jul 8;100(14):8466-71.

23. Dahlman I, Forsgren M, Sjogren A, Nordstrom EA, Kaaman M, Naslund E, et al. Downregulation of electron transport chain genes in visceral adipose tissue in type 2 diabetes independent of obesity and possibly involving tumor necrosis factor-alpha. Diabetes. 2006 Jun;55(6):1792-9.

24. Takamura T, Honda M, Sakai Y, Ando H, Shimizu A, Ota T, et al. Gene expression profiles in peripheral blood mononuclear cells reflect the pathophysiology of type 2 diabetes. Biochem Biophys Res Commun. 2007 Sep 21;361(2):379-84.

25. Zhou Y, Park SY, Su J, Bailey K, Ottosson-Laakso E, Shcherbina L, et al. TCF7L2 is a master regulator of insulin production and processing. Hum Mol Genet. 2014 Dec 15;23(24):6419-31.

26. Cauchi S, El Achhab Y, Choquet H, Dina C, Krempler F, Weitgasser R, et al. TCF7L2 is reproducibly associated with type 2 diabetes in various ethnic groups: a global meta-analysis. J Mol Med (Berl). 2007 Jul;85(7):777-82.

27. Grant SF, Thorleifsson G, Reynisdottir I, Benediktsson R, Manolescu A, Sainz J, et al. Variant of transcription factor 7-like 2 (TCF7L2) gene confers risk of type 2 diabetes. Nat Genet. 2006 Mar;38(3):320-3.

28. Saxena R, Gianniny L, Burtt NP, Lyssenko V, Giuducci C, Sjogren M, et al. Common single nucleotide polymorphisms in TCF7L2 are reproducibly associated with type 2 diabetes and reduce the insulin response to glucose in nondiabetic individuals. Diabetes. 2006 Oct;55(10):2890-5.

29. Lyssenko V, Lupi R, Marchetti P, Del Guerra S, Orho-Melander M, Almgren P, et al. Mechanisms by which common variants in the TCF7L2 gene increase risk of type 2 diabetes. The Journal of clinical investigation. 2007 Aug;117(8):2155-63.

30. Garagnani P, Giuliani C, Pirazzini C, Olivieri F, Bacalini MG, Ostan R, et al. Centenarians as super-controls to assess the biological relevance of genetic risk factors for common age-related diseases: a proof of principle on type 2 diabetes. Aging (Albany NY). 2013 May;5(5):373-85.

31. Perez-Enciso M, Tenenhaus M. Prediction of clinical outcome with microarray data: a partial least squares discriminant analysis (PLS-DA) approach. Hum Genet. 2003 May;112(5-6):581-92.

32. Erhardt E, Czako M, Csernus K, Molnar D, Kosztolanyi G. The frequency of Trp64Arg polymorphism of the beta3-adrenergic receptor gene in healthy and obese Hungarian children and its association with cardiovascular risk factors. Eur J Clin Nutr. 2005 Aug;59(8):955-9.

33. Speliotes EK, Willer CJ, Berndt SI, Monda KL, Thorleifsson G, Jackson AU, et al. Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index. Nat Genet. 2010 Nov;42(11):937-48.

34. Thorleifsson G, Walters GB, Gudbjartsson DF, Steinthorsdottir V, Sulem P, Helgadottir A, et al. Genome-wide association yields new sequence variants at seven loci that associate with measures of obesity. Nat Genet. 2009 Jan;41(1):18-24.

35. Willer CJ, Speliotes EK, Loos RJ, Li S, Lindgren CM, Heid IM, et al. Six new loci associated with body mass index highlight a neuronal influence on body weight regulation. Nat Genet. 2009 Jan;41(1):25-34.

36. Pais R, Gribble FM, Reimann F. Stimulation of incretin secreting cells. Ther Adv Endocrinol Metab. 2016 Feb;7(1):24-42.

37. McClean PL, Irwin N, Cassidy RS, Holst JJ, Gault VA, Flatt PR. GIP receptor antagonism reverses obesity, insulin resistance, and associated metabolic disturbances induced in mice by prolonged consumption of high-fat diet. American journal of physiology Endocrinology and metabolism. 2007 Dec;293(6):E1746-55.

38. Miyawaki K, Yamada Y, Ban N, Ihara Y, Tsukiyama K, Zhou H, et al. Inhibition of gastric inhibitory polypeptide signaling prevents obesity. Nat Med. 2002 Jul;8(7):738-42.

39. Gogebakan O, Osterhoff MA, Schuler R, Pivovarova O, Kruse M, Seltmann AC, et al. GIP increases adipose tissue expression and blood levels of MCP-1 in humans and links high energy diets to inflammation: a randomised trial. Diabetologia. 2015 Aug;58(8):1759-68.

40. Lee AW, Hengstler H, Schwald K, Berriel-Diaz M, Loreth D, Kirsch M, et al. Functional inactivation of the genome-wide association study obesity gene neuronal growth regulator 1 in mice causes a body mass phenotype. PLoS One. 2012;7(7):e41537.

41. Chanclon B, Luque RM, Cordoba-Chacon J, Gahete MD, Pozo-Salas AI, Castano JP, et al. Role of endogenous cortistatin in the regulation of ghrelin system expression at pancreatic level under normal and obese conditions. PLoS One. 2013;8(2):e57834.

42. Rafacho A, Castellano-Muñoz M, Alonso-Magdalena P, Irles E, Bello M, Vetorazzi J, et al. Cortistatin hyperpolarizes pancreatic beta cell membrane and reduces glucose-stimulated insulin secretion. Diabetology & Metabolic Syndrome. 2015;7(Suppl 1):A250-A.

Figure Legends

Figure 1. Heat map representing individual expression data of genes differentially expressed in PBCs between normal and overweight children from the IDEFICS study. Rows represent the set of 1,077 differentially expressed genes, sorted by fold change. Columns represent log2-transformed gene expression values for each gene in each individual.

Figure 2. A) Partial least square discriminant analysis (PLS-DA) plots involving 1,077 genes differentially expressed between normal weight and overweight. B) PLS-DA plot including the genes with the highest loading score from the first PLS-DA (Supplemental Table 2). C) Plot with variable loading from PLS-DA showed in C with the 10 variables that contributed the most to variability of individual expression data. PLS-DA analysis was performed using the Excel add-in Multibase package (Numerical Dynamics, Japan).

Figure 3. Partial least square discriminant analysis (PLS-DA) for discrimination of individuals with high and low levels of triglycerides (HighTG; LowTG) within normal weight and overweight children (NW; OW), including 1,077 genes differentially expressed between normal weight and overweight.

Figure 4. Triglyceride levels in children with low and high expression levels of specific genes in blood cells. Genes selected were: ZNF418, NPPA, POLR1C, ADRB3, P2RX2, DNAL4, CORT or TFF1. Results are mean ± standard error of the mean. Statistics: W, effect of BMI; E, effect of expression levels (low vs high), WxE, interactive effect of BMI and expression levels ($p<0.05$ two-way ANOVA);*$p<0.05$ high vs low expression levels (Student's t-test); $p<0.05$ overweight vs normal weight (Student's t-test).

Table 1. General characteristics, anthropometry and biochemical parameters in the population analysed in the study. BMI, body mass index. BMI z-score according to Cole et al (20), Subcutaneous fat represents the sum of four different skin folds (subscapular, biceps, triceps and suprailliac crest). Statistics: *,** differences between overweight and normal weight children ($p<0.05$ and $p<0.01$, respectively, by Student's t-test).

	Normal weight (n=17; 9 ♂+ 8♀)	**Overweight (n=15; 7 ♂+ 8♀)**
Weight(kg)	19.7 ± 0.7	36.4 ± 1.2**
BMI (kg/m^2)	14.6 ± 0.3	22.5 ± 0.4**
BMI z-score (kg/m^2)	-0.56 ± 0.21	2.71 ± 0.10**
Age (years)	6.3 ± 0.3	7.1 ± 0.2*
Subcutaneous fat (mm)	22.8 ± 2.3	66.2 ± 2.6**
Waist/height Ratio	0.441 ± 0.008	0.563 ± 0.010**
Insulin (µIU/mL)	3.96 ± 0.76	8.27 ± 0.89**
Glucose (mg/dL)	83.4 ± 2.2	87.2 ± 1.7
HOMA index	0.84 ± 0.17	1.76 ± 0.17 **
Triglycerides (mg/dL)	51.1 ±4.0	67. 9± 6.1*
HDL cholesterol (mg/dL)	53.9 ± 4.4	49.5 ± 3.4
LDL cholesterol (mg/dL)	87.6 ± 4.9	90.5 ± 7.9
Total cholesterol (mg/dL)	151.7 ± 6.3	153.5 ± 8.6
Total cholesterol/HDL cholesterol	3.0 ± 0.2	3.3 ± 0.2

Table 2. Top 10 ranked pathways from MetaCore pathway analysis. Pathways are ranked based upon p-value.

	Pathway	pValue	FDR	Ratio
1	**Oxidative phosphorylation**	**7.937E-09**	**5.413E-06**	**20/105**
2	Development Role of HDAC and calcium/calmodulin-dependent kinase (CaMK) in control of skeletal myogenesis	9.963E-06	3.397E-03	11/54
3	Cell cycle_Initiation of mitosis	4.890E-05	8.513E-03	7/25
4	Signal transduction_AKT signaling	5.149E-05	8.513E-03	9/43
5	Immune response_IL-4 signaling pathway	6.241E-05	8.513E-03	9/44
6	Cell cycle_Chromosome condensation in prometaphase	1.527E-04	1.529E-02	6/21
7	Regulation of degradation of deltaF508-CFTR in CF	1.570E-04	1.529E-02	8/39
8	Cell cycle_Sister chromatid cohesion	2.027E-04	1.675E-02	6/22
9	Signal transduction_Activation of PKC via G-Protein coupled receptor	2.418E-04	1.675E-02	9/52
10	Immune response_IL-15 signaling via JAK-STAT cascade	2.648E-04	1.675E-02	6/23

Figure 1

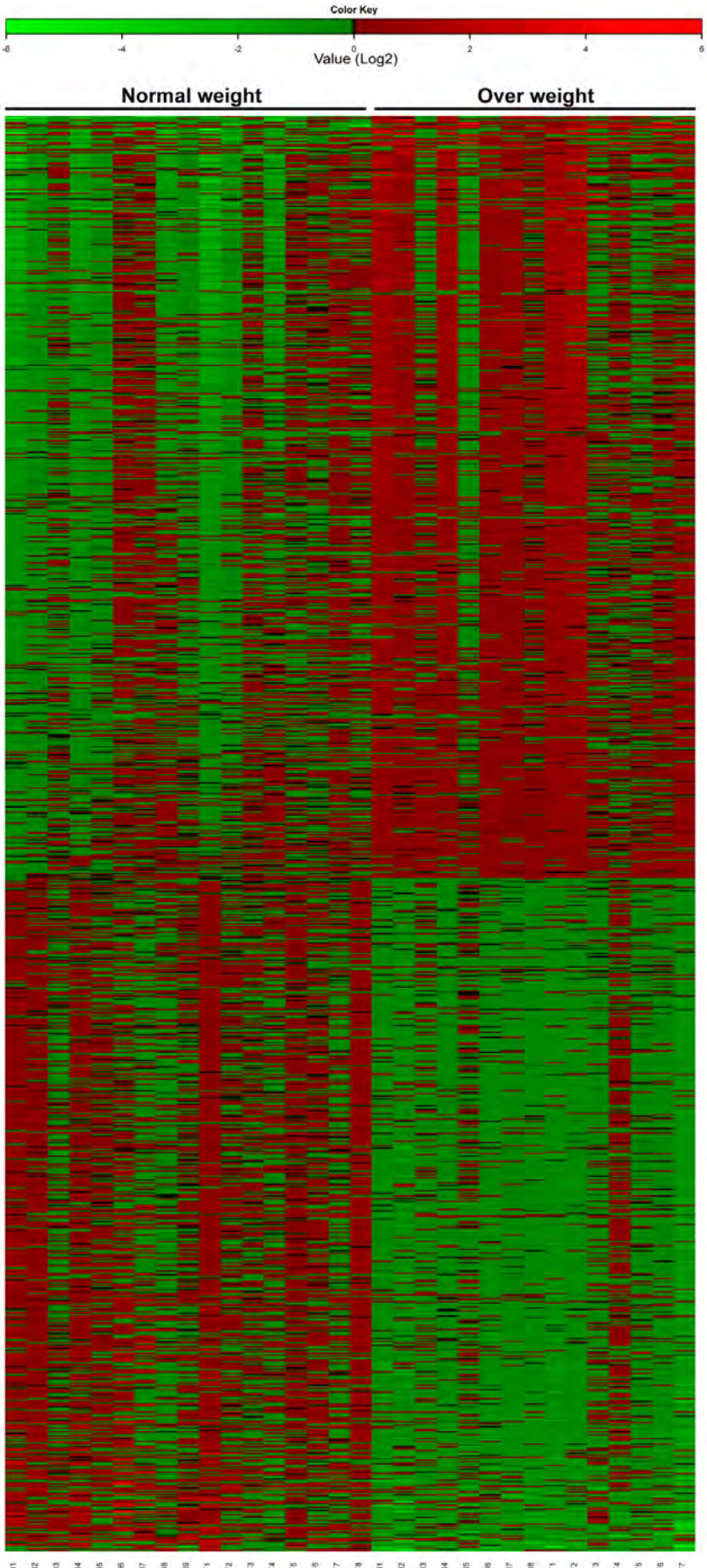
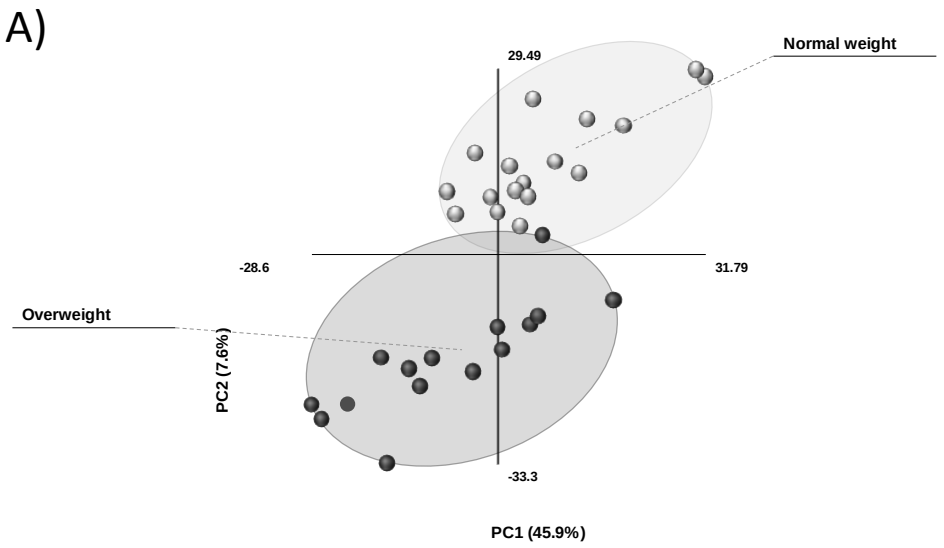
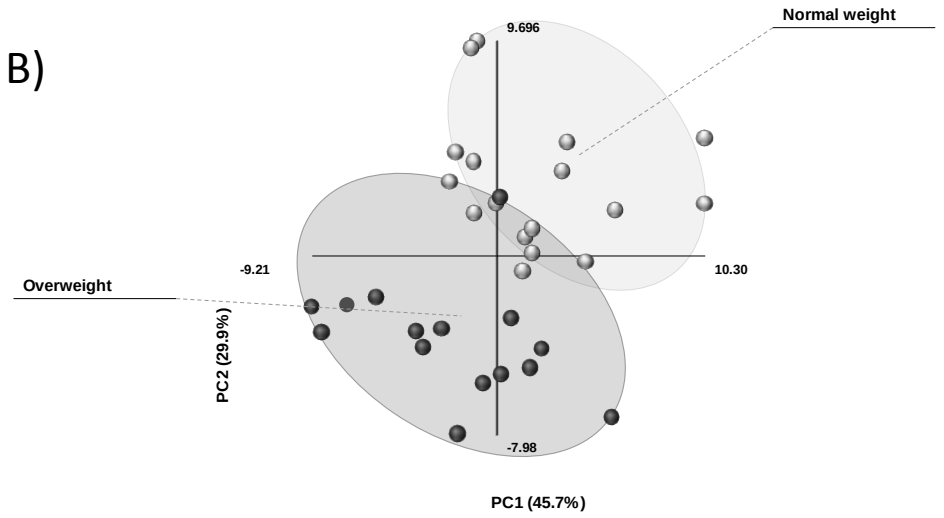


Figure 2





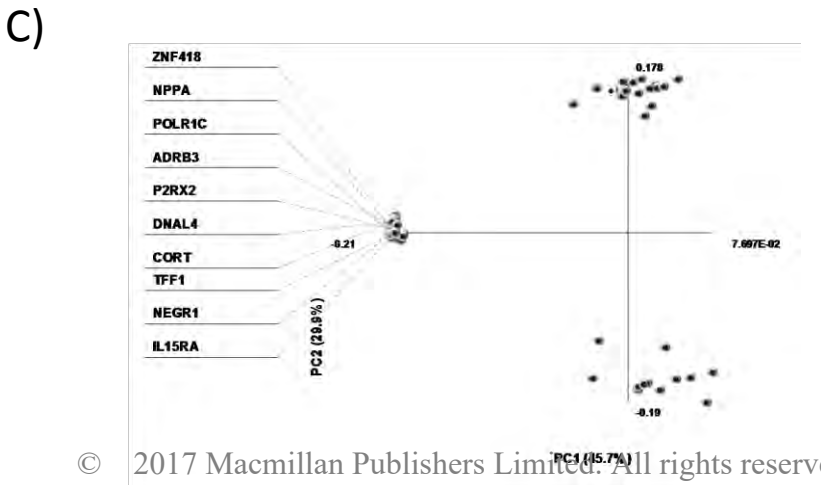


Figure 3

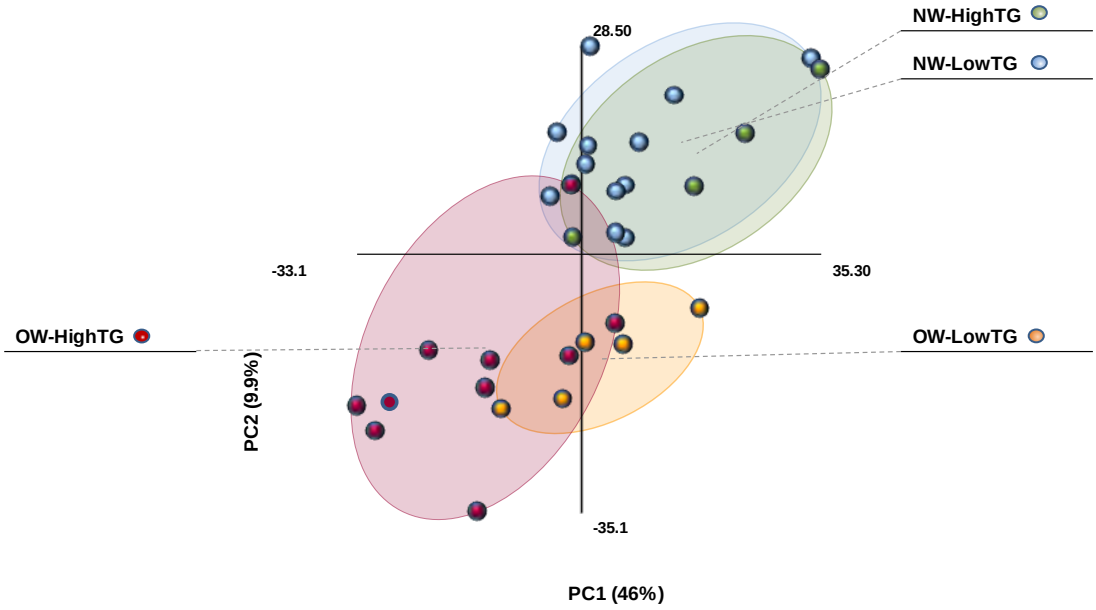


Figure 4