**Supplementary table 5. Gene function and associated phenotypes in EWAS catalog**

|  |  |  |  |
| --- | --- | --- | --- |
| **Gene Name** | **Gene Function** | **Associated Phenotypes** | **Phenotypes linked to lifestyle and cardiometabolic factors.** |
| **Top genes identified in our EWAS** | | | |
| *NLRC5* | This gene plays a role in cytokine response and antiviral immunity through its inhibition of NF-kappa-B activation and negative regulation of type I interferon signaling pathways. | * Clear cell renal carcinoma * Fetal vs adult liver * Tumor necrosis factor * Fetal vs adult liver * HIV infection * Age * Primary Sjogren’s syndrome * Soluble tumor necrosis factor receptor 2 * Autoantibody production in systemic lupus erythematosus * Sex * Serum total cholesterol * Alcohol consumption per day * Aging * Crohn's disease * parity * Smoking * Papuan ancestry * Alzheimer's disease * Multiple sclerosis * Inflammatory bowel disease | * Fetal vs adult liver * Tumor necrosis factor * Serum total cholesterol * Alcohol consumption per day * Smoking |
| *IMPAD1* | Adenosine monophosphate deaminase 1 catalyzes the deamination of AMP to IMP in skeletal muscle and plays an important role in the purine nucleotide cycle. | * Non-syndromic cleft lip, palate * Age group * Rheumatoid arthritis * Child abuse * Migration in Italy * Gestational age * HIV infection * Sex | * Migration in Italy |
| *CPLX2* | This gene encodes complexin/synaphin gene family are cytosolic proteins that function in synaptic vesicle exocytosis. These proteins bind syntaxin, part of the SNAP receptor. The protein product of this gene binds to the SNAP receptor complex and disrupts it, allowing transmitter release. | * Age 4 vs age 0 * Clear cell renal carcinoma * Pancreatic ductal adenocarcinoma * HIV infection * Gestational age * Rheumatoid arthritis * Ulcerative colitis * Cholesterol esters in chylomicrons and extremely large VLDL * Response to antidepressants | * Cholesterol esters in chylomicrons and extremely large VLDL * Response to antidepressants |
| *EIF4E3* | This gene belongs to the EIF4E family of translational initiation factors that interact with the 5-prime cap structure of mRNA and recruit mRNA to the ribosome. Among its related pathways are Interferon gamma signaling and Innate Immune System. | * Age 4 vs age 0 * Fetal vs adult liver * Clear cell renal carcinoma * Rheumatoid arthritis * HIV infection * Gestational age * Alcohol consumption per day * Sex * Gestational age * Papuan ancestry * Lung function decline * 3-methyl-2-oxobutyrate | * Fetal vs adult liver * Alcohol consumption per day * 3-methyl-2-oxobutyrate |
| *TBL1XR1* | This gene is a member of the WD40 repeat-containing gene family and shares sequence similarity with transducin (beta)-like 1X-linked (TBL1X). The protein encoded by this gene is thought to be a component of both nuclear receptor corepressor (N-CoR) and histone deacetylase 3 (HDAC 3) complexes, and is required for transcriptional activation by a variety of transcription factors | * Clear cell renal carcinoma * Gestational age * Clear cell renal carcinoma * Fetal vs adult liver * Rheumatoid arthritis * HIV infection * Alcohol consumption per day * Ischemic stroke * Hypertensive disorders of pregnancy * Free cholesterol to total lipids ratio in medium HDL * Free cholesterol to total lipids ratio in small VLDL * Total fat free mass * Total lean mass * Phospholipids to total lipids ratio in small VLDL * Smoking * Maternal underweight | * Fetal vs adult liver * Alcohol consumption per day * Ischemic stroke * Free cholesterol to total lipids ratio in medium HDL * Free cholesterol to total lipids ratio in small VLD * Total fat free mass * Total lean mass * Phospholipids to total lipids ratio in small VLDL * Smoking |
| *BLK* | This gene encodes a nonreceptor tyrosine-kinase of the src family of proto-oncogenes that are typically involved in cell proliferation and differentiation. The protein has a role in B-cell receptor signaling and B-cell development. The protein also stimulates insulin synthesis and secretion in response to glucose and enhances the expression of several pancreatic beta-cell transcription factors. | * Clear cell renal carcinoma * Smoking * Gestational age * HIV infection * Crohn's disease * Inflammatory bowel disease * Sex * Mortality * Air pollution exposure * Alzheimer’s diseases * Age | * Smoking * Mortality * Air pollution exposure |
| *SKI* | This gene encodes the nuclear protooncogene protein homolog of avian sarcoma viral (v-ski) oncogene. It functions as a repressor of TGF-beta signaling, and may play a role in neural tube development and muscle differentiation | * Smoking * Clear cell renal carcinoma * Age * Pancreatic ductal adenocarcinoma * HIV infection * Gestational age * Primary Sjogren’s syndrome * Smoking * Ischemic stroke * HIV infection * Alcohol consumption per day * Gestational age * Rheumatoid arthritis * Pubertal timing * BMI * Hepatic Fat * Diastolic Blood Pressure * Attention deficit hyperactivity disorder trajectory. * Aging * Gestational age * C-reactive protein * Pubertal timing * Sex * Systolic Blood Pressure * Child abuse * Ischemic stroke * fasting glucose * Sex * Educational attainment * Papuan ancestry * Long-term prenatal exposure to paracetamol in children with ADHD versus controls * Arsenic exposure * Triglycerides in very small VLDL * Air pollution exposure * Triglycerides in small VLDL * Free cholesterol to total lipids ratio in large LDL * BMI * Clear cell renal carcinoma * Fetal vs adult liver * Age 4 vs age 0 | * Smoking * Ischemic stroke * Alcohol consumption per day * BMI * Hepatic Fat * Diastolic Blood Pressure * C-reactive protein * Systolic Blood Pressure * Fasting glucose * Triglycerides in very small VLDL * Air pollution exposure * Fetal vs adult liver |
| *MEF2D* | This gene is a member of the myocyte-specific enhancer factor 2 (MEF2) family of transcription factors. Members of this family are involved in control of muscle and neuronal cell differentiation and development and are regulated by class II histone deacetylases. | * Fetal vs adult liver * Clear cell renal carcinoma * Pancreatic ductal adenocarcinoma * Smoking * Gestational age * HIV infection * Rheumatoid arthritis * Age * Fetal alcohol spectrum disorder * Arm total mass * Preeclampsia * Inflammatory bowel disease * Gestational age * Alcohol consumption * Birthweight * Nitrogen dioxide exposure | * Fetal vs adult liver * Smoking * Alcohol consumption * Nitrogen dioxide exposure |
| *TLX3* | The protein encoded by this gene is an orphan homeobox protein that encodes a DNA-binding nuclear transcription factor | * Age 4 vs age 0 * Clear cell renal carcinoma * Smoking * Body mass index * Triglycerides to total lipids ratio in IDL * Pancreatic ductal adenocarcinoma * Fetal vs adult liver * HIV infection * Alcohol consumption | * Smoking * Body mass index * Triglycerides to total lipids ratio in IDL * Alcohol consumption * Fetal vs adult liver |
| *ST8SIA1* | The protein encoded by this gene is a type II membrane protein that catalyzes the transfer of sialic acid from CMP-sialic acid to GM3 to produce gangliosides GD3 and GT3.Gangliosides are membrane-bound glycosphingolipids containing sialic acid. Ganglioside GD3 is known to be important for cell adhesion and growth of cultured malignant cells. | * Age 4 vs age 0 * Clear cell renal carcinoma * Pancreatic ductal adenocarcinoma * Gestational age * Sex * Rheumatoid arthritis * HIV infection * Alcohol consumption per day * Papuan ancestry * Smoking * Cholesterol esters to total lipids ratio in large HDL * Total cholesterol to total lipids ratio in large HDL | * Alcohol consumption per day * Smoking * Cholesterol esters to total lipids ratio in large HDL * Total cholesterol to total lipids ratio in large HDL |
| *ANG* | The protein encoded by this gene is an exceedingly potent mediator of new blood vessel formation. It hydrolyzes cellular tRNAs resulting in decreased protein synthesis and is similar to pancreatic ribonuclease. | * Fetal vs adult liver * Primary Sjogren’s syndrome * Rheumatoid arthritis * Gestational age * Papuan ancestry * HIV infection * Cleft lip vs cleft lip palate * Total lipids in chylomicrons and extremely large VLDL * Total cholesterol to total lipids ratio in very large VLDL * Phospholipids in chylomicrons and extremely large VLDL * Smoking * Nitrogen dioxide pollution | * Fetal vs adult liver * Total lipids in chylomicrons and extremely large VLDL * Total cholesterol to total lipids ratio in very large VLDL * Phospholipids in chylomicrons and extremely large VLDL * Smoking * Nitrogen dioxide pollution |
| *CHRM3* | This gene encodes muscarinic receptors. The muscarinic cholinergic receptors belong to a larger family of G protein-coupled receptors. The functional diversity of these receptors is defined by the binding of acetylcholine and includes cellular responses such as adenylate cyclase inhibition, phosphoinositide degeneration, and potassium channel mediation. Muscarinic receptors influence many effects of acetylcholine in the central and peripheral nervous system. | * Clear cell renal carcinoma * Pancreatic ductal adenocarcinoma * Gestational age * Rheumatoid arthritis * Age * HIV infection * Schizophrenia * Conjugated linoleic acid. * Ratio of conjugated linoleic acid to total fatty acids * Smoking | * Conjugated linoleic acid. * Ratio of conjugated linoleic acid to total fatty acids * Smoking |
| *HOXD1* | This gene is a member of the Antp homeobox family and encodes a protein with a homeobox DNA-binding domain. This nuclear protein functions as a sequence-specific transcription factor that is involved in differentiation and limb development | * Age * Pancreatic ductal adenocarcinoma * Rheumatoid arthritis * Smoking * HIV infection | * Smoking |
| *TRIOBP* | his gene encodes a protein with an N-terminal pleckstrin homology domain and a C-terminal coiled-coil region. The protein interacts with trio, which is involved with neural tissue development and controlling actin cytoskeleton organization, cell motility and cell growth. The protein also associates with F-actin and stabilizes F-actin structures. | * Age 4 vs age 0 * Fetal vs adult liver * Rheumatoid arthritis * Smoking * Gestational age * Ethnicity * HIV infection * Type II diabetes mellitus * Alzheimer's disease * Crohn's disease * HIV infection * Total cholesterol in very small VLDL * Alcohol consumption | * Smoking * Type II diabetes mellitus * Alcohol consumption |
| VIPR2 | This gene encodes a receptor for vasoactive intestinal peptide, a small neuropeptide. Vasoactive intestinal peptide is involved in smooth muscle relaxation, exocrine and endocrine secretion, and water and ion flux in lung and intestinal epithelia. | * Age 4 vs age 0 * Clear cell renal carcinoma * Pancreatic ductal adenocarcinoma * Rheumatoid arthritis * Smoking * Mortality * Phospholipids to total lipids ratio in very large VLDL * Sex * Trajectories of social communication deficits * Alcohol consumption * HIV infection * Gestational age * Crohn's disease * Papuan ancestry * HIV infection * Cortical surface area | * Smoking * Phospholipids to total lipids ratio in very large VLDL * Alcohol consumption |