Identification of novel nutrient-sensitive gene regulatory networks in amniotic fluid from fetuses with spina bifida using miRNA and transcription factor network analysis

Marina White[1], Jayden Arif-Pardy[1], *Kristin L Connor[1]

[1] Health Sciences, Carleton University, Ottawa, ON K1S 5B6, Canada
[*] Correspondence: kristin.connor@carleton.ca

Abstract

Background: Neural tube defects (NTDs) remain among the most common congenital anomalies. Contributing risk factors include genetics and nutrient deficiencies, however, a comprehensive assessment of nutrient-gene interactions in NTDs is lacking. We hypothesised that multiple nutrient-gene interactions would be evident in NTD-associated gene signatures.

Methods: We applied a novel, nutrient-focused gene expression analysis pipeline to identify nutrient-sensitive gene regulatory networks in amniocyte gene expression data (GSE4182) from fetuses with NTDs (cases; n=3) and fetuses with no congenital anomalies (controls; n=5). Differentially expressed genes (DEGs) were identified and screened for having nutrient cofactors. Transcription factors (TFs) with nutrient cofactors that regulated DEGs, and nutrient-sensitive miRNAs that had a previous link to NTDs, were identified and used to construct DEG regulatory networks.

Results: Of the 880 DEGs in cases (vs. controls), 10% had at least one nutrient cofactor. DEG regulatory network analysis revealed that 39% and 52% of DEGs in cases were regulated by 22 nutrient-sensitive miRNAs and 10 nutrient-dependent TFs, respectively. Zinc- and B vitamin- dependent genes and gene regulatory networks (Zinc: 10 TFs targeting 50.6% of DEGs; B vitamins: 4 TFs targeting 37.7% of DEGs, 9 miRNAs targeting 17.6% of DEGs) were dysregulated in cases. Two nutrient-dependent TFs predicted to target DEGs in cases (Tumor Protein 63 and Churchill Domain Containing 1) have not been previously linked to NTDs.

Conclusions: We identified multiple novel nutrient-sensitive gene regulatory networks associated with NTDs, which may relate to NTD pathogenesis, and indicate new targets to explore for NTD prevention or to optimise fetal development.

Keywords: Neural tube defects, spina bifida, nutrients, gene regulatory networks, network analysis

Introduction

Neural tube defects (NTDs) are among the most prevalent congenital anomalies, affecting around 19 per 10,000 livebirths globally[1,2]. Complex genetic and environmental contributions underly around 75% of NTDs, which have multifactorial causation[3,4] and can associate with multiple comorbidities, including placental maldevelopment and function[5], fetal growth restriction[6] and preterm birth[7]. Nutritional factors have a well-recognised contribution to the etiology of isolated NTDs, namely folate and vitamin B12[8], and to a lesser extent, inositol[9], vitamin D[10], choline[11] and essential trace elements[12,13]. Accordingly, widespread and targeted nutritional prevention strategies, such as fortification of food products with folic acid and periconceptional folic acid supplementation, have decreased NTD incidence[14]. However, that more than 300,000 infants remain affected by NTDs annually[1] highlights the importance of improving knowledge on their causation to inform further strategies for prevention.

The role of genetics in NTD etiology is supported by associations between NTDs and chromosomal anomalies, as well as heritable susceptibility to NTDs[15]. Distinct genetic signatures that associate with NTDs have also been revealed through large-scale genomic analysis[16-21], which can provide an integrated, global view of the mechanisms regulating normal and pathological development and aid in the identification of biomarkers for prevention and diagnosis[22]. Some research on genetic signatures in NTDs has been performed using amnioctyes[20,21], which are a heterogeneous population of fetal stem cells in the amniotic fluid, suggested to have some capacity to differentiate into multiple fetal cell lineages[23,24]. Importantly, amniocyte transcriptome (gene expression) profiling has detected the expression of genes with known tissue-specific expression in numerous fetal organ systems, including the brain and placenta[24], suggesting that amniocytes may be uniquely positioned to inform on genetic signatures in NTDs, as well as potentially placental-related comorbidities. While transcriptome profiling can help provide a gene-, pathway-, and network-level understanding of the genetic signatures that associate with NTDs, there is also a need to integrate knowledge on gene-environment, including nutrient, interactions to work towards a more complete understanding of NTD phenotypes, and to identify target candidates for prevention.

Adequate nutritional resources are essential for ensuring normal embryonic and fetoplacental growth and development[25]. Nutrients play important roles in DNA stability and repair and gene expression regulation[26], and host genetics influence nutrient metabolism and bioavailability[26]. When considering gene-nutrient interactions in NTDs, research to-date has largely focused on the contribution of genetic variants in the folate pathway to NTD risk[15,27]. However, knowledge is limited on whether, and the extent to which, *other* nutrients and their interactions with genes can explain NTD etiology. This is a critical gap to fill, given that fetal NTDs persist even in countries with folic acid food fortification[28], and in pregnant people with normal blood folate levels[29].

Comorbidities associated with NTDs, including altered placental phenotype[5], poor fetal growth[6], and early birth[7], may also be responsive to, or driven by, altered nutrient bioavailability[30,31]. Thus, understanding nutrient-gene interactions in NTDs is critical for improving knowledge on the pathogenesis of NTDs and their comorbidities. However, to date, no study has integrated transcriptome network analysis with field knowledge on nutrient-gene interactions in the context of NTDs.

Here, we developed and applied a novel nutrient-focused gene expression analysis pipeline to help address knowledge gaps on the role of nutrient-gene interactions in NTDs. Using data from a case-control study, we first aimed to examine the degree to which differentially expressed genes (DEGs) in fetuses with spina bifida (cases) were micronutrient dependent. We next aimed to discover novel nutrient-sensitive gene regulatory networks dysregulated in NTDs, through the identification of nutrient-responsive or dependent miRNAs and transcription factors (TFs) targeting co-regulated DEGs in cases. We hypothesised that multiple nutrient-gene interactions would be evident in gene signatures associated with NTDs, and that clear nutrient-sensitive miRNA and TF 'regulons' (a group of genes belonging to a common regulatory network) would emerge among DEGs.

Methods

Data source and pre-processing

Amniocyte gene expression data for fetuses with spina bifida (cases; n=4) and fetuses with no congenital anomalies (controls; n=5) were obtained from Gene Expression Omnibus (GEO; GSE4182) on June 28, 2021, and re-analysed. These microarray data were sequenced using the Affymetrix Human Genome U133 Plus 2.0 Array as part of a case-control study in Budapest, Hungary, and deposited in GEO in November 2006[19,21]. Available sample characteristics are provided in Supplementary Table S1. All three cases had a spina bifida diagnoses, and one case also had anencephaly. Gestational age at time of amniotic fluid collection ranged from 13-19 weeks for cases and 17-19 weeks and 5 days for the control group. Each control sample was pooled from two patient amniotic fluid samples from pregnancies undergoing routine amniocentesis because of a maternal age >35 years, and thus included amniotic fluid from 10 controls total.

All data pre-processing were performed in R software (version 3.6.1) using the Bioconductor (version 3.12) Affy package[32]. Raw microarray data (.CEL files) were downloaded from GEO and normalised using the robust multi-array average method, as recommended for Affymetrix microarrays[33]. Expression values were visualised using box and whisker plots and outliers were identified using sample clustering and uniform manifold approximation and projection (UMAP)[34]. Where multiple probesets corresponded to a single gene, a representative probeset was chosen for each gene, based on recommended methods[35] (Supplementary Table S2). In brief, preference was given to probesets where each probe in the set contains the sequence of a single known transcript, excluding probes that hit alternate transcripts from the same gene or transcripts from different genes. Where more than one probeset remained, the probeset with the highest absolute expression level across all samples was retained for further analyses. Individual probeset identifiers were annotated with Entrez gene IDs[36].

Differential gene expression analysis

DEGs between cases and controls were identified using the Limma package and empirical Bayes method in R (version 3.6.1)[37]. Statistical significance was determined after false discovery rate correction (Benjamini-Hochberg method[38]) at q value <0.05 and an absolute log fold change (FC)≥2. Functional analysis of the DEG list was performed using the Protein ANalysis THrough Evolutionary Relationships (PANTHER) classification system (version 16.0)[39]. While these data were first analysed to examine differential gene expression in cases compared to controls by the original authors[19], standard approaches to gene expression analysis have advanced in the time since, and it was necessary for us to generate a list of DEGs to use in our subsequent analyses.

Identifying nutrient cofactors of differentially expressed genes

To determine which of the DEGs coded for proteins that were micronutrient dependent (i.e., they coded for proteins with one or more micronutrient [vitamin or mineral] cofactor), we created a DEG-nutrient interaction network. Using a previously published dataset of protein/gene-nutrient (cofactor) interactions[40], we identified DEGs coding for a cofactor-interacting protein and their associated cofactors. This dataset on protein-nutrient interaction networks integrates knowledge on nutrient-protein (gene) interactions from the Metal MACiE25, Uniprot23, Expasy24 and EBI CoFactor databases. Applying a systems-approach to nutrient-gene interaction analysis is key for understanding the ways in which multiple micronutrients, their metabolic products, and interacting proteins function in biological processes, and for obtaining a more complete view of nutritional contributions to health and disease[40]. The PANTHER classification system (version 16.0) was then used to functionally annotate DEGs that were nutrient-dependent[39], and we visualised DEG-nutrient relations as a network using Cytoscape (version 3.8.2)[41].

Nutrient-sensitive miRNA targetome analysis

Next, to construct nutrient-sensitive miRNA regulatory networks for DEGs in cases, we first curated a list of nutrients that have known relevance to NTDs. This included cofactors for one carbon metabolism and nucleotide synthesis (choline[42], vitamins B1[43-45], B6[42,44,45], B9 [folate][44,45], and B12[42,44]), as well as inositol[46], iron[45], and vitamin D[10]. We next used miRWalk (version 2.0), a platform that integrates data on validated and predicted miRNA-target (gene) interactions, to identify miRNAs that are known to interact with biological processes related to these nutrients, or with activity known to be sensitive to these nutrients[47]. We then cross-referenced the list of miRNAs sensitive to, or associated with, these NTD-related nutrients against a curated inventory of miRNAs that have been associated with NTDs or neural tube closure in previous animal and human studies, which was also built using miRWalk and through manual searching in PubMed (search string: "neural tube*" AND "miRNA"). Targetome files with miRNA-target gene network data (miRNA targetomes), limited to experimentally validated miRNA-target interactions (miRTarBase)[47,48], were then retrieved for the miRNAs that were found in both lists (i.e., miRNAs whose activity was sensitive to the NTD-related nutrients, that also had a previous link to NTDs). Our rationale to limit our focus to miRNAs with a previous NTD link when constructing nutrient-sensitive miRNA regulatory networks was driven by the high number of miRNAs associated with the nutrients of interest. For example, approximately 1000 miRNAs were identified to be sensitive to each of choline, inositol and iron. Lastly, the targetome files were used to construct nutrient-sensitive miRNA-DEG regulatory networks in Cytoscape (version 3.8.2)[41].

Nutrient-sensitive transcription factor regulatory network analysis

Next, we used iRegulon (version 1.3) to identify transcription factors (TFs) predicted to co-regulate DEGs in cases[49]. Specifically, enriched TF binding motifs, or CHIP-seq peaks (derived from ChIP-seq data sets, also referred to as tracks), were

identified in the regulatory sequences in the 20 kb (kilobase; 1000 base pairs of DNA) surrounding the transcription start site (TSS) of each DEG (from −10 kb to +10 kb of the TSS)[49]. Candidate TFs associated with enriched motifs and tracks were identified using default search parameters and an enrichment score threshold of 3.0[49]. Two lists of candidate TFs were produced: One representing TFs targeting up-regulated genes in cases, and the other targeting down-regulated genes in cases. Within these lists, TFs with nutrient cofactors were identified using the same dataset of protein (gene)-nutrient interactions used above, to identify DEGs with nutrient cofactors[40]. Nutrient-sensitive TFs and their direct DEG targets were then exported as network files, separately for up- and down-regulated DEG lists. Nutrient-sensitive TF-DEG regulatory networks were visualised in Cytoscape (version 3.8.2)[41].

Results

Data pre-processing

Expression value distribution, density, and intensity were similar across all nine samples (Supplementary Figure S1A). One case sample (GSM94601) identified as an outlier in UMAP analysis was excluded (Supplementary Figure S1A), consistent with approaches taken in previous analyses of this dataset[20]. The remaining eight samples were then re-clustered using UMAP, and complete separation of the case and control groups was evident (Supplementary Figure S1B). Three cases and five control samples were retained for subsequent analyses.

Amniocyte gene expression differs in cases compared to controls

In a comparison of global amniocyte gene expression, 1,155 genes (2.11% of probesets sequenced) were differentially expressed between cases and controls ($q<0.05$; absolute log fold change ≥2). After the removal of duplicate probesets and genes with no available Entrez ID (Supplementary Figure S2 and Table S2), 880 DEG (downregulated: n=725, upregulated: n=155) were carried forward for further analyses (Figure 1).

After functional annotation of the DEGs using PANTHER (Supplementary Figure S3), the top identified GO molecular functions of DEGs in cases were binding (n=274 genes), catalytic activity (n=152) and regulatory functions (n=120). The top identified biological processes of DEGs in cases were cellular processes (n=431 genes), biological regulation (n=290), and metabolic processes (n=226). The top protein classes of DEGs were protein modifying enzymes (n=71 genes), gene-specific transcriptional regulators (n=63), and metabolite interconversion enzymes (n=49). The full DEG list is provided in Supplementary Table S3.

Novel nutrient-gene relationships identified in genes dysregulated in cases

We next investigated whether DEGs in case amniocytes coded for proteins that had nutrient cofactors. Of the DEGs in cases, 87 (9.9%) had at least one nutrient cofactor (downregulated: n=69, upregulated: n=18; Table 2). Cofactors included calcium (for n=14 genes), chloride (n=1), iron (n=6), heme (n=4), magnesium (n=22), manganese (n=4), S-Adenosyl methionine (n=3), vitamins B2 (n=4), B3 (n=7), B5 (n=2), B6 (n=1), and B9 (folate; n=1), vitamin C (n=1), zinc (n=19), and metal cations (n=9; Figure 2A). In nearly all DEG clusters interacting with a nutrient cofactor, the majority of DEGs were downregulated, apart from DEGs interacting with calcium (9 of 14 were upregulated) and folate (Folate receptor beta [*FOLR2*] was upregulated and the only folate dependent DEG; Figure 2A).

Functional annotation of nutrient-dependent DEGs in case amniocytes revealed the top GO molecular functions to be catalytic activity (n=46 genes), binding (n=20), and transporter and molecular transducer activity (n=4 each; Supplementary Figure S4). The top identified biological processes of DEGs with a nutrient cofactor were cellular process (n=51 genes), metabolic process (n=33), and biological regulation (n=33), and the top protein classes were metabolite interconversion enzymes (n=23 genes), protein modifying enzymes (n=14), and calcium-binding proteins (n=6).

The top dysregulated genes with a nutrient cofactor were involved in lipid and purine nucleotide metabolic processes, post-translational protein modifications, and cellular homeostasis and motility-related processes (Table 1). Specifically, the three most downregulated, nutrient-dependent genes in cases were Ethanolamine-Phosphate Phospho-Lyase (*ETNPPL*; vitamin B6-dependent; FC=-5.14, q=0.002), involved in lipid metabolism, Baculoviral IAP Repeat Containing 3 (*BIRC3*; zinc-dependent; FC=-3.40, q=0.002), associated with apoptosis inhibition, inflammatory signalling pathways, and cell proliferation, and Polypeptide N-Acetylgalactosaminyltransferase 3 (*GALNT3*; manganese-dependent; FC=-3.31, $q=4.57e^{-4}$), involved in O-linked oligosaccharide biosynthesis. The three most upregulated genes were Adenylosuccinate Synthase 1 (*ADSSL1*; magnesium-dependent; FC=3.23, q=0.009), which codes for an enzyme involved in purine nucleotide synthesis and metabolism, Matrix metallopeptidase 16 (*MMP16*; zinc- and calcium-dependent; FC=3.17, q=0.002), involved in extracellular matrix degradation in normal physiological and disease processes, and Ectonucleotide Pyrophosphatase/Phosphodiesterase 2 (*ENPP2*; zinc- and calcium-dependent; FC=2.83, q=0.006), associated with motility-related processes, including angiogenesis and neurite outgrowth.

Overall, nutrient-dependent, downregulated genes in cases were involved in immune-related processes and nutrient metabolism and homeostasis (Table 1). This included eight genes from the metal cation-binding Metallothionein (*MT*) gene family (*MT1E, MT1F, MT1G, MT1H, MT1HL1, MT1M, MT1X* and *MT2A*), associated with intracellular metal homeostasis, as well as 16 zinc-interacting genes involved in immune and inflammatory processes, antioxidant activity, mitochondrial morphology, the renin-angiotensin pathway, and cell signaling (Table 2). Downregulated genes coding for proteins with B vitamin cofactors included Riboflavin Kinase (*RFK*), which is essential for vitamin B2 utilization, *ETNPPL* and Glycerol-3-Phosphate Acyltransferase 3 (*AGPAT9*), which are involved in lipid metabolism, and Methylsterol Monooxygenase 1 (*MSMO1*) and Isopentenyl-Diphosphate Delta Isomerase 1 (*IDI1*), which are involved in the biosynthesis of cholesterol.

Upregulated genes in cases included those involved in immune and inflammatory processes, such as Arachidonate 5-lipoxygenase (*ALOX5*; calcium- and iron-dependent), Colony Stimulating Factor 1 Receptor (*CSF1R*; magnesium-dependent), *FOLR2* (folate-dependent), and NLR Family Apoptosis Inhibitory Protein (*NAIP*; zinc-dependent; Table 2). Calcium-dependent Neuropilin 1 (*NRP1*) and Ectonucleotide Pyrophosphatase/Phosphodiesterase 2 (*ENPP2*), and iron- and heme-dependent Prostaglandin-Endoperoxide Synthase 1 *(PTGS1*), were also upregulated in cases and are involved in angiogenesis (Table 1).

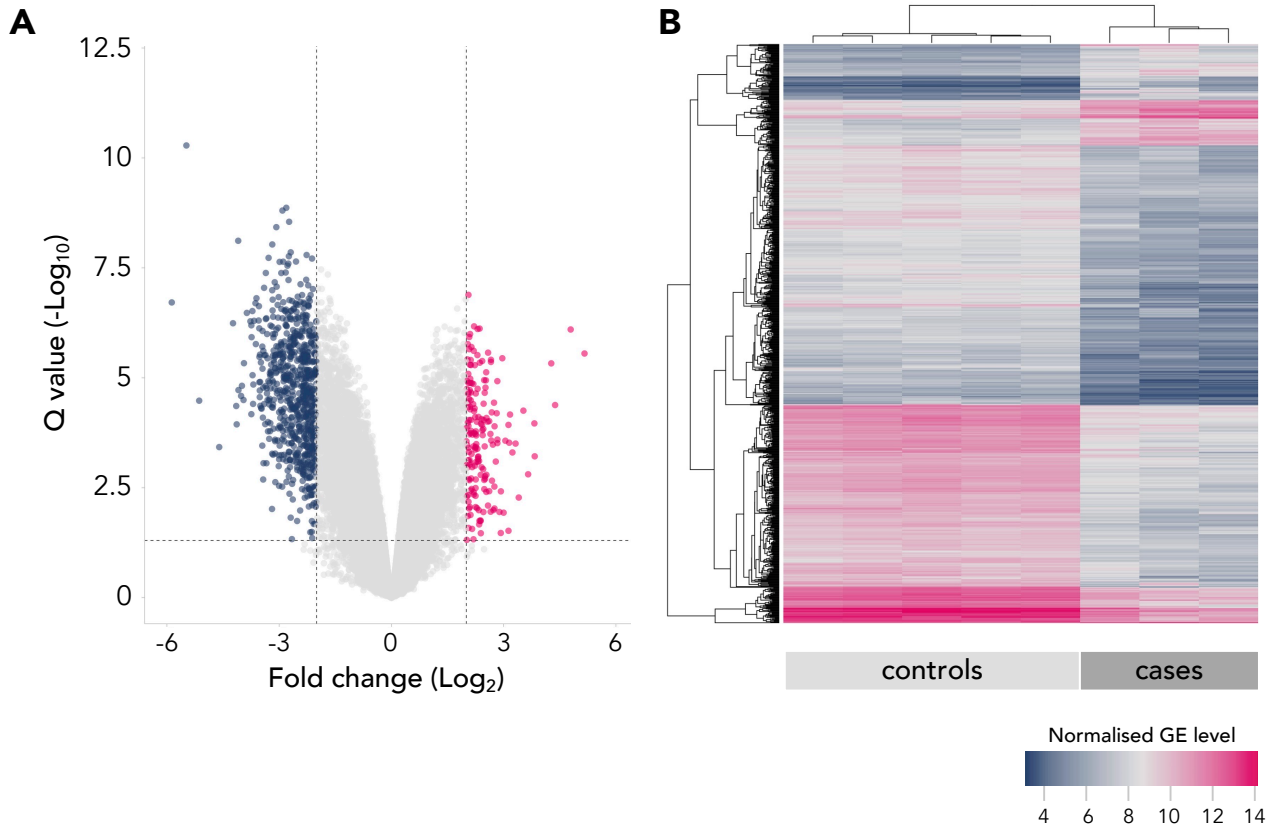

Figure 1. Amniocyte gene expression is different in fetuses with NTDs compared to controls. (A) Volcano plot and **(B)** heat map of genes with increased (n=155) and decreased (n=725) expression in cases compared to controls (clustering method: average linkage; distant measurement method: Euclidean; clustering applies to rows and columns). Differential expression was determined at absolute fold change > 2.0 and q < .05. GE = gene expression.

Nutrient-sensitive miRNAs target differentially regulated genes in cases

We identified 22 unique miRNAs that had a previous link to NTDs and were sensitive to the nutrients of interest identified a priori (Figure 2B). Collectively, these 22 miRNAs targeted 39% (n=344) of DEGs in cases (19% [n=155] for B vitamin-sensitive miRNAs, 30% [n=263] for choline-sensitive miRNAs, 23% [n=203] for vitamin D-sensitive miRNAs, 33% [n=289] for inositol sensitive miRNAs, and 33% [n=293] for iron-sensitive miRNAs).

The largest proportion of DEGs were targeted by *MIR-124-3p* (n=87 DEGs), a known suppressor of tumor proliferation and invasion, and neural inflammation[50-52], with activity sensitive to choline, inositol, iron, and vitamin D. *MIR-17* and *MIR-20a*, both belonging to the miR-17/92 cluster with known roles in cell cycle, proliferation and apoptosis, immune function, and neurodegeneration[53], targeted the next largest proportions of DEGs in amniocytes of cases (n=80 DEGs for *MIR-17*, sensitive to choline, inositol, iron, and vitamin D; n=71 DEGs for *MIR-20a*; sensitive to choline, inositol and iron; Figure 2B). Two of the miRNAs (*MIR142-3p* and *MIR-200a*; both involved in epithelial-to-mesenchymal transition[54,55]), only targeted downregulated genes in cases, while the remaining 20 miRNAs identified targeted both up- and downregulated DEGs in cases (Figure 2B).

Network analysis revealed that seven of these 22 unique miRNAs were known to be sensitive to only one nutrient of interest. This included inositol-sensitive miRNAs: *MIR-23a-3p*, *MIR-142-3p*, *MIR-144*, *MIR-185*; iron-sensitive miRNAs: *MIR-200a*, *MIR-301*; and vitamin B6-sensitive miRNA: *MIR-30e-3p* (Figure 3A-B). Iron, inositol, and choline- sensitive miRNAs targeted the largest proportion of DEGs in cases (Figure 3C). miRNA-targetome networks for DEG in amniocytes from cases compared to controls revealed a high degree of overlap in DEGs targeted by the 22 unique nutrient-sensitive miRNAs (Figure 3D).

Nutrient-dependent transcription factors target co-regulated differentially expressed genes

iRegulon analysis revealed 125 and 106 unique TFs predicted to target up- and down- regulated genes, respectively, in case amniocytes (Supplementary Table S4). Of these, 16 TFs had nutrient cofactors and target down- (n=5) and up-regulated (n=10) DEGs, or both (n=1; Table 2). Nutrient cofactors of these TFs included calcium (for n=1 TF), magnesium (n=1), metal cations (n=1), vitamins A (n=1), B2 (n=1), B3 (n=2), B5 (n=1), D (n=1) and zinc (n=10; Figure 4A). Collectively, these 16 TFs are predicted to regulate the expression of 52% (n=460) of DEGs (Figure 4B).

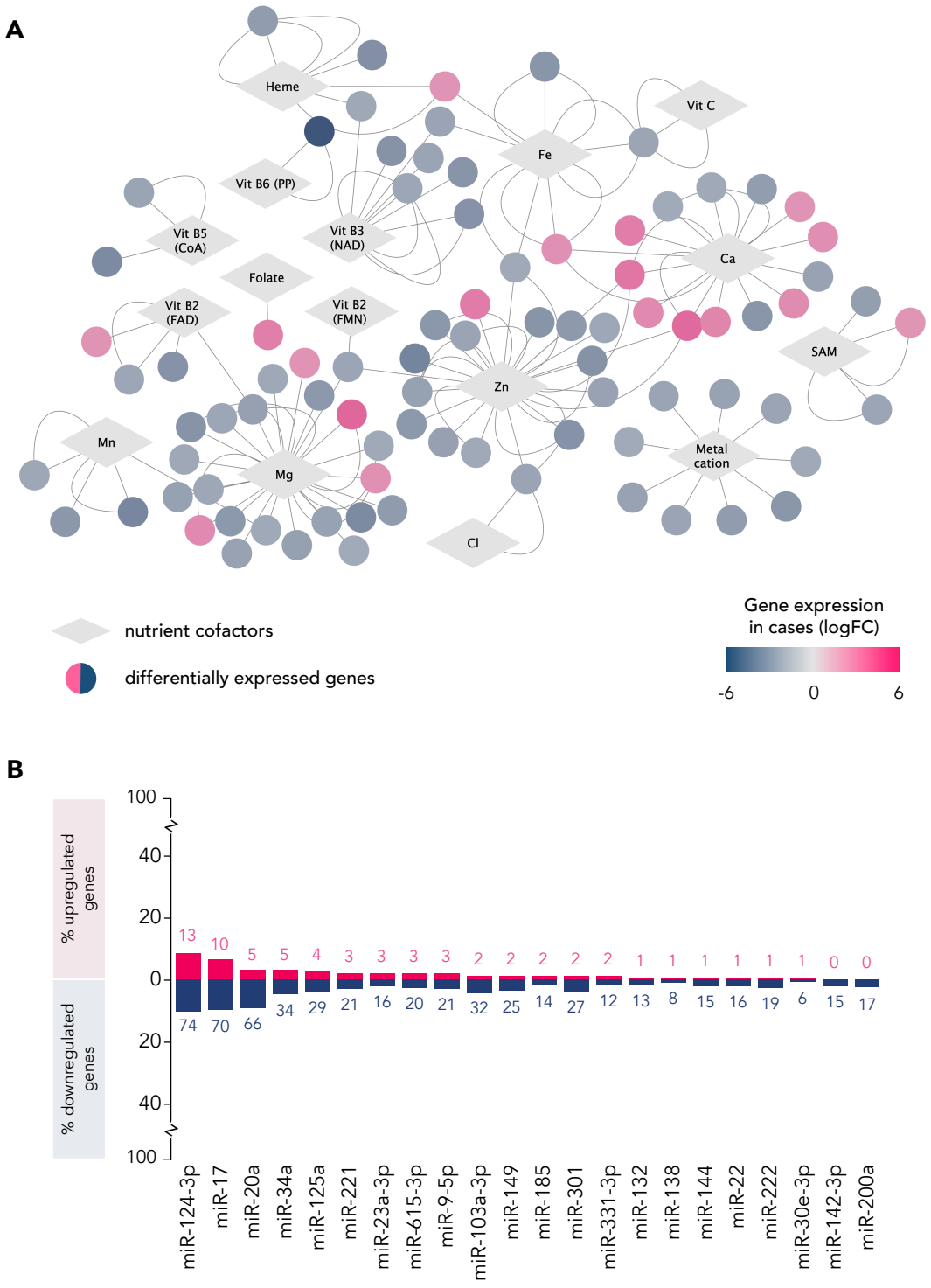

Figure 2. Genes with differential expression in cases compared to controls have (A) nutrient cofactors and (B) are targeted by nutrient-sensitive miRNAs. (A) Genes (circle nodes) that had increased (pink) or decreased (blue) expression in cases compared to controls and have a known nutrient cofactor (grey diamonds) are represented in this network diagram. Data on nutrient-gene (protein) interactions was obtained from Scott-Boyer et al. 2016. (B) The largest proportion of DEG are targeted by miR-124-3p, miR-17 (both sensitive to choline, inositol, iron and vitamin D), and miR-20a (sensitive to choline, inositol and iron; miRWalk2.0). Red = upregulated, blue = downregulated, where number = number of DEG targeted by that miRNA. CoA = Coenzyme A. FAD = Flavin adenine dinucleotide. FMN = Flavin mononucleotide. logFC = log fold change. NAD = Nicotinamide adenine dinucleotide. PP = Pyridoxal phosphate. SAM = S-Adenosyl methionine. Vit = vitamin.

Table 1. Genes with a nutrient cofactor that were differentially expressed in cases (vs. controls).

Nutrient	Gene (n=87)	Entrez ID	Gene title	logFC	q value	Key functions and processes *
Calcium	ALOX5	240	Arachidonate 5-lipoxygenase	2.12	0.03	• Enzyme required for leukotriene biosynthesis, involved in immune and inflammatory processes
	ANXA8	653145	Annexin A8	-2.77	0.02	• Anticoagulant protein
	COLEC12	81035	Collectin Subfamily Member 12	2.66	0.0005	• Involved in the recognition, internalization, and degradation of oxidatively modified low-density lipoprotein (oxLDL) by vascular endothelial cells
	CRNN	49860	Cornulin	-2.49	0.003	• Promotes the G1/S cycle of cell proliferation, involved in regulation proinflammatory cytokine-induced cell proliferation
	ENPP2	5168	Ectonucleotide Pyrophosphatase/ Phosphodiesterase 2	2.83	0.006	• Involved in cell motility (including angiogenesis, neurite outgrowth), stimulates microtubule formation and migration of smooth muscle
	ENTPD3	956	Ectonucleoside Triphosphate Diphosphohydrolase 3	-2.29	0.0002	• Involved in hydrolysis of adenosine triphosphate (ATP) to regulate extracellular levels of ATP
	GAS6	2621	Growth Arrest Specific 6	2.03	0.0005	• Involved in cell growth, survival, adhesion, and migration
	ITGB2	3689	Integrin Subunit Beta 2	2.24	0.008	• Involved in cell adhesion and signaling
	MMP16	4325	Matrix metallopeptidase 16	3.17	0.002	• Involved in extracellular matrix degradation
	NRP1	8829	Neuropilin 1	2.32	0.01	• Involved in cell migration, cardiovascular system function, angiogenesis, and neuronal circuit formation
	NRP2	8828	Neuropilin 2	2.19	0.002	• Involved in cell migration, suspected involvement in cardiovascular development, axon guidance, and tumorigenesis
	PLA2G12A	81579	Phospholipase A2 Group XIIA	-2.04	0.02	• Cleaves arachidonic acid from phospholipids in the membrane
	PRSS3	5646	Serine Protease 3	-2.02	0.004	• Digestive protease that cleaves proteins at the serine residue
	SLC8A1	6546	Solute Carrier Family 8 Member A1	2.42	0.003	• Membrane-bound ion exchange channel in cardiac tissue, required for proper calcium dependent cardiac muscle contraction
Chloride	ACE2	59727	Angiotensin Converting Enzyme 2	-2.24	0.003	• Involved in renin-angiostatin hormone system regulation
Iron	ALOX12	239	Arachidonate 12-Lipoxygenase, 12S Type	-2.79	0.003	• Oxidizes polyunsaturated fats using molecular oxygen to create lipid hydroperoxides
	ALOX5	240	Arachidonate 5-lipoxygenase	2.12	0.03	• Enzyme required for leukotriene biosynthesis, involved in immune and inflammatory processes
	MSMO1	6307	Methylsterol Monooxygenase 1	-2.24	0.003	• Involved in the biosynthesis of cholesterol and can metabolize eldecalcitol (a vitamin D analog)
	PHYH	5264	Phytanoyl-CoA 2-Hydroxylase	-2.21	0.004	• Involved in alpha oxidation of fatty acids and is important for lipid metabolism
	PPP3CB	5532	Protein Phosphatase 3 Catalytic Subunit Beta	-2.04	0.005	• Involved in dephosphorylating proteins involved in the transduction of intracellular Ca^{2+}-mediated signals
	PTGS1	5742	Prostaglandin-Endoperoxide Synthase 1	2.02	0.007	• Involved in prostaglandin synthesis, and is also involved in angiogenesis in endothelial cells
Heme	DUOX2	50506	Dual Oxidase 2	-2.09	0.006	• Involved in thyroid hormone synthesis and antimicrobial defence at the cell mucosa level
	NR1D2	9975	Nuclear Receptor Subfamily 1 Group D Member 2	-3.00	0.03	• A transcriptional repressor that regulates circadian rhythm and other metabolic pathways (lipid metabolism, inflammatory response)
	PTGS1	5742	Prostaglandin-Endoperoxide Synthase 1	2.02	0.007	• Involved in prostaglandin synthesis, and is also involved in angiogenesis in endothelial cells
	TDO2	6999	Tryptophan 2,3-Dioxygenase	-2.52	0.02	• Involved in the kynurenine pathway (first and rate-limiting step in tryptophan metabolism)

Metal cation	GDPD3	79153	Glycerophosphodiester Phosphodiesterase Domain Containing 3	-2.24	0.0004	• Involved in the hydrolysis of lysoglycerophospholipids to become lysophosphatidic acid (LPA)
	MT1E	4493	Metallothionein 1E	-2.31	0.01	• Acts as an antioxidant and is involved in copper homeostasis
	MT1F	4494	Metallothionein 1F	-2.25	0.002	• Acts as an antioxidant and is involved in copper homeostasis, contains cysteine residues to bind to metals
	MT1G	4495	Metallothionein 1G	-2.32	0.005	• Acts as an antioxidant and is involved in metal homeostasis (analog to MTIE), transcriptionally regulated by metals and corticosteroids
	MT1H	4496	Metallothionein 1H	-2.25	0.002	• Acts as an antioxidant, paralog of MT1M, transcriptionally regulated by metal cations and corticosteroids
	MT1HL1	645745	Metallothionein 1H Like 1	-2.05	0.003	• Acts as an antioxidant, transcriptionally regulated by metals and corticosteroids
	MT1M	4499	Metallothionein 1M	-2.74	0.009	• Acts as an antioxidant and involved in mineral absorption
	MT1X	4501	Metallothionein 1X	-2.52	0.002	• Acts as an antioxidant, may be involved in anti-apoptotic pathways, transcriptionally regulated by metals and corticosteroids
	MT2A	4502	Metallothionein 2A	-2.02	0.0024	• Acts as an antioxidant and is involved in metal homeostasis, apoptotic and autophagy pathways, involved in regulating intracellular zinc levels, transcriptionally regulated by metals and corticosteroids
Magnesium	ADCY9	115	Adenylate Cyclase 9	-2.83	0.001	• Involved in cyclic adenosine monophosphate (cAMP) from ATP formation, corticosteroid regulation
	ADSSL1	122622	Adenylosuccinate Synthase 1	3.23	0.009	• Involved in the conversion of inosine phosphate (IMP) to adenosine monophosphate (AMP), first step in the purine nucleotide cycle
	ATP10B	23120	ATPase Phospholipid Transporting 10B (Putative)	-2.31	0.0002	• Involved in the maintenance of lysosomal membrane integrity and cortical neuron function
	ATP8B1	5205	ATPase Phospholipid Transporting 8B1	-2.68	0.0003	• Involved in the transfer of aminophospholipids
	ATP8B4	79895	ATPase Phospholipid Transporting 8B4 (Putative)	2.07	0.001	• Involved in the transfer of aminophospholipids and may be involved in vesicle formation and in the uptake of lipid signaling molecules
	CDS1	1040	CDP-Diacylglycerol Synthase 1	-2.32	0.0002	• Involved in phosphatidylinositol regulation and regulation of cell growth, calcium metabolism, and protein kinase C activity
	CSF1R	1436	Colony Stimulating Factor 1 Receptor	2.30	0.008	• Regulates the production, differentiation, and function of macrophages
	ENO2	2026	Enolase 2	2.14	0.009	• Promotes cell survival of neocortical neurons
	ERN1	2081	Endoplasmic Reticulum to Nucleus Signaling 1	-3.17	0.002	• Involved in protein turnover, specifically in degrading misfolded proteins, regulating protein synthesis, and activating chaperones
	IDI1	3422	Isopentenyl-Diphosphate Delta Isomerase 1	-2.37	0.001	• Involved in cholesterol synthesis
	KIT	3815	KIT Proto-Oncogene, Receptor Tyrosine Kinase	-2.39	0.01	• Involved in the phosphorylation of intracellular proteins, hematopoiesis, stem cell maintenance, gametogenesis, melanogenesis
	LATS2	26524	Large Tumor Suppressor Kinase 2	-2.12	0.01	• Involved in localizing on the centromeres during interphase and metaphase and is required for spindle formation, may be involved in the cytoskeleton damage response
	MAP3K9	4293	Mitogen-Activated Protein Kinase 9	-2.69	0.001	• Involved in the MAPK signal transduction pathway, involved in mitochondrial death and resulting apoptosis by stimulating the release of Cytochrome c
	MAPK8	5599	Mitogen-Activated Protein Kinase 8	-2.07	0.002	• Involved in the MAPK signal transduction pathway, apoptosis via stimulation of Cytochrome c release, and TNF-α and UV-related apoptosis
	NT5C2	22978	5'-Nucleotidase, Cytosolic II	-2.06	0.001	• Involved in purine metabolism.
	NT5C3A	51251	5'-Nucleotidase, Cytosolic IIIA	-2.55	0.005	• Involved in nucleoside metabolism
	PPM1A	5494	Protein Phosphatase, Mg^{2+}/Mn^{2+} Dependent 1A	-2.06	0.005	• Involved in the negative regulation of the cell stress response pathways and cell cycle regulation
	PPM1K	152926	Protein Phosphatase, Mg^{2+}/Mn^{2+} Dependent 1K	-2.15	0.0008	• Involved in the regulation of the mitochondrial permeability transition pore

	PRKCQ	5588	Protein Kinase C Theta	-2.35	0.002	• Involved in NF-kappaB and AP-1 signaling, T-cell activation
	RFK	55312	Riboflavin Kinase	-2.19	0.018	• Catalyzes the conversion of riboflavin (vitamin B2) to flavin mononucleotide (FMN), role in vitamin B2 utilisation
	SIK3	23397	SIK Family Kinase 3	-2.10	0.004	• Positive regulator of mTOR signaling
	SRXN1	140809	Sulfiredoxin 1	-2.00	0.005	• Involved in anti-inflammatory actions through the reduction of cysteine-sulfinic acid produced by oxidants in certain peroxisomes
Manganese	B3GALT2	8707	Beta-1,3-Galactosyltransferase 2	-2.71	0.015	• Involved in the synthesis of glycolipids and glycoproteins (specifically, carbohydrate moieties)
	B4GALT4	8702	Beta-1,4-Galactosyltransferase 4	-2.12	0.007	• Involved in the synthesis of glycolipids and glycoproteins (specifically, carbohydrate moieties)
	GALNT3	2591	Polypeptide N-Acetylgalactosaminyltransferase 3	-3.31	0.0005	• Involved in phosphate homeostasis and O-linked oligosaccharide biosynthesis
	PPM1A	5494	Protein Phosphatase, Mg^{2+}/Mn^{2+} Dependent 1A	-2.06	0.005	• Involved in cell cycle regulation
S-Adenosyl methionine	HNMT	3176	Histamine N-Methyltransferase	2.01	0.004	• Involved in histamine inactivation
	PRMT5	10419	Protein Arginine Methyltransferase 5	-2.18	0.001	• Involved in the addition of methyl groups from S-adenosylmethionine (SAM) to arginine residues in proteins
	VCPKMT	79609	Valosin Containing Protein Lysine Methyltransferase	-2.47	0.01	• Involved in general protein metabolism and methylation
Vitamin B2 (FAD)	CYBB	1536	Cytochrome B-245 Beta Chain	2.04	0.002	• Involved in the pH maintenance of phagocytic cells
	IDI1	3422	Isopentenyl-Diphosphate Delta Isomerase 1	-2.37	0.001	• Involved in cholesterol synthesis
	QSOX1	5768	Quiescin Sulfhydryl Oxidase 1	-2.16	0.002	• Involved in growth regulation of fibroblasts as they exit the proliferative cycle and enter quiescence
	STEAP4	79689	STEAP4 Metalloreductase	-2.90	0.0005	• Involved in the regulation of the inflammatory cytokine response, metabolic homeostasis, and cell proliferation
Vitamin B2 (FMN)	RFK	55312	Riboflavin Kinase	-2.19	0.018	• Catalyzes the conversion of riboflavin (vitamin B2) to form flavin mononucleotide (FMN), role in vitamin B2 utilisation
Vitamin B3 (NAD)	ALDH1A3	220	Aldehyde Dehydrogenase 1 Family Member A3	-2.88	0.0007	• Involved in synthesis of retinoic acid
	DUOX2	50506	Dual Oxidase 2	-2.09	0.006	• Involved in thyroid hormone synthesis and antimicrobial defence at the cell mucosa level
	GMDS	2762	GDP-Mannose 4,6-Dehydratase	-2.15	0.01	• Involved in fructose and mannose metabolism (carbohydrate metabolism)
	HPGD	3248	15-Hydroxyprostaglandin Dehydrogenase	-2.90	0.002	• Involved in prostaglandin inactivation and regulation of the inflammatory response
	MSMO1	6307	Methylsterol Monooxygenase 1	-2.24	0.003	• Involved in the biosynthesis of cholesterol and can metabolize eldecalcitol (a vitamin D analog)
	RDH10	157506	Retinol Dehydrogenase 10	-2.22	0.001	• Involved in the synthesis of all-trans-retinal
	SDR16C5	195814	Short Chain Dehydrogenase/ Reductase Family 16C Member 5	-2.81	0.001	• Responsible for oxidation of retinol to retinaldehyde
Vitamin B5 (CoA)	ABHD5	51099	Abhydrolase Domain Containing 5, Lysophosphatidic Acid Acyltransferase	-3.17	0.0002	• Places an acyl group on a lysophosphatidic acid
	AGPAT9	84803	Glycerol-3-Phosphate Acyltransferase 3	-2.25	0.03	• Involved in the formation of lysophosphatidic acid, and lipodystrophy and lipid metabolism

Vitamin B6 (PP)	ETNPPL	64850	Ethanolamine-Phosphate Phospho-Lyase	-5.14	0.002	• Involved in lipid metabolism
Vitamin B9 (THF)	FOLR2	2350	Folate Receptor Beta	2.57	0.003	• Involved in the transfer of 5-methyltetrahydrofolate to the interior of cells
Vitamin C (Ascorbic Acid)	PHYH	5264	Phytanoyl-CoA 2-Hydroxylase	-2.21	0.004	• Involved in the alpha oxidation of fatty acids with a 3-methyl branch, lipid metabolism
Zinc	ACE2	59727	Angiotensin Converting Enzyme 2	-2.24	0.0031	• Involved in renin-angiostatin hormone system regulation
	ADAM12	8038	ADAM Metallopeptidase Domain 12	-2.05	0.02	• Involved in muscle development, neurogenesis, and fertilization, and macrophage-derived giant and osteoclast cell differentiation
	APOBEC3A	200315	Apolipoprotein B MRNA Editing Enzyme Catalytic Subunit 3A	-2.92	0.006	• Involved in the immune response, prevents the integration of foreign DNA
	APOBEC3B	9582	Apolipoprotein B MRNA Editing Enzyme Catalytic Subunit 3B	-2.73	0.009	• Involved in the immune response, prevents the integration of foreign DNA
	BIRC3	330	Baculoviral IAP Repeat Containing 3	-3.40	0.002	• Involved in apoptosis inhibition, inflammatory signalling pathways, cell proliferation, and cell invasion during metastasis
	BRAF	673	B-Raf Proto-Oncogene, Serine/Threonine Kinase	-2.68	0.001	• Involved in MAPK signalling pathway, influences cell differentiation, division, and secretion
	C9orf3	84909	Aminopeptidase O (Putative)	-2.22	0.0001	• Involved in the angiotensin pathway
	CA13	377677	Carbonic Anhydrase 13	-2.35	0.001	• Involved in pH maintenance and carbon metabolism.
	CRYAB	1410	Crystallin Alpha B	-2.30	0.0005	• Facilitates protein renaturation
	ENPP2	5168	Ectonucleotide Pyrophosphatase/ Phosphodiesterase 2	2.83	0.006	• Angiogenic factor involved in cell motility, including in angiogenesis and neurite outgrowth
	GNE	10020	Glucosamine (UDP-N-Acetyl)-2-Epimerase/N-Acetylmannosamine Kinase	-2.37	0.0002	• Involved in cell adhesion, signal transduction, early development, and healthy sialylation in hematopoietic cells
	MMP16	4325	Matrix Metallopeptidase 16	3.17	0.002	• Involved in extracellular matrix degradation
	NAIP	4671	NLR Family Apoptosis Inhibitory Protein	2.68	0.002	• Involved apoptosis inhibition, cytokine production and macrophage pyroptosis
	PDLIM1	9124	PDZ And LIM Domain 1	-2.11	0.001	• Involved in maintenance of the cytoskeleton and of the cell polarity in fibroblasts
	PPP3CB	5532	Protein Phosphatase 3 Catalytic Subunit Beta	-2.04	0.005	• Involved in dephosphorylating proteins involved in the transduction of intracellular Ca^{2+}-mediated signals
	RFK	55312	Riboflavin Kinase	-2.19	0.018	• Catalyzes the conversion of riboflavin (vitamin B2) to flavin mononucleotide (FMN), role in vitamin B2 utilisation
	S100A12	6283	S100 Calcium Binding Protein A12	-2.80	0.024	• Proinflammatory mediator involved in the regulation of the inflammatory processes that recruits leukocytes and upregulates cytokine production
	S100A7	6278	S100 Calcium Binding Protein A7	-2.63	0.02	• Involved in cell cycle progression and cell adhesion, and the immune response (proinflammatory role in the IL-7 signalling pathway)
	YME1L1	10730	YME1 Like 1 ATPase	-2.89	0.003	• Involved in mitochondrial morphology and prevents the mitochondria from oxidatively damaging proteins, and promotes anti-apoptotic signals

[*] All key information on gene (protein) functions and processes were obtained from genecards.org. logFC = log fold change. FAD = Flavin adenine dinucleotide. NAD = Nicotinamide adenine dinucleotide. CoA = Coenzyme A. PP = Pyridoxal phosphate. THF = Tetrahydrofolate.

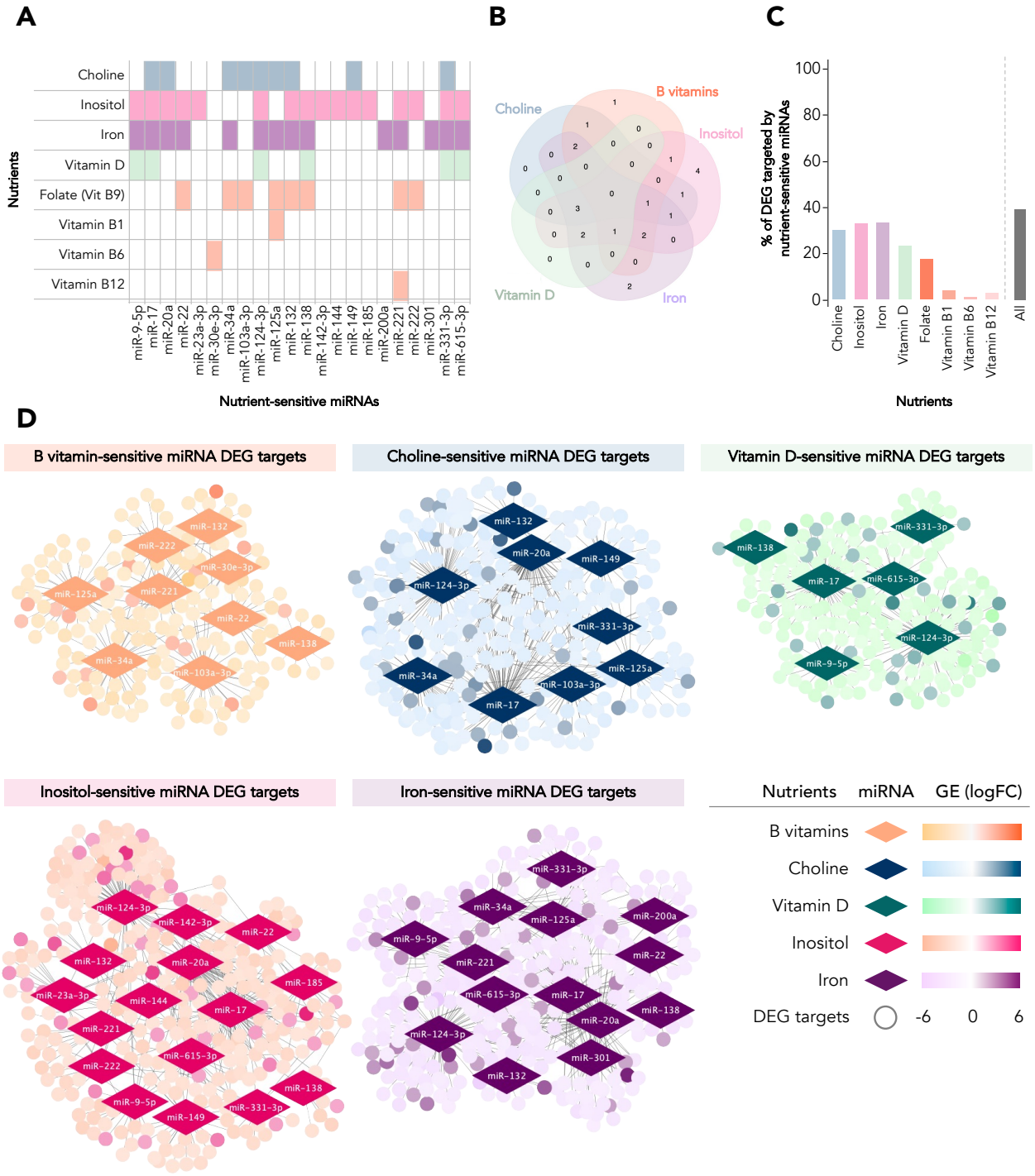

Figure 3. Nutrient-sensitive miRNA-DEG network analysis. (A-B) 22 unique miRNAs known to be sensitive to the activity of nutrients of interest (identified a priori, miRWalk2.0) with a previous link to NTDs. Of the 22 miRNAs, seven were sensitive to only one nutrient of interest (inositol-sensitive: miR-23a-3p, miR-142-3p, miR-144, miR-185; iron-sensitive: miR-200a, miR-301; vitamin B6-sensitive: miR-30e-3p). (C) Iron, inositol and choline- sensitive miRNAs targeted the largest proportion of DEGs in cases, respectively. Collectively, the 22 unique miRNAs targeted 39% of DEG in cases (n=344; B vitamin-sensitive miRNAs: 19% [n=155], choline-sensitive miRNAs: 30% [n=263], vitamin D-sensitive miRNAs: 23% [n=203], inositol-sensitive miRNAs: 33% [n=289], iron-sensitive miRNAs: 33% [n=293]). (D) Nutrient-sensitive miRNA-targetome networks for DEG in cases compared to controls. Diamonds = nutrient-sensitive miRNAs. Circles (nodes) = DEG targets, where lighter shades = downregulated in cases, darker shades = upregulated in cases. DEG = differentially expressed genes. GE = gene expression. FC = fold change. Vit = vitamin.

E1A Binding Protein P300 (*EP300*), a transcription factor with known roles in cell growth, differentiation, and NTDs, was the top nutrient-dependent regulator of DEG transcription identified and the only TF identified that putatively targeted both up- and down-regulated genes. *EP300* had 311 expected targets (upregulated: n=9, downregulated: n=302; 35.3% of total DEGs; Table 2), and zinc and vitamin B5 cofactors. The zinc-requiring P53 family proteins (*TP53*, *TP63*, and *TP73*), involved in cell proliferation and apoptosis, were the next-largest putative regulators of DEGs, potentially targeting 252 downregulated genes in cases (28.6% of total DEGs).

Of TFs identified as potential regulators of upregulated genes, the metal-cation reliant TROVE domain family, member 2 (*TROVE2*) gene, a known regulator of immune and inflammatory processes, was associated with the largest proportion (n=35 [4.0%] upregulated genes in cases). Quiescin Sulfhydryl Oxidase 1 (*QSOX1;* vitamin B2-dependent) which targeted the next-largest group of DEGs (n=25 upregulated genes), was also downregulated in cases (FC=-2.16, q=0.002; Table 2). One of the P53 family proteins (*TP63*), as well as an additional zinc-dependent TF identified (Churchill Domain Containing 1 [*CHURC1*]), have not been previously linked to NTDs.

Nutrient-sensitive TF-gene networks were visualised for enriched TFs predicted to target downregulated and upregulated genes (Figure 4C-D). All TFs predicted to target downregulated genes had zinc as a cofactor, and one (*EP300*) was also vitamin B5-dependent (Figure 4C). TFs targeting upregulated genes were dependent on calcium, magnesium, metal cations, vitamins A, B3, B5, D, and zinc (Figure 4D). In the TF-gene network figures, it is visually evident that the regulons of downregulated genes targeted by nutrient-sensitive TFs were much larger than the upregulated regulons, but were fewer in number (Figure 4C-D).

Discussion

Using a novel nutrient-focused gene expression analysis pipeline, we identified multiple nutrient-sensitive genes and gene regulatory networks that were dysregulated in amniocytes from fetuses with NTDs. We found that over half of the dysregulated genes were predicted to be co-regulated by TFs and miRNAs that are sensitive to, or dependent on, nutrients with known roles in NTD etiology, namely zinc, B vitamins, and B vitamin-like nutrients (choline and inositol). These findings suggest that nutrient-sensitive genes and gene regulatory networks may explain an important portion of genetic signatures associated with NTDs. Profiling nutrient-transcriptome interaction networks in the context of NTDs may help to illustrate, in part, the gene-environment interactions that contribute to NTD pathogenesis, improving understanding of causal pathways.

Several zinc-dependent genes and gene regulatory networks were dysregulated in case amniocytes compared to controls. Zinc is a key regulator of gene expression, and is essential for the rapid DNA synthesis, cell growth, and proliferation that accompany embryonic development, including neural tube closure[56,57]. Low maternal zinc status may be associated with increased NTD risk[58], and zinc imbalance in animal studies leads to NTDs, possibly through modulation of *p53* activity[13]. Here, three zinc-dependent P53 family proteins were identified as transcription factors predicted to target over a quarter of DEGs in case amniocytes, and eight Metallothionein genes, which are essential for maintaining zinc homeostasis and controlling zinc-dependent cellular signaling[59], were downregulated in cases. Altogether, our findings support a role for zinc-associated gene dysregulation in NTD genetic phenotypes, in alignment with previous animal studies, and illustrate novel zinc-dependent gene networks in an NTD-affected a human cohort.

One zinc- and vitamin B5- dependent TF, *EP300*, was predicted to target over one third of DEGs in case amniocytes. *EP300* has been studied in mouse models of NTDs and is predicted to activate pathways that promote fusion of the neural tube[60,61]. In humans, rare deleterious variants in the *EP300* gene have recently been associated with NTDs[62], and here, *EP300* was the dominant, nutrient-sensitive focal point for DEG regulation in cases. Two additional nutrient-dependent TFs identified here further support a role for both zinc and B vitamin-interacting gene dysregulation in cases: Sirtuin 6 (*SIRT6*; vitamin B3- and zinc- dependent) and *CHURC1* (zinc-dependent). *SIRT6* suppression by high glucose levels is a proposed to mechanism linking maternal diabetes to increased risk of fetal NTDs[63], however, *CHURC1* has not previously been associated with NTDs. In zebrafish and xenopus, *churc1* is known to play a role in embryogenesis and neuroectodermal development[64,65]. In humans, *CHURC1* has been investigated in relation to autism, where a microdeletion in a gene region containing the genes for both *CHURC1* and methylenetetrahydrofolate dehydrogenase 1 (*MTHFD1*), a folate metabolism enzyme with known links to NTD etiology[15,27], associated with increased autism risk[66]. While research on zinc-B vitamins interactions is limited, there is some evidence that zinc may have a role in niacin (vitamin B3) metabolism through interaction with pyridoxine (vitamin B6)[67,68]. Overall, the dominance of dysregulated zinc- and B vitamin- dependent gene networks in our results suggest that both zinc- and B vitamin-gene interactions in NTDs should be a focus of further study.

To date, folates (vitamin B9), and more recently, folate one-carbon metabolism (OCM) cofactors (namely vitamins B12 and B6), have been the central nutrient targets for NTD prevention. In agreement with the established role of folates and OCM cofactors in NTDs, we identified nine miRNAs with known associations to NTDs that were also B vitamin-sensitive, eight being sensitive to folate, and were predicted to regulate the expression of nearly one fifth of DEGs in case amniocytes. Further, the three miRNAs that targeted the largest proportion of DEGs in cases (*MIR-124-3p*, *MIR-17*, and *MIR-20a*) were each sensitive to both choline and inositol, water soluble B vitamin-like molecules with known roles in NTD etiology[9,11], and four TFs with B vitamin cofactors (*EP300* and *SIRT6*, as well as C-terminal-binding protein 2 [*CTBP2*] and *QSOX1*) predicted to target DEGs in cases were identified. Both *QSOX1* and *CTBP2* have inferred associations with NTDs via their interactions with valproic acid and folic acid[69], and the role of the *CTBP2* gene in the Wnt and Notch signaling pathways[70]. However, no studies have provided direct evidence linking *QSOX1* or *CTBP2* to NTD pathogenesis or phenotype in humans, and here, *QSOX1* expression was also downregulated in case amniocytes compared to controls. Taken together, our findings advance beyond a folic acid (vitamin B9)-centric view of the role of B vitamins in NTDs and highlight B vitamin-sensitive or dependent gene candidates for further study in NTDs.

Additionally, we found that six miRNAs with known associations to NTDs were vitamin D-sensitive, and collectively, were predicted to target nearly one quarter of DEGs in case amniocytes. Vitamin D receptor (VDR) was also identified as a transcription factor targeting DEGs, and Methylsterol monooxygenase 1 (*MSMO1*), associated with vitamin D analog metabolism, was downregulated in cases. Vitamin D plays important anti-inflammatory and

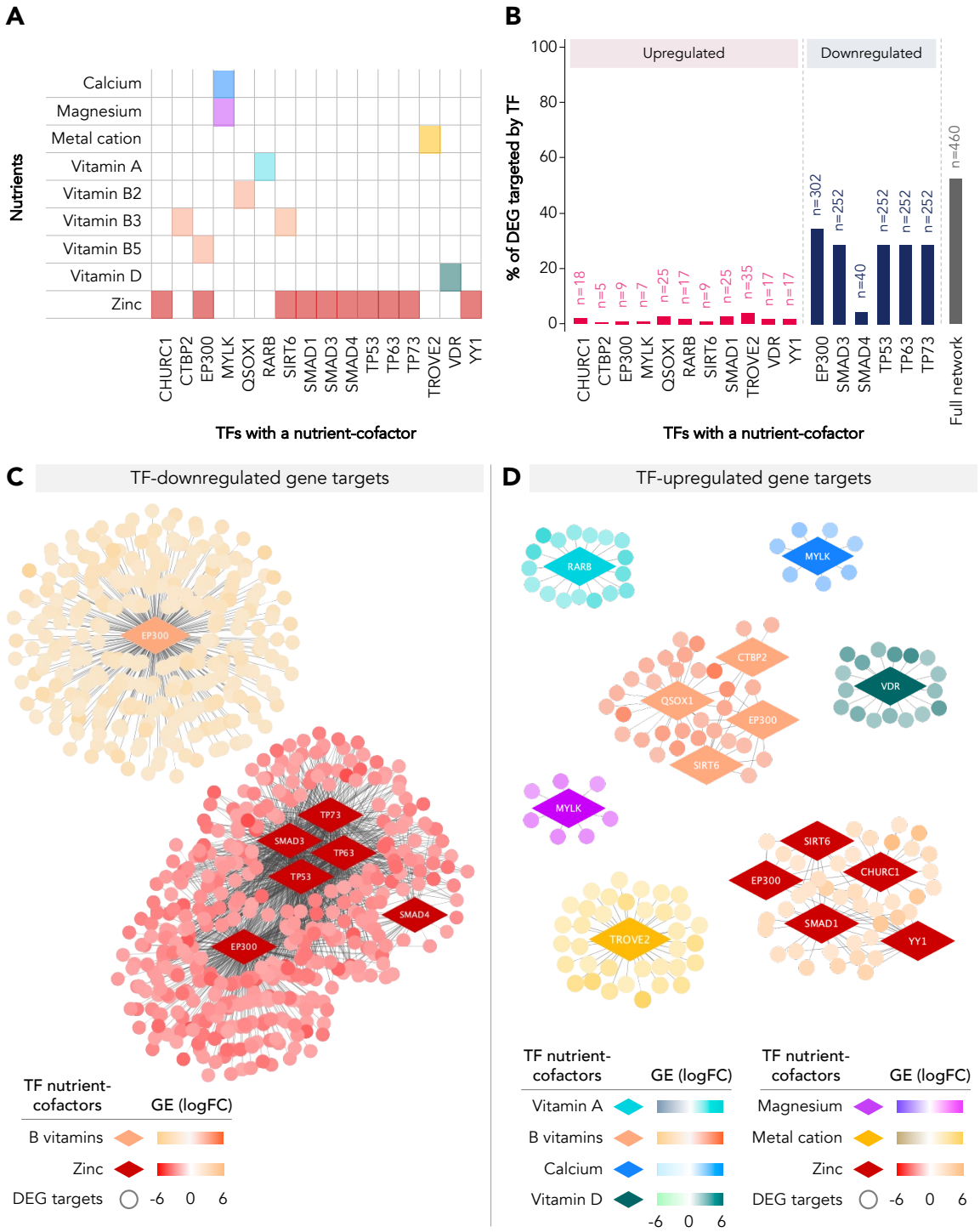

Figure 4. Nutrient-sensitive transcription factor regulatory networks for DEGs in cases compared to controls. (A) 16 transcription factors with nutrient-cofactors that are predicted to regulate DNA motifs or tracks of DEG were identified. Nutrient cofactors of these TFs included calcium (for n=1 TF), magnesium (n=1), metal cations (n=1), vitamins A (n=1), B2 (n=1), B3 (n=2), B5 (n=1), D (n=1) and zinc (n=10). **(B)** Collectively, these 16 TFs are predicted to regulate the expression of 52% (n=460) of DEGs. Red = upregulated in cases, blue = downregulated in cases, grey = total proportion of DEG targeted by a nutrient-sensitive TF. Nutrient-sensitive TF-gene networks were visualised for TFs predicted to target **(C)** downregulated genes (n=5 unique TFs, left-bottom panel) and **(D)** upregulated genes **(**n=10 unique TFs, right-bottom panel) separately. One TF (E1A Binding Protein P300; EP300) is predicated to target both up- and down- regulated genes in cases. Diamonds = nutrient-sensitive TFs. Circles (nodes) = DEG targets, where colours on the left of the colour palette = downregulated in cases, right of the colour palette = upregulated in cases. DEG = differentially expressed genes. GE = gene expression. FC = fold change. TF = transcription factor.

Table 2. Transcription factors with a nutrient cofactor that are predicted to target genes that are differentially expressed in cases compared to controls.

Transcription factor	NES *	# Of targets (DEG)	# Of motifs/ tracks	Cluster code **	Nutrient cofactor(s)	Linked to NTDs?
Putatively targeting upregulated genes in cases:						
CHURC1	3.27	18	1/0	M5	Zinc	No
CTBP2	3.13	5	0/1	T6	Vit B3 (NAD)	Yes
EP300	3.33	9	1/0	M5	Vit B5 (CoA), Zinc	Yes
MYLK	3.28	7	1/0	M14	Mg, Ca	Yes
QSOX1	3.69	25	2/0	M5	Vit B2 (FAD)	Yes
RARB	3.69	17	1/0	M5	Vit A	Yes
SIRT6	3.33	9	1/0	M5	Vit B3 (NAD), Zinc	Yes †
SMAD1	3.69	25	2/0	M5	Zinc	Yes
TROVE2	3.42	35	1/0	M12	Metal cation	Yes
VDR	3.69	17	1/0	M5	Vit D	Yes
YY1	3.69	17	1/0	M5	Zinc	Yes
Putatively targeting downregulated genes in cases:						
EP300	3.64	302	0/3	T2	Vit B5 (CoA), Zinc	Yes
SMAD3	3.49	252	5/0	M1	Zinc	Yes
SMAD4	3.19	40	1/0	M8	Zinc	Yes
TP53	3.49	252	5/0	M1	Zinc	Yes
TP63	3.49	252	5/0	M1	Zinc	No
TP73	3.49	252	4/0	M1	Zinc	Yes

* Maximal enrichment score for the motif(s) or track(s) predicted to be targeted by the TF. ** Cluster codes indicate the motif (M) or track (T) clusters a TF is associated with. † Previous link to NTDs in animal models only. NES = Normalized enrichment score. DEG = differentially expressed genes.

immunomodulatory roles[71], and low maternal vitamin D levels have been associated with fetal NTDs[10]. In mice, vitamin D3 supplementation prevents lipopolysaccharide-induced NTDs via increased folate transport from the maternal circulation to the embryo[72]. Thus, our findings support vitamin D as a nutrient of interest for further research on nutrient-gene interactions in NTD pathogenesis.

The small size of this cohort is a key limitation to our study that may limit the validity of our findings, however, the rarity of NTDs and the challenges that accompany amniotic fluid collection make these data a valuable resource for examining the genomic signatures associated with isolated NTDs. Data on cohort characteristics were also limited, and it is plausible that there are unexplored variables that could help explain the differences observed here, including maternal clinical or demographic characteristics, as well as dietary intakes; namely of nutrients that have known associations with NTDs. Fortification of wheat products with folic acid had also not been introduced at the time of data collection, and while the original authors did report that all study participants were taking a multivitamin supplement throughout pregnancy[21], it remains difficult to speculate about the nutrient status, including folate, of the study's participants. Lastly, our scope for miRNA-targetome analysis was narrowed to focus on only miRNAs with a previous link to NTDs. This may have prevented the identification of novel miRNAs associated with NTDs, and future research with higher computational resources should consider an expanded focus.

Building on the work of authors who previously analysed these data to understand gene- and gene pathway dysregulation in isolated NTDs[20,21], our nutrient-focused network analysis approach integrates knowledge on nutrient-gene interactions with gene expression network analysis to examine the functional networks underlying NTD phenotypes. Network medicine is an essential tool for identifying the functional networks that drive health and disease states[22], and we are the first to apply this approach in the context of nutrient-gene interactions in NTDs. The nutrient-focused gene expression analysis pipeline applied here may also be of use for research in other areas at the intersection of nutrition and health and disease, beyond NTDs. Integrated nutrient-transcriptome network analysis may provide insights into nutrient-sensitive mechanisms that underly disease processes, which may aid in the identification of novel nutrition targets for the prevention of disease and promotion of optimal health and developmental outcomes.

Acknowledgments

We thank the authors who originally sequenced these data for making them publicly available in GEO.

Data availability statement

The microarray data used in this analysis were obtained from GEO (GSE4182). The code used for microarray pre-processing has been deposited in a GitHub repository (https://github.com/marina-white/ntd_af_nutrient-gene_study.git).

Funding

This research is supported by the Canadian Institutes of Health Research (CIHR PJT-175161). MW is supported by an Ontario Graduate Scholarship and JAP is supported by a Dean's Summer Research Internship, Faculty of Science, Carleton University.

References

1. Christianson, A., Howson, C.P. & Modell, B. Global report on birth defects: The hidden toll of dying and disabled children. (ed. Foundation, M.o.D.B.D.) (2006).
2. Blencowe, H., Kancherla, V., Moorthie, S., Darlison, M.W. & Modell, B. Estimates of global and regional prevalence of neural tube defects for 2015: a systematic analysis. *Ann N Y Acad Sci* **1414**, 31-46 (2018).

3. Copp, A.J., Stanier, P. & Greene, N.D.E. Genetic Basis of Neural Tube Defects. in *Textbook of Pediatric Neurosurgery* 1-28 (Srpinger, Cham., 2017).
4. Greene, N.D. & Copp, A.J. Neural tube defects. *Annu Rev Neurosci* **37**, 221-242 (2014).
5. White, M., Grynspan, D., Van Mieghem, T. & Connor, K.L. Isolated fetal neural tube defects associate with increased risk of placental pathology: Evidence from the Collaborative Perinatal Project. *Placenta* **114**, 56-67 (2021).
6. Furness, D.*, et al.* Folate, Vitamin B12, Vitamin B6 and homocysteine: impact on pregnancy outcome. *Maternal & Child Nutrition* **9**, 155-166 (2013).
7. Siega-Riz, A.M., Savitz, D.A., Zeisel, S.H., Thorp, J.M. & Herring, A. Second trimester folate status and preterm birth. *American Journal of Obstetrics and Gynecology* **191**, 1851-1857 (2004).
8. Finnell, R.H.*, et al.* Gene Environment Interactions in the Etiology of Neural Tube Defects. *Front Genet* **12**, 659612 (2021).
9. D'Souza, S.W., Copp, A.J., Greene, N.D.E. & Glazier, J.D. Maternal Inositol Status and Neural Tube Defects: A Role for the Human Yolk Sac in Embryonic Inositol Delivery? *Adv Nutr* **12**, 212-222 (2021).
10. Nasri, K.*, et al.* Maternal 25-hydroxyvitamin D level and the occurrence of neural tube defects in Tunisia. *Int J Gynaecol Obstet* **134**, 131-134 (2016).
11. Shaw, G.M., Carmichael, S.L., Yang, W., Selvin, S. & Schaffer, D.M. Periconceptional dietary intake of choline and betaine and neural tube defects in offspring. *Am J Epidemiol* **160**, 102-109 (2004).
12. Tian, T.*, et al.* Single and mixed effects of metallic elements in maternal serum during pregnancy on risk for fetal neural tube defects: A Bayesian kernel regression approach. *Environ Pollut* **285**, 117203 (2021).
13. Kakebeen, A.D. & Niswander, L. Micronutrient imbalance and common phenotypes in neural tube defects. *Genesis* **59**, e23455 (2021).
14. De Wals, P.*, et al.* Reduction in Neural-Tube Defects after Folic Acid Fortification in Canada. *New England Journal of Medicine* **357**, 135-142 (2007).
15. Bassuk, A.G. & Kibar, Z. Genetic basis of neural tube defects. *Semin Pediatr Neurol* **16**, 101-110 (2009).
16. Krupp, D.R.*, et al.* Transcriptome profiling of genes involved in neural tube closure during human embryonic development using long serial analysis of gene expression (long-SAGE). *Birth Defects Res A Clin Mol Teratol* **94**, 683-692 (2012).
17. Price, E.M.*, et al.* Profiling placental and fetal DNA methylation in human neural tube defects. *Epigenetics & chromatin* **9**, 6-6 (2016).
18. Shang, Z.*, et al.* Single-cell RNA-seq reveals dynamic transcriptome profiling in human early neural differentiation. *Gigascience* **7**(2018).
19. Barrett, T.*, et al.* NCBI GEO: archive for functional genomics data sets--update. *Nucleic Acids Res* **41**, D991-995 (2013).
20. Li, Z., Feng, J. & Yuan, Z. Key Modules and Hub Genes Identified by Coexpression Network Analysis for Revealing Novel Biomarkers for Spina Bifida. *Front Genet* **11**, 583316 (2020).
21. Nagy, G.R.*, et al.* Use of routinely collected amniotic fluid for whole-genome expression analysis of polygenic disorders. *Clin Chem* **52**, 2013-2020 (2006).
22. Sonawane, A.R., Weiss, S.T., Glass, K. & Sharma, A. Network Medicine in the Age of Biomedical Big Data. *Front Genet* **10**, 294 (2019).
23. Maguire, C.T.*, et al.* Genome-wide analysis reveals the unique stem cell identity of human amniocytes. *PLoS One* **8**, e53372 (2013).
24. Kang, J.H.*, et al.* Comparative Transcriptome Analysis of Cell-Free Fetal RNA from Amniotic Fluid and RNA from Amniocytes in Uncomplicated Pregnancies. *PLoS One* **10**, e0132955 (2015).
25. Godfrey, K.M. & Barker, D.J. Maternal nutrition in relation to fetal and placental growth. *European journal of obstetrics, gynecology, and reproductive biology* **61**, 15-22 (1995).
26. Paoloni-Giacobino, A., Grimble, R. & Pichard, C. Genetics and nutrition. *Clin Nutr* **22**, 429-435 (2003).
27. Zhang, T.*, et al.* Genetic variants in the folate pathway and the risk of neural tube defects: a meta-analysis of the published literature. *PLoS One* **8**, e59570 (2013).
28. Bentley, T.G., Willett, W.C., Weinstein, M.C. & Kuntz, K.M. Population-level changes in folate intake by age, gender, and race/ethnicity after folic acid fortification. *Am J Public Health* **96**, 2040-2047 (2006).
29. Ray, J.G.*, et al.* Vitamin B12 and the risk of neural tube defects in a folic-acid-fortified population. *Epidemiology* **18**, 362-366 (2007).
30. Belkacemi, L., Nelson, D.M., Desai, M. & Ross, M.G. Maternal undernutrition influences placental-fetal development. *Biol Reprod* **83**, 325-331 (2010).
31. Bloomfield, F.H. How is maternal nutrition related to preterm birth? *Annu Rev Nutr* **31**, 235-261 (2011).
32. Gautier, L., Cope, L., Bolstad, B.M. & Irizarry, R.A. affy--analysis of Affymetrix GeneChip data at the probe level. *Bioinformatics* **20**, 307-315 (2004).
33. Irizarry, R.A.*, et al.* Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics* **4**, 249-264 (2003).
34. McInnes, L., Healy, J., Saul, N. & Lukas, G. UMAP: Uniform Manifold Approximation and Projection. *Journal of Open Source Software* **3**, 861 (2018).
35. Stalteri, M.A. & Harrison, A.P. Interpretation of multiple probe sets mapping to the same gene in Affymetrix GeneChips. *BMC Bioinformatics* **8**, 13 (2007).
36. Maglott, D., Ostell, J., Pruitt, K.D. & Tatusova, T. Entrez Gene: gene-centered information at NCBI. *Nucleic Acids Res* **39**, D52-57 (2011).
37. Ritchie, M.E.*, et al.* limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* **43**, e47 (2015).
38. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B* **57**, 289-300 (1995).
39. Mi, H., Muruganujan, A., Casagrande, J.T. & Thomas, P.D. Large-scale gene function analysis with the PANTHER classification system. *Nat Protoc* **8**, 1551-1566 (2013).
40. Scott-Boyer, M.P.*, et al.* A network analysis of cofactor-protein interactions for analyzing associations between human nutrition and diseases. *Sci Rep* **6**, 19633 (2016).
41. Shannon, P.*, et al.* Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* **13**, 2498-2504 (2003).
42. Petersen, J.M.*, et al.* One-Carbon Cofactor Intake and Risk of Neural Tube Defects Among Women Who Meet Folic Acid Recommendations: A Multicenter Case-Control Study. *Am J Epidemiol* **188**, 1136-1143 (2019).
43. Kerns, J.C. & Gutierrez, J.L. Thiamin. *Adv Nutr* **8**, 395-397 (2017).
44. Lyon, P., Strippoli, V., Fang, B. & Cimmino, L. B Vitamins and One-Carbon Metabolism: Implications in Human Health and Disease. *Nutrients* **12**(2020).
45. Chandler, A.L.*, et al.* Neural tube defects and maternal intake of micronutrients related to one-carbon metabolism or antioxidant activity. *Birth Defects Res A Clin Mol Teratol* **94**, 864-874 (2012).
46. D'Souza, S.W., Copp, A.J., Greene, N.D.E. & Glazier, J.D. Maternal Inositol Status and Neural Tube Defects: A Role for

the Human Yolk Sac in Embryonic Inositol Delivery? *Adv Nutr* (2020).

47. Sticht, C., De La Torre, C., Parveen, A. & Gretz, N. miRWalk: An online resource for prediction of microRNA binding sites. *PLoS One* **13**, e0206239 (2018).
48. Huang, H.Y.*, et al.* miRTarBase 2020: updates to the experimentally validated microRNA-target interaction database. *Nucleic Acids Res* **48**, D148-D154 (2020).
49. Janky, R.*, et al.* iRegulon: from a gene list to a gene regulatory network using large motif and track collections. *PLoS Comput Biol* **10**, e1003731 (2014).
50. Huang, S.*, et al.* Increased miR-124-3p in microglial exosomes following traumatic brain injury inhibits neuronal inflammation and contributes to neurite outgrowth. *FASEB J* **32**, 512-528 (2018).
51. Wang, Y.*, et al.* miR-124-3p functions as a tumor suppressor in breast cancer by targeting CBL. *BMC Cancer* **16**, 826 (2016).
52. Majid, A.*, et al.* miR-124-3p Suppresses the Invasiveness and Metastasis of Hepatocarcinoma Cells. *Front Mol Biosci* **7**, 223 (2020).
53. Mogilyansky, E. & Rigoutsos, I. The miR-17/92 cluster: a comprehensive update on its genomics, genetics, functions and increasingly important and numerous roles in health and disease. *Cell Death Differ* **20**, 1603-1614 (2013).
54. Eades, G.*, et al.* miR-200a regulates SIRT1 expression and epithelial to mesenchymal transition (EMT)-like transformation in mammary epithelial cells. *J Biol Chem* **286**, 25992-26002 (2011).
55. Xie, N.*, et al.* MicroRNA-142-3p suppresses cell proliferation, invasion and epithelial-to-mesenchymal transition via RAC1-ERK1/2 signaling in colorectal cancer. *Mol Med Rep* **24**(2021).
56. Cousins, R.J. A role of zinc in the regulation of gene expression. *Proc Nutr Soc* **57**, 307-311 (1998).
57. MacDonald, R.S. The role of zinc in growth and cell proliferation. *J Nutr* **130**, 1500S-1508S (2000).
58. Dey, A.C.*, et al.* Maternal and neonatal serum zinc level and its relationship with neural tube defects. *J Health Popul Nutr* **28**, 343-350 (2010).
59. Krężel, A. & Maret, W. The Functions of Metamorphic Metallothioneins in Zinc and Copper Metabolism. *Int J Mol Sci* **18**(2017).
60. Bhattacherjee, V.*, et al.* CBP/p300 and associated transcriptional co-activators exhibit distinct expression patterns during murine craniofacial and neural tube development. *Int J Dev Biol* **53**, 1097-1104 (2009).
61. Lee, S. & Gleeson, J.G. Closing in on Mechanisms of Open Neural Tube Defects. *Trends Neurosci* **43**, 519-532 (2020).
62. Au, K.S.*, et al.* Human myelomeningocele risk and ultra-rare deleterious variants in genes associated with cilium, WNT-signaling, ECM, cytoskeleton and cell migration. *Sci Rep* **11**, 3639 (2021).
63. Yu, J., Wu, Y. & Yang, P. High glucose-induced oxidative stress represses sirtuin deacetylase expression and increases histone acetylation leading to neural tube defects. *J Neurochem* **137**, 371-383 (2016).
64. Taibi, A.*, et al.* Zebrafish churchill regulates developmental gene expression and cell migration. *Dev Dyn* **242**, 614-621 (2013).
65. Snir, M., Ofir, R., Elias, S. & Frank, D. Xenopus laevis POU91 protein, an Oct3/4 homologue, regulates competence transitions from mesoderm to neural cell fates. *EMBO J* **25**, 3664-3674 (2006).
66. Griswold, A.J.*, et al.* A de novo 1.5 Mb microdeletion on chromosome 14q23.2-23.3 in a patient with autism and spherocytosis. *Autism Res* **4**, 221-227 (2011).
67. Vannucchi, H. & Moreno, F.S. Interaction of niacin and zinc metabolism in patients with alcoholic pellagra. *Am J Clin Nutr* **50**, 364-369 (1989).
68. Vannucchi, H., Kutnink, M.D., Sauberlich, M. & Howerde, E. Interaction among niacin, vitamin B6 and zinc in rats receiving ethanol. *Int J Vitam Nutr Res* **56**, 355-362 (1986).
69. Davis, A.P.*, et al.* Comparative Toxicogenomics Database (CTD): update 2021. *Nucleic Acids Res* **49**, D1138-D1143 (2021).
70. Van Hateren, N., Shenton, T. & Borycki, A.G. Expression of avian C-terminal binding proteins (Ctbp1 and Ctbp2) during embryonic development. *Dev Dyn* **235**, 490-495 (2006).
71. Aranow, C. Vitamin D and the immune system. *J Investig Med* **59**, 881-886 (2011).
72. Chen, Y.-H.*, et al.* Supplementation With Vitamin D3 During Pregnancy Protects Against Lipopolysaccharide-Induced Neural Tube Defects Through Improving Placental Folate Transportation. *Toxicological Sciences* **145**, 90-97 (2015).