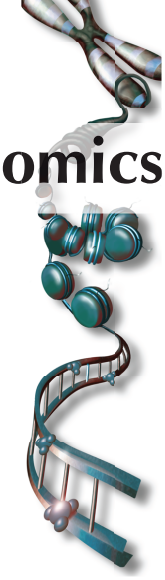
Research Article

For reprint orders, please contact: reprints@futuremedicine.com

Epigenomics



Differential long noncoding RNA profiling of BMI in twins

Weilong Li[1], Jan Baumbach[2,3], Martin J Larsen[4,5], Afsaneh Mohammadnejad[1], Jesper Lund[1], Kaare Christensen[1,5], Lene Christiansen[1,6] & Qihua Tan*[,1,5]

[1]Unit of Epidemiology, Biostatistics & Biodemography, Department of Public Health, University of Southern Denmark, Odense, Denmark

[2]Computational BioMedicine, Department of Mathematics & Computer Science, University of Southern Denmark, Odense, Denmark

[3]Chair of Experimental Bioinformatics, TUM School of Life Sciences, Technical University of Munich, Munich, Germany

[4]Department of Clinical Genetics, Odense University Hospital, Odense, Denmark

[5]Unit of Human Genetics, Department of Clinical Research, University of Southern Denmark, Odense, Denmark

[6]Department of Clinical Immunology, Copenhagen University Hospital, Rigshospitalet, Copenhagen, Denmark

*Author for correspondence: Tel.: +45 6550 3536; qtan@health.sdu.dk

Aim: Many efforts have been deployed to identify genetic variants associated with BMI. Alternatively, we explore epigenetic contribution to BMI variation by focusing on long noncoding RNAs (lncRNAs) which represents a key layer of epigenetic control. **Materials & methods:** We analyzed lncRNA expression in whole blood of 229 monozygotic twin pairs in association with BMI using generalized estimating equations. **Results & conclusion:** Six lncRNA probes were identified as significant (false discovery rate <0.05), with BMI showing causal effects on the expression of the significant lncRNAs. Functional annotation of differential profiles identified Gene ontology biological processes including kidney development, regulations of lipid biosynthetic process, circadian rhythm, notch signaling, etc. Whole blood lncRNAs are significantly expressed in response to BMI variation.

First draft submitted: 22 January 2020; Accepted for publication: 3 June 2020; Published online: 9 September 2020

Keywords: BMI • causal inference • lncRNA • twin study

The BMI quantifies the amount of tissue mass including muscle, fat and bone in an individual, and can be an indicator of high body fatness. It is calculated as weight in kg divided by the square of height in m (kg/m^2) and it is a simple and inexpensive surrogate measurement for body mass and obesity. BMI is connected to diabetes and cardiovascular diseases which have impact on life quality and mortality [1–3].

Long noncoding RNA (lncRNA) transcripts are longer than 200 nucleotides and are not translated into protein [4]. They have a large range of functions: molecular and biochemical mechanisms, from *cis*- to *trans*-regulation of gene expression, and from epigenetic modulation in the nucleus to post-transcriptional control in the cytoplasm [5]. There are a large number of traditional epidemiology studies focusing on the association between BMI and individual's occupation [6], incomes [7] and many other factors, and many studies, including genome-wide association studies and epigenome-wide association studies (EWAS) have been carried out on BMI [8–11], very few have studied the roles of lncRNA in the regulation of BMI [12,13].

Monozygotic (MZ) twins represent an ideal population for controlling the genetic background in investigating epigenetic associations, as they share similar genetic makeups [14]. We have demonstrated the massive value of the twin design in previous studies and showed a significantly higher statistical power over traditional case–control design using unrelated individuals in complex disease studies [15]. Moreover, the use of twins also enables inference on causal relationships in epigenetic association studies using a cross-sectional setup [16].

To utilize the power of twin design in epigenetic twin studies, the phenotype under investigation is preferable with modest high heritability as we have showed in our previous simulation study [15]. Twin studies have estimated high heritability of up to 60% indicating a large contribution of genetic factors to the variation of BMI [17], even though it is argued that the variance explained by genetic factors is prone to overestimated in twin studies [18]. In

Future Medicine

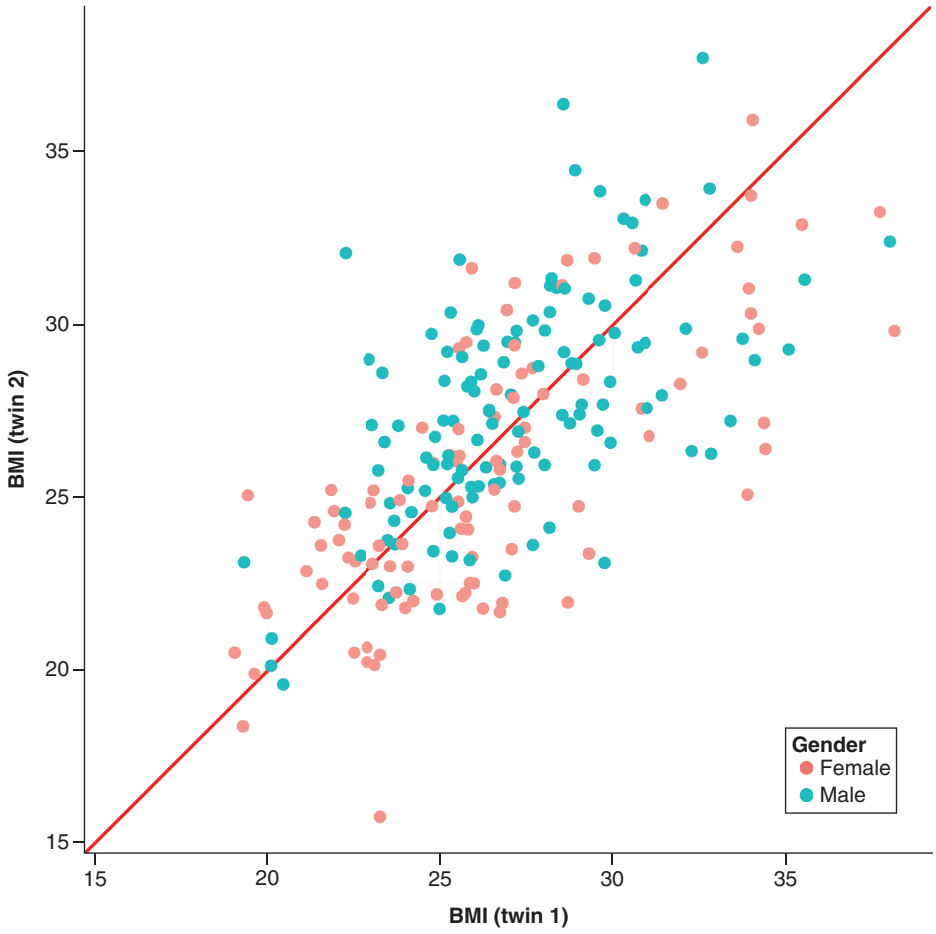


Figure 1. Scatter plot displaying intrapair BMI discordance by plotting BMI measurements for twin 1 on x-axis & for twin 2 on y-axis. All twin pairs distance themselves from the diagonal line of equal BMI.

fact, only a limited proportion of BMI variation (2–21%) was explained by current genome-wide association studies (GWAS) [19–21]. Recently, this estimate has been updated to 40% with whole genome sequencing data assuming additive genetic effects [22]. On the other hand, the epigenetic contribution to BMI variation on top of genetics could represent an important layer of molecular mechanism implicated in BMI development, as a considerable proportion of BMI variation is due to nongenetic factors. This study applies the powerful twin designs in studying the association between lncRNA expression and BMI in MZ twin pairs to identify significant and causal regulatory molecular markers of BMI followed by functional annotations.

Materials & methods

Samples

This study used a dataset of the Middle Aged Danish Twins (MADT) study from the Danish Twin Register [23]. There were 220 complete MZ twin pairs included in this analysis, and whole blood samples were taken over the period from 2008 to 2011. Samples include 242 males and 198 females (Figure 1). The age ranges from 56 to 80 and BMI ranges from 15.77 to 38.15 across the samples. The mean and standard deviation for BMI are 66.49 and 6.12, respectively. Blood cell counts are available for all the samples.

RNA extraction & gene expression analysis

Whole blood was collected in PAXgene Blood RNA tubes (PreAnalytiX GmbH; Hombrechtikon, Switzerland) and total RNA was extracted using the PAXgene Blood miRNA kit (QIAGEN, Hilden, Germany) according to the manufacturer's protocol. The concentration of the extracted RNA was determined using a NanoDrop

spectrophotometer ND-8000 (NanoDrop Technologies, DE, USA), and the quality was assessed by the Agilent 2100 Bioanalyzer (Agilent Technologies, CA, USA).

Gene expression analysis was performed using the Agilent SurePrint G3 Human GE 8 × 60 K Microarray (Agilent Technologies). Sample labeling, hybridization, washing, scanning and quantification were performed according to the array manufacturer's recommendations.

Data preprocessing

We used similar data preprocessing procedure as described in our previous publication [24]. Briefly, the array images were analyzed using Agilent Feature Extraction software v. 10.7.3.1 (Agilent Technologies), and raw intensity was background corrected using the NormExp method; normalized was performed using loess normalization method for within-array normalization and quantile normalization for between-array normalization [25–27]. Prior to statistical analysis, the coefficient of variation (CV) was calculated for each probe, and probes with CV < 0.1 were excluded in the further analysis. All the probes on the Agilent SurePrint G3 array were re-annotated (hg19) using GENCODE v.25 gene annotation database (www.gencodegenes.org).

Statistical analysis

The normalized expression data was first adjusted for cell compositions by using linear regression. Then the association of lncRNA expression and BMI was tested using generalized estimating equations method, with family as cluster, adjusting age and sex. The analysis was carried out using *geeglm()* function from R package *geepack* [28]. The model tests if the coefficient for BMI is different from zero (the null hypothesis) with a significant positive/negative coefficient indicating positive/negative correlation between expression/activity of a lncRNA and BMI measurement. False discovery rate (FDR) was calculated to account for multiple testing [29]. FDR < 0.05 was used as a cut off for defining significance.

Casual inference

For lncRNAs showing significant association with BMI (FDR < 0.05), causal relationship with BMI was investigated by the Inference about Causation through Examination of Familial Confounding method, which is a regression-based method for causal inference in association studies using twins or family data [16,30–34]. In their approach, 'familial' means factors (both genetic and environmental) shared by relatives, which is essential for Inference about Causation through Examination of Familial Confounding to make explicit causal inference. As described by Li *et al.* [16], inference on causal relationship from X to Y can be made by examining the changes in the correlation in the following models:

$$E\left(Y_{self}\right) = \alpha + \beta_{self} X_{self} \quad \text{(Eq. 1)}$$

$$E\left(Y_{self}\right) = \alpha + \beta_{co-twin} X_{co-twin} \quad \text{(Eq. 2)}$$

$$E\left(Y_{self}\right) = \alpha + \beta'_{self} X_{self} + \beta'_{co-twin} X_{co-twin} \quad \text{(Eq. 3)}$$

By adjusting $X_{co\text{-}twin}$ in Equation 3: if the association is due to the family confounding, both X_{self} and $X_{co\text{-}twin}$ would have nearly same effects on Y_{self}, and a similar extent of decrease would be observed in $\beta'_{co\text{-}twin}$ and β'_{self} in comparison to $\beta_{co\text{-}twin}$ and β_{self}, respectively. Thus, a small $|\beta_{co\text{-}twin} - \beta'_{co\text{-}twin}|$ in comparison with $|\beta_{self} - \beta'_{self}|$ would be observed; and if the association is causal with X causing changes in Y, the correlation in Equation 2 would be negated after conditioning on X_{self}, in Equation 3, in other words, large $|\beta_{co\text{-}twin} - \beta'_{co\text{-}twin}|$ in comparison with $|\beta_{self} - \beta'_{self}|$. In brief, if the associations between Y_{self} and the predictors, X_{self} and $X_{co\text{-}twin}$, remain unchanged before and after adjusting for each other, then no evidence of causal relationship is given. On the other hand, if there is a significant attenuation of $X_{co\text{-}twin}$ association after conditioning on X_{self}, there is an evidence consistent with X causing Y. Likewise, the generalized estimating equations model was applied for parameter estimation with twin pairs set as clusters.

Functional annotations

LncRNA probes with p-value < 0.01 have been passed to Genomic Regions Enrichment of Annotations Tool [35] to map the binding sites of the differentially expressed lncRNAs to the regulatory domain of the human genome (UCSC.hg19) to identify enrichments in biological processes and functional clusters implicated in BMI.

Table 1. Top 20 probes from lncRNA analysis using generalized estimating equation.

Probe ID	Gene symbol	Genomic coordinate (hg19)	Estimate	p-value	FDR
A_33_P3815064	*LOC283575*	chr14:77535783-77535842	-0.021	1.23E-07	0.001
A_21_P0011999	*LINC00570*	chr2:11542175-11542234	0.046	1.93E-07	0.002
A_19_P00318883	*APTR*	chr7:77314638-77314579	-0.009	1.37E-06	0.012
A_21_P0006309	*lnc-AL162389.1-1*	chr9:110477317-110477258	0.010	2.98E-06	0.027
A_24_P17870	*HCP5*	chr6:31432645-31432704	-0.020	3.23E-06	0.029
A_21_P0000382	*SNORD71*	chr16:71792377-71792318	0.031	4.63E-06	0.042
A_32_P115558		chr17:016285923-016285864	0.025	6.16E-06	0.056
A_21_P0008280	*LOC102724190*	chr14:77535582-77535523	-0.015	7.77E-06	0.071
A_21_P0008448	*LOC101927856*	chr14:65170570-65170511	-0.011	1.89E-05	0.172
A_33_P3298750		chr20:47319537-47319596	0.013	1.96E-05	0.179
A_21_P0000223	*SNORD15B*	chr11:75115551-75115610	0.041	2.16E-05	0.197
A_33_P3401284	*RMRP*	chr9:35657998-35657939	0.027	2.35E-05	0.215
A_21_P0000618	*EIF1B-AS1*	chr3:40214727-40214668	-0.008	2.36E-05	0.216
A_21_P0014322	*LOC101928595*	chr16:30116145-30116204	-0.014	3.20E-05	0.292
A_21_P0000334	*SNORA49*	chr12:132515831-132515890	0.013	3.43E-05	0.313
A_21_P0004567	*lnc-ARRDC3-1*	chr5:90607514-90606842	-0.015	3.63E-05	0.331
A_21_P0000489	*SNORD104*	chr17:62223458-62223517	0.026	3.64E-05	0.331
A_21_P0000351	*SCARNA22*	chr4:1976428-1976487	0.033	4.14E-05	0.378
A_33_P3329462	*DLEU1-AS1*	chr13:51095345-51095286	0.012	4.24E-05	0.386
A_33_P3282394	*MLLT1*	chr19:6210517-6210458	-0.014	4.47E-05	0.408

FDR: False discovery rate.

Results

Of the 14,832 lncRNA probes measured on the microarray, 438 probes are on the sex chromosomes and 5260 probes with CV < 0.1 (house-keeping lncRNAs). After removal of sex chromosome probes and invariant probes, a total of 9134 probes remained for subsequent analysis. Statistical testing detected 1374 probes with p-value < 0.05 and 575 probes with p-value < 0.01. The top 20 probes (ranked by p-value) are shown in Table 1, and the complete result is provided in Supplementary Table 1. There are six probes with FDR < 0.05, and they are annotated to *LOC283575* (*p* 1.23E-07), *LINC00570* (*p* 1.93E-07), *APTR* (*p* 1.37E-06), *lnc-AL162389.1-1* (*p* 2.98E-06), *HCP5* (*p* 3.23E-06) and *SNORD71* (*p* 4.63E-06), respectively.

The results of casual inference on the six significant probes are provided in Supplementary Table 2. Figure 2 visualizes the results of β changes (both self change and co-twin change) for following assumptions: lncRNA expression causes BMI in Figure 2A and BMI causes lncRNA expression in Figure 2B. The permutation p-values for β changes under the two assumptions are shown in Supplementary Figure 1 where the changes in $\beta_{\text{co-twin}}$ are all significant with p-value < 0.05 while no change in β_{self} is significant. Interestingly, a causal effect of BMI on lncRNA expression is clearly supported by the patterns in Figure 2B for all the six significant lncRNA probes (large changes in $\beta_{\text{co-twin}}$, small changes of nearly zero in β_{self}) while the pattern in Figure 2A does not support a causal relationship from lncRNA expression to BMI (changes in $\beta_{\text{co-twin}}$ and in β_{self} are nearly equal).

Results from annotation of probes with p-value < 0.01 using Genomic Regions Enrichment of Annotations Tool are displayed in Table 2, and corresponding acyclic graph is in Figure 3. There are 21 Gene ontology (GO) biological processes identified with FDR < 0.05. Most of the GO biological processes are related to kidney including metanephric nephron development, renal tubule development, loop of Henle development, glomerulus vasculature development and cell differentiation involved in kidney development. Other GO biological processes are either directly or indirectly related to obesity such as to regulation of lipid biosynthetic process, phagocytosis, regulation of circadian rhythm, positive regulation of notch signaling pathway, negative regulation of auditory receptor cell differentiation, retinal blood vessel morphogenesis and pharyngeal system development.

Discussion

By controlling genetic effects on BMI through twin design, we have performed an exploratory analysis on the association between lncRNA expression and BMI, identifying six significant probes after correcting for multiple

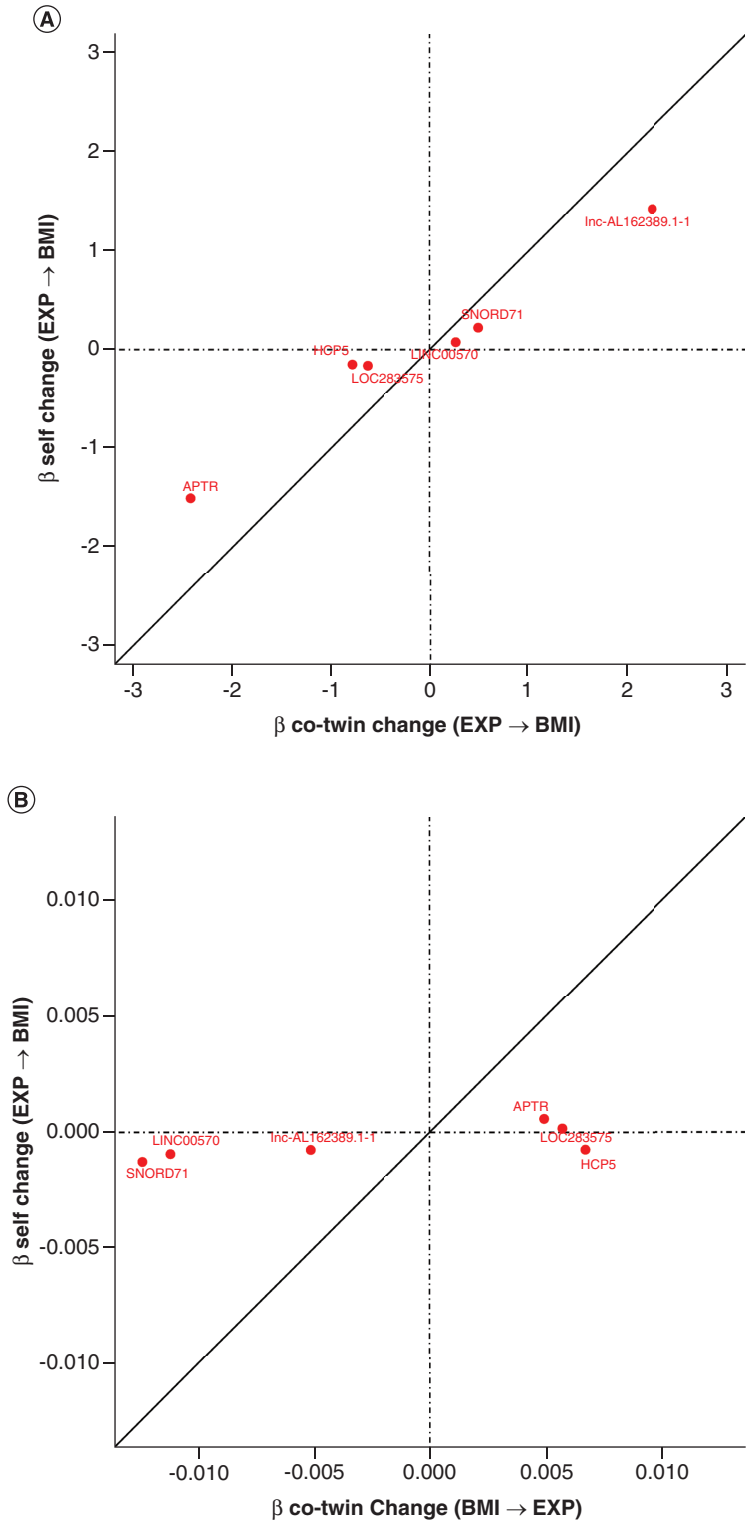


Figure 2. Comparison between β co-twin change ($|\beta_{co\text{-}twin} - \beta'_{co\text{-}twin}|$) & β self change ($|\beta_{self} - \beta'_{self}|$). **(A)** Outcome Y = BMI, and exposure X = expression. **(B)** Outcome Y = expression, and exposure X = BMI. β_{self} is the estimation of the overall correlation including the casual proportion and family confounding proportion (Equation 1); $\beta_{co\text{-}twin}$ estimates only the family confounding proportion of the correlation (Equation 2); β'_{self} and $\beta'_{co\text{-}twin}$ are estimation of full model (Equation 3) combing Equation 1 and Equation 2. If $|\beta_{co\text{-}twin} - \beta'_{co\text{-}twin}|$ is similar to $|\beta_{self} - \beta'_{self}|$, then the association is due to family confounding (Figure 2A). If $|\beta_{co\text{-}twin} - \beta'_{co\text{-}twin}|$ is much larger than $|\beta_{self} - \beta'_{self}|$ (ratio >1.5), then it indicates a causal effect (Figure 2B). In the figures, we observed that BMI changes cause the lncRNA expression change. EXP: Gene expression.

testing. From the causal inference analysis, we found that all six significant lncRNA probes are expressed in response to BMI changes, but not *vice versa*. Among the six significant lncRNAs, *LOC283575* harbors SNP *rs1986116* which has been found as one of the independent SNPs most highly associated with sleep quality [36] and which could indirectly affect BMI [37]. *LINC00570* was found downregulated in osteoarthritis patients who had a high BMI [38]. Gene *APTR* was upregulated in fibrotic liver samples and activated hepatic stellate cells, which was indicated

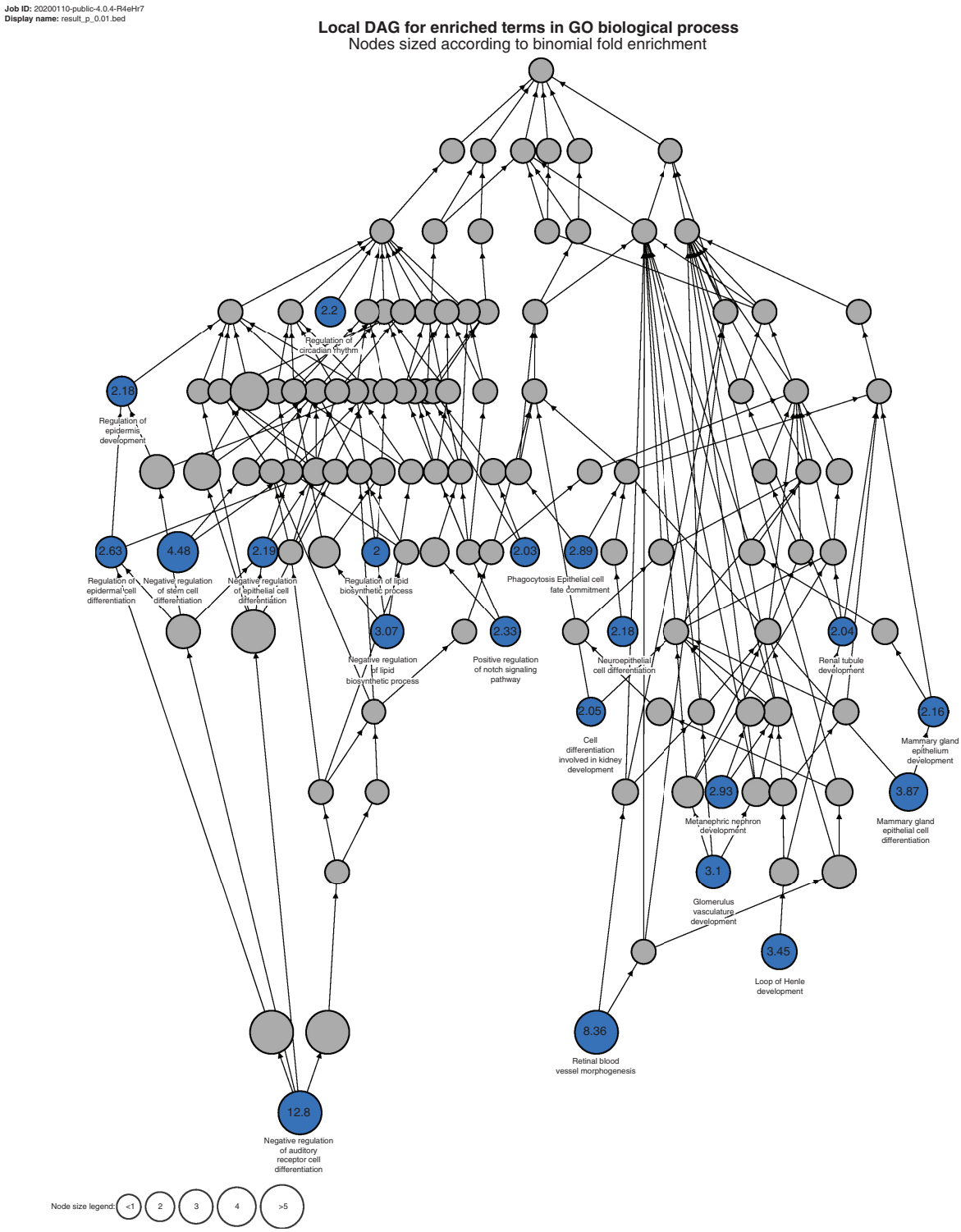


Figure 3. Directed acyclic graph generated from Genomic Regions Enrichment of Annotations Tool based on the enriched terms. It is the representation of hierarchy of Gene ontology database. Enriched terms are shown in blue, and their ancestor terms are shown in grey. Nodes have been sized according to binomial fold enrichment. DAG: Directed acyclic graph; GO: Gene ontology.

Table 2. Significant functional clusters biological process identified by Genomic Regions Enrichment of Annotations Tool.

Term name	Binom raw p-value	Binom FDR q-val	Binom fold enrichment	Binom observed region hits
Negative regulation of auditory receptor cell differentiation	6.05E-08	1.66E-05	12.797	9
Metanephric nephron development	3.67E-07	8.05E-05	2.926	30
Negative regulation of stem cell differentiation	1.13E-06	1.85E-04	4.481	16
Mammary gland epithelial cell differentiation	1.91E-06	2.92E-04	3.873	18
Renal tubule development	2.84E-06	3.93E-04	2.035	51
Phagocytosis	3.15E-06	4.27E-04	2.028	51
Regulation of epidermal cell differentiation	4.40E-06	5.36E-04	2.630	29
Negative regulation of lipid biosynthetic process	6.17E-06	6.99E-04	3.066	22
Regulation of circadian rhythm	7.35E-06	8.05E-04	2.201	39
Regulation of lipid biosynthetic process	1.60E-05	1.59E-03	2.001	45
Regulation of epidermis development	5.50E-05	4.46E-03	2.181	32
Loop of Henle development	8.71E-05	6.23E-03	3.445	14
Mammary gland epithelium development	8.80E-05	6.22E-03	2.155	31
Retinal blood vessel morphogenesis	1.02E-04	6.95E-03	8.363	6
Epithelial cell fate commitment	3.20E-04	1.61E-02	2.893	15
Neuroepithelial cell differentiation	3.35E-04	1.68E-02	2.183	25
Glomerulus vasculature development	4.10E-04	1.98E-02	3.104	13
Cell differentiation involved in kidney development	6.47E-04	2.77E-02	2.050	26
Negative regulation of epithelial cell differentiation	7.07E-04	2.91E-02	2.190	22
Positive regulation of notch signaling pathway	7.90E-04	3.07E-02	2.328	19
Pharyngeal system development	1.48E-03	4.70E-02	2.322	17

FDR: False discovery rate.

as a potential biomarker for liver cirrhosis [39], and liver cirrhosis is highly associated with obesity [40]. *HCP5* is an important gene that associated with AIDS, and it was found hypermethylated in obese Ghanaian African migrants [41]. During adipocyte differentiation, *SNORD71* was found upregulated when compared with obese over lean individuals, and it is more expressed in both pre-adipocytes and subcutaneous fat tissue [42].

The enriched GO biological processes suggest that lncRNAs are involved in kidney development associated with BMI. One of the strong indicators of chronic kidney disease is high BMI. A compensatory hyperfiltration occurs to meet the heightened metabolic demands for obese individuals [43]. It has been shown that incidence of chronic kidney disease increases with higher BMI in large cohort studies [44,45]. There are several renal alterations and impairments that are related to obesity: increased kidney weight, glomerular hypertrophy, tubular hypertrophy, hemodynamic changes, increased salt sensitivity, renin–angiotensin–aldosterone system activation as well as changes of glucose metabolism and adipose-derived inflammation and deposition of lipid components [46–48].

Our results from over-representation analysis suggest that the lipid biosynthetic process related to obesity could be medicated by lncRNA expression. Lipids are essential for many vital functions such as storing energy and forming cell membranes. However, exceeding lipid could result in high body fat such as triglycerides, which is the main form of stored energy and main constituents of body fat [49–51]. It is thus highly sensible that our identified lncRNA genes and GO biological processes are frequently linked to internal organs such as the liver and kidneys as visceral fat content is strongly associated with metabolic disorders [52].

Other GO biological processes are also highly related to obesity. Many studies have reported that circadian rhythm such as abnormal sleep/wake patterns are associated with obesity since sleep is an important modulator of neuroendocrine function and glucose metabolism [53–55]. Our finding suggests that lncRNA could also be involved in regulation of circadian rhythm. Phagocytosis is a critical part of the immune system, and obesity and diabetes could

decrease phagocytosis capacities [56–58]. Notch signaling is crucial for cell–cell communication and development, and it has been found important for metabolism that improves glucose tolerance, insulin sensitivity and ameliorates obesity and atherosclerosis [59]. There are also reports of changes in auditory [60–62], sight [63,64] and pharyngeal function [65,66] that relate to high BMI or obesity.

The top significant probes are all inferred as differentially expressed in response to BMI but not as causing changes in BMI. Interestingly, by causal inference on the association between peripheral blood DNA methylation and BMI, Li *et al.* [16] reported that BMI has a causal effect on DNA methylation which means methylation changes in response to BMI. Although our study focused on a different epigenetic mechanism (lncRNA regulation of gene activity), the two approaches both revealed similar nature of causal relationship between BMI and epigenetic regulation in peripheral blood cells.

Even with the powerful twin design, we identified only six significant lncRNAs associated with BMI after FDR correction. Although the heritability for BMI might be overestimated, still the moderate–high heritability estimates indicate that a large proportion of BMI variation could be explained by DNA sequence variations [20]. Meanwhile, different from DNA methylation, another epigenetic mechanisms that are well studied, to what extend that lncRNA is involved, and the function of lncRNA in regarding to BMI variation is still unknown and requires further investigation. Nevertheless, our exploratory analysis identified lncRNAs associated with BMI in MZ twins. Annotations based on regulatory domain were able to reveal significant biological processes potentially implicated in BMI and obesity. Importantly, the use of twins enables causal inference on significant findings in cross-sectional studies.

Conclusion

Our differential expression profiling in whole blood of MZ twin pairs identified six significant lncRNAs in association with BMI, all expressed in response to changes in BMI. Functional analysis showed significant enrichment of GO biological processes involving kidney development, regulations of lipid biosynthetic process, etc. Overall, the lncRNAs represent alternative epigenetic markers associated with BMI.

Future perspective

In this study we used whole blood samples as sources for lncRNA expression profiling. Twin-based lncRNA profiling on BMI using other relevant tissues (muscles or adipose) should help with validating and verifying our results. Furthermore, future studies correlating lncRNA profiling with genetic sequence variations and gene expression that integrates multiple omics would be valuable to investigate the functions of lnRNA in relation to BMI.

Summary points

- Whole blood samples from 220 complete monozygotic twin pairs were used in this study.
- Agilent SurePrint G3 Human GE 8 × 60 K Microarray was used for long noncoding RNA (lncRNA) profiling.
- Discordant twin design was applied to investigate the association between lncRNA expression and BMI variation.
- Causal inferences were made using Inference about Causation through Examination of Familial Confounding method for the top significant lncRNA probes.
- Result from causal inference suggests that six significant lncRNA probes are expressed in response to BMI changes.
- LncRNAs are involved in kidney development associated with BMI.
- The lipid biosynthetic process related to obesity could be medicated by lncRNA expression.
- Gene ontology biological processes including regulation of circadian rhythm, regulation of notch signaling were identified.

Supplementary data

To view the supplementary data that accompany this paper please visit the journal website at: www.futuremedicine.com/doi/suppl/10.2217/epi-2020-0033

Financial & competing interests disclosure

This project was supported by the EFSD/CDS/Lilly Collaborative Diabetes Research Programme (2013), the Lundbeck Foundation (grant number R170-2014-1353); the DFF research project 1 from the Danish Council for Independent Research, Medical Sciences (DFF-FSS) (grant number DFF – 6110-00114); the Novo Nordisk Foundation Medical and Natural Sciences Research Grant (grant

number NNF13OC0007493) and by the National Natural Science Foundation of China (grant number 81773506). J Baumbach's work was support by VILLUM Young Investigator (grant number 13154) and H2020 project REPO-TRIAL (grant number 777111). The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

No writing assistance was utilized in the production of this manuscript.

References

Papers of special note have been highlighted as: • of interest; •• of considerable interest

1. Di Cesare M, Bentham J, Stevens GA *et al.* Trends in adult body-mass index in 200 countries from 1975 to 2014: a pooled analysis of 1698 population-based measurement studies with 19.2 million participants. *Lancet* 387(10026), 1377–1396 (2016).
2. Gaziano TA, Opie LH. Body-mass index and mortality. *Lancet* 374(9684), 113–114 (2009). www.ncbi.nlm.nih.gov/pubmed/19595336
3. De Gonzalez AB, Hartge P, Cerhan JR *et al.* Body-mass index and mortality among 1.46 million white adults. *N. Engl. J. Med.* 363(23), 2211–2219 (2010).
4. Perkel JM. Visiting "Noncodarnia". *Biotechniques* 54(6), 301–304 (2013).

•• **An introduction of long noncoding RNA: its function and role in disease.**

5. Sun M, Kraus WL. From discovery to function: the expanding roles of long noncoding RNAs in physiology and disease. *Endocr. Rev.* 36(1), 25–64 (2015).
6. Eriksen D, Rosthøj S, Burr H, Holtermann A. Sedentary work-associations between five-year changes in occupational sitting time and body mass index. *Prev. Med.* 73, 1–5 (2015).
7. García Villar J, Quintana-Domeque C. Income and body mass index in Europe. *Econ. Hum. Biol.* 7(1), 73–83 (2009).
8. Locke AE, Kahali B, Berndt SI *et al.* Genetic studies of body mass index yield new insights for obesity biology. *Nature* 518(7538), 197–206 (2015).
9. Dick KJ, Nelson CP, Tsaprouni L *et al.* DNA methylation and body-mass index: a genome-wide analysis. *Lancet* 383(9933), 1990–1998 (2014).
10. Wilson LE, Harlid S, Xu Z, Sandler DP, Taylor JA. An epigenome-wide study of body mass index and DNA methylation in blood using participants from the Sister study cohort. *Int. J. Obes. (Lond.)* 41(1), 194–199 (2017).
11. Wahl S, Drong A, Lehne B *et al.* Epigenome-wide association study of body mass index, and the adverse outcomes of adiposity. *Nature* 541, 81–86 (2017).
12. Schmidt E, Dhaouadi I, Gaziano I *et al.* LincRNA H19 protects from dietary obesity by constraining expression of monoallelic genes in brown fat. *Nat. Commun.* 9, 3622 (2018).
13. Wei S, Du M, Jiang Z, Hausman GJ, Zhang L, Dodson M V. Long noncoding RNAs in regulating adipogenesis: new RNAs shed lights on obesity. *Cell. Mol. Life Sci.* 73(10), 2079–2087 (2016).

• **Reviews application of discordant twin design in twins.**

14. Tan Q. Epigenetic epidemiology of complex diseases using twins. *Med. Epigenet.* 1(1), 46–51 (2013).

• **Shows power advantage in applying discordant twin design in epigenetic studies.**

15. Li W, Christiansen L, Hjelmborg J, Baumbach J, Tan Q. On the power of epigenome-wide association studies using a disease-discordant twin design. *Bioinformatics* 34(23), 4073–4078 (2018).

• **Explains the Inference about Causation through Examination of Familial Confounding method and its application for BMI using methylation data.**

16. Li S, Wong EM, Bui M *et al.* Inference about causation between body mass index and DNA methylation in blood from a twin family study. *Int. J. Obes. (Lond.)* 43(2), 243–252 (2019).

• **Discusses the recent development of estimating heritability of BMI by analyzing 40 twin cohorts.**

17. Silventoinen K, Jelenkovic A, Sund R *et al.* Differences in genetic and environmental variation in adult BMI by sex, age, time period, and region: an individual-based pooled analysis of 40 twin cohorts. *Am. J. Clin. Nutr.* 106(2), 457–466 (2017).

•• **Discusses about heritability in twin study, GWAS and future challenge.**

18. Mayhew AJ, Meyre D. Assessing the heritability of complex traits in humans: methodological challenges and opportunities. *Curr. Genomics* 18(4), 332–340 (2017).
19. Bouret SG. Development of hypothalamic circuits that control food intake and energy balance. Chapter 7. In: *Appetite and Food Intake: Central Control (2nd Edition)*. FL, USA (2017).

20. Elks CE, Hoed den M, Zhao JH *et al.* Variability in the heritability of body mass index: a systematic review and meta-regression. *Front. Endocrinol. (Lausanne)* 3, 29 (2012).
21. Yengo L, Sidorenko J, Kemper KE *et al.* Meta-analysis of genome-wide association studies for height and body mass index in ~700000 individuals of European ancestry. *Hum. Mol. Genet.* 27(20), 3641–3649 (2018).

- **Recent development of GWAS that could potentially bridge the gap between heritability estimated from twin study and heritability estimated from GWAS using whole genome sequencing.**

22. Wainschtein P, Jain DP, Yengo L *et al.* Recovery of trait heritability from whole genome sequence data. *bioRxiv.* (2019).
23. Gaist D, Bathum L, Skytthe A *et al.* Strength and anthropometric measures in identical and fraternal twins: no evidence of masculinization of females with male co-twins. *Epidemiology* 11(3), 340–343 (2000).
24. Larsen MJ, Kruse TA, Tan Q *et al.* Classifications within molecular subtypes enables identification of BRCA1/BRCA2 mutation carriers by RNA tumor profiling. *PLoS ONE* 8(5), e64268 (2013).
25. Yang Y, Dudoit S, Luu P, Lin D. Normalization for cDNA microarray data: a robust composite method addressing single and multiple slide systematic variation. *Nucleic Acids Res.* 30(4), e15 (2002).
26. Yang YH, Thorne NP. Normalization for two-color cDNA microarray data. Science and Statistics: A Festschrift for Terry Speed, 40, 403–418 (2008).
27. Ritchie ME, Silver J, Oshlack A *et al.* A comparison of background correction methods for two-colour microarrays. *Bioinformatics* 23, 2700–2707 (2007).
28. Halekoh U, Højsgaard S, Yan J. The R package geepack for generalized estimating equations. *J. Stat. Softw.* 15(2), 1–11 (2006).
29. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Series B (Methodol.)* 57(1), 289–300 (1995).
30. Hopper JL, Bui QM, Erbas B *et al.* Does eczema in infancy cause hay fever, asthma, or both in childhood? Insights from a novel regression model of sibling data. *J. Allergy Clin. Immunol.* 130(5), 1117–1122 (2012).
31. Stone J, Dite GS, Giles GG, Cawson J, English DR, Hopper JL. Inference about causation from examination of familial confounding: application to longitudinal twin data on mammographic density measures that predict breast cancer risk. *Cancer Epidemiol. Biomarkers Prev.* 21(7), 1149–1204 (2012).
32. Bui M, Bjørnerem Å, Ghasem-Zadeh A, Dite GS, Hopper JL, Seeman E. Architecture of cortical bone determines in part its remodelling and structural decay. *Bone* 55(2), 353–361 (2013).
33. Dite GS, Gurrin LC, Byrnes GB *et al.* Predictors of mammographic density: insights gained from a novel regression analysis of a twin study. *Cancer Epidemiol. Biomarkers Prev.* 17(12), 3474–3481 (2008).
34. Davey CG, López-Solà C, Bui M *et al.* The effects of stress-tension on depression and anxiety symptoms: evidence from a novel twin modelling analysis. *Psychol. Med.* 46(15), 3213–3218 (2016).
35. McLean CY, Bristor D, Hiller M *et al.* GREAT improves functional interpretation of *cis*-regulatory regions. *Nat. Biotechnol.* 28(5), 495–501 (2010).
36. Byrne EM, Gehrman PR, Medland SE *et al.* A genome-wide association study of sleep habits and insomnia. *Am. J. Med. Genet. B Neuropsychiatr. Genet.* 162B(5), 439–490 (2013).

- **Reviews the relationship between short sleep duration, poor sleep quality and obesity.**

37. Beccuti G, Pannain S. Sleep and obesity. *Curr. Opin. Clin. Nutr. Metab. Care* 14(4), 402–412 (2011). https://insights.ovid.com/article/00075197-201107000-00016
38. Kuttapitiya A, Assi L, Laing K *et al.* Microarray analysis of bone marrow lesions in osteoarthritis demonstrates upregulation of genes implicated in osteochondral turnover, neurogenesis and inflammation. *Ann. Rheum. Dis.* 76(10), 1764–1773 (2017).
39. Yu F, Zheng J, Mao Y *et al.* Long non-coding RNA APTR promotes the activation of hepatic stellate cells and the progression of liver fibrosis. *Biochem. Biophys. Res. Commun.* 463, 679–685 (2015).
40. Chiang DJ, Pritchard MT, Nagy LE. Obesity, diabetes mellitus, and liver fibrosis. *Am. J. Physiol. Gastrointest. Liver Physiol.* 300(5), G697–702 (2011).
41. Meeks KAC, Henneman P, Venema A *et al.* An epigenome-wide association study in whole blood of measures of adiposity among Ghanaians: the RODAM study. *Clin. Epigenetics* 9, 103 (2017).
42. Ortega FJ, Moreno-Navarrete JM, Pardo G *et al.* MiRNA expression profile of human subcutaneous adipose and during adipocyte differentiation. *PLoS ONE* 5, e9022 (2010).
43. Kovesdy CP, Furth S, Zoccali C. Obesity and kidney disease: hidden consequences of the epidemic. *Iran. J. Kidney Dis.* 11(2), 99–108 (2017).
44. Chang Y, Ryu S, Choi Y *et al.* Metabolically healthy obesity and development of chronic kidney disease: a cohort study. *Ann. Intern. Med.* 164(5), 305–317 (2016).
45. Stengel B, Tarver-Carr ME, Powe NR, Eberhardt MS, Brancati FL. Lifestyle factors, obesity and the risk of chronic kidney disease. *Epidemiology* 14(4), 479–487 (2003).

46. Tsuboi N, Okabayashi Y, Shimizu A, Yokoo T. The renal pathology of obesity. *Kidney Int. Rep.* 2(2), 251–260 (2017). https://doi.org/10.1016/j.ekir.2017.01.007
47. Wisse BE. The inflammatory syndrome: the role of adipose tissue cytokines in metabolic disorders linked to obesity. *J. Am. Soc. Nephrol.* 15(11), 2792–2800 (2004).
48. Després JP, Lemieux I. Abdominal obesity and metabolic syndrome. *Nature* 444, 881–887 (2006).
49. Lamb MM, Ogden CL, Carroll MD, Lacher DA, Flegal KM. Association of body fat percentage with lipid concentrations in children and adolescents: United States, 1999–2004. *Am. J. Clin. Nutr.* 94, 887 (2011).
50. Arrese EL, Soulages JL. Insect fat body: energy, metabolism, and regulation. *Annu. Rev. Entomol.* 55, 207–225 (2009).
51. Walton C, Lees B, Crook D, Godsland IF, Stevenson JC. Relationships between insulin metabolism, serum lipid profile, body fat distribution and blood pressure in healthy men. *Atherosclerosis* 118(1), 35–43 (1995).
52. Kitchlew R, Chachar AZ, Latif S. Body mass index; visceral fat and total body fat distribution and its relation to body mass index in clinical setting using bioimpedance body composition monitor. *Prof. Med. J.* 24, 326–334 (2017).
53. Pagano ES, Spinedi E, Gagliardino JJ. White adipose tissue and circadian rhythm dysfunctions in obesity: pathogenesis and available therapies. *Neuroendocrinology* 104(4), 347–363 (2017).
54. Bray MS, Young ME. Circadian rhythms in the development of obesity: potential role for the circadian clock within the adipocyte. *Obes. Rev.* 8, 169–181 (2007).
55. Pivik R, Gu Y, Downs H *et al.* Autonomic activity in overweight-obese children during the morning transition in cardiac circadian regulation (P21-065-19). *Curr. Dev. Nutr.* 3, P21–065&–19 (2019).
56. Serena C, Keiran N, Ceperuelo-Mallafre V *et al.* Obesity and Type 2 diabetes alters the immune properties of human adipose derived stem cells. *Stem Cells* 34(10), 2559–2573 (2016).
57. Hellmann J, Zhang MJ, Tang Y, Rane M, Bhatnagar A, Spite M. Increased saturated fatty acids in obesity alter resolution of inflammation in part by stimulating prostaglandin production. *J. Immunol.* 19, 1383–1392 (2013).
58. Morceli G, Honorio-França AC, Fagundes DLG, Calderon IMP, França EL. Antioxidant effect of melatonin on the functional activity of colostral phagocytes in diabetic women. *PLoS ONE* 8(2), e56915 (2013).
59. Bi P, Kuang S. Notch signaling as a novel regulator of metabolism. *Trends Endocrinol. Metab.* 26(5), 248–255 (2015).
60. Hwang JH, Hsu CJ, Yu WH, Liu TC, Yang WS. Diet-induced obesity exacerbates auditory degeneration via hypoxia, inflammation, and apoptosis signaling pathways in CD/1 mice. *PLoS ONE* 8(4), e60730 (2013).
61. Dhanda N, Taheri S. A narrative review of obesity and hearing loss. *Int. J. Obes. (Lond.)* 41(7), 1066–1073 (2017).
62. Hu H, Tomita K, Kuwahara K *et al.* Obesity and risk of hearing loss: a prospective cohort study. *Clin. Nutr.* 39(3), 870–875 (2019).
63. Kurniawan ED, Cheung CY, Tay WT *et al.* The relationship between changes in body mass index and retinal vascular caliber in children. *J. Pediatr.* 165(6), 1166–1171 (2014).
64. Köchli S, Endes K, Infanger D, Zahner L, Hanssen H. Obesity, blood pressure, and retinal vessels: a meta-analysis. *Pediatrics* 141(6), e20174090 (2018).
65. Thapa A, Jayan B, Nehra K, Agarwal SS, Patrikar S, Bhattacharya D. Pharyngeal airway analysis in obese and non-obese patients with obstructive sleep apnea syndrome. *Med. J. Armed Forces India* 71, S369–S375 (2015). https://linkinghub.elsevier.com/retrieve/pii/S0377123714001142
66. Genta PR, Schorr F, Eckert DJ *et al.* Upper airway collapsibility is associated with obesity and hyoid position. *Sleep.* 37(10), 1673–1678 (2014).