# 1. Gene Aliases

ALDH1A1, Aldehyde Dehydrogenase 1 Family Member A1, RALDH1, ALDH1, PUMB1, 3-Deoxyglucosone Dehydrogenase, Retinaldehyde Dehydrogenase 1, Aldehyde Dehydrogenase 1A1, Retinal Dehydrogenase 1, EC 1.2.1.36, ALDH-E1, RALDH 1, ALHDII, ALDC, Epididymis Secretory Sperm Binding Protein Li 53e, Aldehyde Dehydrogenase 1 Family, Member A1, Aldehyde Dehydrogenase Family 1 Member A1, Aldehyde Dehydrogenase, Liver Cytosolic, Aldehyde Dehydrogenase 1, Soluble, Aldehyde Dehydrogenase, Cytosolic, Epididymis Luminal Protein 12, Acetaldehyde Dehydrogenase 1, Epididymis Luminal Protein 9, ALDH Class 1, EC 1.2.1.19, EC 1.2.1.28, EC 1.2.1.3, HEL-S-53e, EC 1.2.1, ALDH11, RalDH1, HEL-9, HEL12

[<https://www.genecards.org/cgi-bin/carddisp.pl?gene=ALDH1A1&keywords=Aldh1a1>].

# 2. Association with Toxicity and/or Disease at a Transcriptional Level

* ALDH1A1 liver gene expression was identified as a potential therapeutic target in the comorbidity of hypertension and nonalcoholic fatty liver disease (NAFLD) [PMID: 34096467].
* In a alpha-naphthyl isothiocyanate (ANIT)-induced cholestasis rat model, the hepatic expression of LRAT, RAR-beta, and ALDH1A1 in cholestatic rats was decreased compared to the control rats, while the CYP26A1 expression of the liver was not altered [PMID: 30373117].
* Hepatocellular carcinoma (HCC) tissues were examined and there was no significant difference in the ALDH1A1-mRNA level between tumorous and non-tumorous tissues. Tumorous ALDH1A1-mRNA level had no relationship with the clinicopathological features of HCC [PMID: 26160842].
* Colorectal cancer (CRC) liver metastasis (CLM) patient tissue samples showed increased expression of ALDH1A1 mRNA and protein compared to primary colorectal cancer tissue and normal colonic epithelium. ALDH1A1 and IGFBP1 are differentially overexpressed in CLM and may play a dual role, functioning as both tumor suppressors and metastasis promoters in CRC [PMID: 27152521].
* Inflammation, induced by lipopolysaccharide treatment in rats, led to reduced ALDH1A1 mRNA in whole liver regardless of the level of vitamin A in their diet. Retinoic acid treatment led to reduced ALDH1A1 expression only in chow-fed rats. The data suggests that acute inflammation rapidly downregulates ALDH1A1 expression in whole liver while increasing its expression in periportal macrophages. Changes in ALDH1A1 expression appear to be part of the early acute-phase inflammatory response, which has been shown to alter the expression of other retinoid homeostatic genes [PMID: 25926859].
* High ALDH1A1 RNA expression was associated with a longer recurrence-free survival in hepatitis B virus (HBV)-related hepatocellular carcinoma patients [PMID: 28792511].
* In hepatocellular carcinoma (HCC) patient tissues, mRNA expressions of most ALDHs family members were significantly downregulated in patients with HCC from the TCGA database. ALDH1A1 RNA expression was not significantly different between tumor and normal tissues. Low protein expressions of ALDH1A1 was found in HCC tissues [PMID: 34928471].

# 3. Summary of Protein Family and Structure

* Protein Accession: P00352
* Size: 501 amino acids
* Molecular mass: 54862 Da
* Domains: Ald\_DH/histidinol\_DH, Ald\_DH\_C, Ald\_DH\_CS\_CYS, Ald\_DH\_CS\_GLU, Ald\_DH\_N, Aldehyde\_DH\_dom
* Blocks: Aldehyde dehydrogenase
* Family: Belongs to the aldehyde dehydrogenase family.
* Cytosolic dehydrogenase that catalyzes the irreversible oxidation of a wide range of aldehydes to their corresponding carboxylic acid [PMID: 19296407, PMID: 12941160, PMID: 15623782, PMID: 17175089, PMID: 26373694, PMID: 25450233]. Functions downstream of retinol dehydrogenases and catalyzes the oxidation of retinaldehyde into retinoic acid, the second step in the oxidation of retinol/vitamin A into retinoic acid. This pathway is crucial to control the levels of retinol and retinoic acid, two important molecules which excess can be teratogenic and cytotoxic. Also oxidizes aldehydes resulting from lipid peroxidation like (E)-4-hydroxynon-2-enal/HNE, malonaldehyde and hexanal that form protein adducts and are highly cytotoxic. By participating for instance to the clearance of (E)-4-hydroxynon-2-enal/HNE in the lens epithelium prevents the formation of HNE-protein adducts and lens opacification [PMID: 19296407, PMID: 12941160, PMID: 15623782].
* Functions downstream of fructosamine-3-kinase in the fructosamine degradation pathway by catalyzing the oxidation of 3-deoxyglucosone, the carbohydrate product of fructosamine 3-phosphate decomposition, which is itself a potent glycating agent that may react with lysine and arginine side-chains of proteins [PMID: 17175089].
* Has an aminobutyraldehyde dehydrogenase activity and is probably part of an alternative pathway for the biosynthesis of GABA/4-aminobutanoate in midbrain, thereby playing a role in GABAergic synaptic transmission [PMID: 30949457].

# 4. Proteins Known to Interact with Gene Product

## Interactions with experimental support

* **ALDH1A1** Retinal dehydrogenase 1; Can convert/oxidize retinaldehyde to retinoic acid. Binds free retinal and cellular retinol-binding protein-bound retinal (By similarity). May have a broader specificity and oxidize other aldehydes in vivo. [PMID: 12081471, PMID: 16189514, PMID: 25416956, PMID: 12081471, PMID: 16189514, PMID: 25416956]
* **ABCC6** Multidrug resistance-associated protein 6; [Isoform 1]: May participate directly in the active transport of drugs into subcellular organelles or influence drug distribution indirectly. Transports glutathione conjugates as leukotriene-c4 (LTC4) and N-ethylmaleimide S-glutathione (NEM-GS). Belongs to the ABC transporter superfamily. ABCC family. Conjugate transporter (TC 3.A.1.208) subfamily. [PMID: 29704455]
* **AGR2** Anterior gradient protein 2 homolog; Required for MUC2 post-transcriptional synthesis and secretion. May play a role in the production of mucus by intestinal cells (By similarity). Proto-oncogene that may play a role in cell migration, cell differentiation and cell growth. Promotes cell adhesion. [PMID: 30575818]
* **ALDH2** Aldehyde dehydrogenase, mitochondrial; Aldehyde dehydrogenase 2 family member; Belongs to the aldehyde dehydrogenase family. [PMID: 21988832]
* **CDK9** Cyclin-dependent kinase 9; Protein kinase involved in the regulation of transcription. Member of the cyclin-dependent kinase pair (CDK9/cyclin-T) complex, also called positive transcription elongation factor b (P-TEFb), which facilitates the transition from abortive to productive elongation by phosphorylating the CTD (C-terminal domain) of the large subunit of RNA polymerase II (RNAP II) POLR2A, SUPT5H and RDBP. This complex is inactive when in the 7SK snRNP complex form. Phosphorylates EP300, MYOD1, RPB1/POLR2A and AR and the negative elongation factors DSIF and NELF. [PMID: 26209609]
* **CEBPB** CCAAT/enhancer-binding protein beta; Important transcription factor regulating the expression of genes involved in immune and inflammatory responses. Plays also a significant role in adipogenesis, as well as in the gluconeogenic pathway, liver regeneration, and hematopoiesis. The consensus recognition site is 5’-T[TG]NNGNAA[TG]-3’. Its functional capacity is governed by protein interactions and post-translational protein modifications. During early embryogenesis, plays essential and redundant functions with CEBPA. [PMID: 24043631]
* **DBH** Soluble dopamine beta-hydroxylase; Conversion of dopamine to noradrenaline. [PMID: 32814053]
* **GFAP** Glial fibrillary acidic protein; GFAP, a class-III intermediate filament, is a cell-specific marker that, during the development of the central nervous system, distinguishes astrocytes from other glial cells. [PMID: 32814053]
* **IL21** Interleukin-21; Cytokine with immunoregulatory activity. May promote the transition between innate and adaptive immunity. Induces the production of IgG(1) and IgG(3) in B-cells (By similarity). May play a role in proliferation and maturation of natural killer (NK) cells in synergy with IL15. May regulate proliferation of mature B- and T-cells in response to activating stimuli. In synergy with IL15 and IL18 stimulates interferon gamma production in T-cells and NK cells. During T-cell mediated immune response may inhibit dendritic cells (DC) activation and maturation. [PMID: 28514442]
* **JPH3** Junctophilin-3; Junctophilins contribute to the formation of junctional membrane complexes (JMCs) which link the plasma membrane with the endoplasmic or sarcoplasmic reticulum in excitable cells. Provides a structural foundation for functional cross-talk between the cell surface and intracellular calcium release channels. JPH3 is brain- specific and appears to have an active role in certain neurons involved in motor coordination and memory. [PMID: 32814053]
* **MUC1** Mucin-1 subunit alpha; The alpha subunit has cell adhesive properties. Can act both as an adhesion and an anti-adhesion protein. May provide a protective layer on epithelial cells against bacterial and enzyme attack. [PMID: 24043631]
* **NOS2** Nitric oxide synthase, inducible; Produces nitric oxide (NO) which is a messenger molecule with diverse functions throughout the body. In macrophages, NO mediates tumoricidal and bactericidal actions. Also has nitrosylase activity and mediates cysteine S-nitrosylation of cytoplasmic target proteins such PTGS2/COX2 (By similarity). [PMID: 23438482]
* **NUPR1** Nuclear protein 1; Transcription regulator that converts stress signals into a program of gene expression that empowers cells with resistance to the stress induced by a change in their microenvironment. Thereby participates in regulation of many process namely cell-cycle, apoptosis, autophagy and DNA repair responses. Controls cell cycle progression and protects cells from genotoxic stress induced by doxorubicin through the complex formation with TP53 and EP300 that binds CDKN1A promoter leading to transcriptional induction of CDKN1A. [PMID: 21988832]
* **POT1** Protection of telomeres protein 1; Component of the telomerase ribonucleoprotein (RNP) complex that is essential for the replication of chromosome termini. Is a component of the double-stranded telomeric DNA-binding TRF1 complex which is involved in the regulation of telomere length by cis- inhibition of telomerase. Also acts as a single-stranded telomeric DNA- binding protein and thus may act as a downstream effector of the TRF1 complex and may transduce information about telomere maintenance and/or length to the telomere terminus. [PMID: 21988832]

## Interactions with text mining support

* **ADH1A** Alcohol dehydrogenase 1A, alpha polypeptide; Belongs to the zinc-containing alcohol dehydrogenase family. [[https://string-db.org/newstring\_cgi/show\_edge\_details.pl?identifiers=9606.ENSP00000297785 9606.ENSP00000209668](https://string-db.org/newstring_cgi/show_edge_details.pl?identifiers=9606.ENSP00000297785%0D9606.ENSP00000209668)]

# 5. Links to Gene Databases

* GeneCards (human): <https://www.genecards.org/cgi-bin/carddisp.pl?gene=ALDH1A1>
* Harmonizome (human): <https://maayanlab.cloud/Harmonizome/gene/ALDH1A1>
* NCBI (human): <https://www.ncbi.nlm.nih.gov/gene/216>
* NCBI (rat): <https://www.ncbi.nlm.nih.gov/gene/24188>
* Ensemble (human): <https://useast.ensembl.org/Homo_sapiens/Gene/Summary?g=ENSG00000165092>
* Ensemble (rat): <https://useast.ensembl.org/Rattus_norvegicus/Gene/Summary?g=ENSRNOG00000017619>
* Rat Genome Database (rat): <https://rgd.mcw.edu/rgdweb/report/gene/main.html?id=2087>
* Uniprot (human): <https://www.uniprot.org/uniprotkb/P00352>
* Uniprot (rat): <https://www.uniprot.org/uniprotkb/P51647>
* Wikigenes (human): <https://www.wikigenes.org/e/gene/e/216.html>
* Wikigenes (rat): <https://www.wikigenes.org/e/gene/e/24188.html>
* Alphafold (human): <https://alphafold.ebi.ac.uk/entry/P00352>
* Alphafold (rat): <https://alphafold.ebi.ac.uk/entry/P51647>
* PDB (human): <https://www.rcsb.org/structure/4WB9>, <https://www.rcsb.org/structure/4WPN>, <https://www.rcsb.org/structure/5AC2>, <https://www.rcsb.org/structure/5L2M>, <https://www.rcsb.org/structure/5L2N>, <https://www.rcsb.org/structure/5L2O>, <https://www.rcsb.org/structure/6DUM>, <https://www.rcsb.org/structure/7UM9>, <https://www.rcsb.org/structure/8D46>, <https://www.rcsb.org/structure/8ENE>
* PDB (mouse): <https://www.rcsb.org/structure/7YOB>
* PDB (rat): none

# 6. GO Terms, MSigDB Signatures, Pathways Containing Gene with Descriptions of Gene Sets

## **Pathways:**

* **Ethanol oxidation**: Ethanol and related alcohols can be ingested as part of the diet and are formed by microorganisms in the intestinal tract. Ethanol oxidation to acetate occurs primarily in liver cells in a multistep process first described by Racker (1949). First, in the cytosol, ethanol is oxidized to acetaldehyde, with the formation of NADH. Second, in the mitochondrion, acetaldehyde is oxidized to acetate with the formation of NADH. Finally, acetate in the mitochondrion can be condensed with coenzyme A to form acetyl CoA. Polymorphisms in the enzymes catalyzing the first two steps are associated with variation in the efficiency of alcohol catabolism in human populations (Chen et al. 1999; Lange et al. 1976; Jornvall 1985). The molecular mechanism by which cytosolic acetaldehyde enters the mitochondrial matrix is not known (Lemasters 2007). Cytosolic enzymes capable of oxidizing acetaldehyde to acetate have also been identified and characterized in vitro (Inoue et al. 1979) so a purely cytosolic pathway for ethanol oxidation to acetate and conversion to acetyl-CoA can be annotated. The role of this pathway in vivo is unclear, though limited attempts to correlate deficiencies in the cytosolic enzyme with alcohol intolerance have yielded suggestive data (Yoshida et al. 1989). Additional peroxisomal and microsomal pathways for the oxidation of ethanol to acetaldehyde have been described; their physiological significance is unclear and they are not annotated here [<https://reactome.org/PathwayBrowser/#/R-HSA-71384>].
* **Fructose catabolism**: Fructose occurs naturally in foods as a free monosaccharide and as a component of the disaccharide sucrose. It is also widely used as a sweetener. In the body, fructose catabolism occurs in the liver and to a lesser extent in the kidney and small intestine. In these tissues, it is converted to dihydroxyacetone phosphate and D-glyceraldehyde 3-phosphate, two intermediates in the glycolytic pathway, in a sequence of three reactions. It is phosphorylated to form fructose 1-phosphate, which is cleaved by aldolase to yield dihydroxyacetone phosphate and D-glyceraldehyde, and the latter compound is phosphorylated to yield D-glyceraldehyde 3-phosphate. Other pathways exist for the conversion of D-glyceraldehyde to intermediates of glycolysis, but these appear to play only a minor role in normal fructose metabolism (Sillero et al. 1969) [<https://reactome.org/PathwayBrowser/#/R-HSA-70350>].
* **RA biosynthesis pathway**: The major activated retinoid, all-trans-retinoic acid (atRA) is produced by the dehydrogenation of all-trans-retinol (atROL) by members of the short chain dehydrogenase/reductase (SDR) and aldehyde dehydrogenase (RALDH) gene families (Das et al. 2014, Napoli 2012) [<https://reactome.org/PathwayBrowser/#/R-HSA-5365859>].

## GO terms:

**9-cis-retinoic acid biosynthetic process** [The chemical reactions and pathways resulting in the formation of 9-cis-retinoic acid, a metabolically active vitamin A derivative. GO:0042904]

**9-cis-retinoic acid metabolic process** [The chemical reactions and pathways involving 9-cis-retinoic acid, a metabolically active vitamin A derivative. GO:0042905]

**apoptotic process** [A programmed cell death process which begins when a cell receives an internal (e.g. DNA damage) or external signal (e.g. an extracellular death ligand), and proceeds through a series of biochemical events (signaling pathway phase) which trigger an execution phase. The execution phase is the last step of an apoptotic process, and is typically characterized by rounding-up of the cell, retraction of pseudopodes, reduction of cellular volume (pyknosis), chromatin condensation, nuclear fragmentation (karyorrhexis), plasma membrane blebbing and fragmentation of the cell into apoptotic bodies. When the execution phase is completed, the cell has died. GO:0006915]

**cellular detoxification of aldehyde** [Any process carried out at the cellular level that reduces or removes the toxicity of an aldehyde. These may include transport of aldehydes away from sensitive areas and to compartments or complexes whose purpose is sequestration of the toxic substance. GO:0110095]

**embryonic eye morphogenesis** [The process occurring in the embryo by which the anatomical structures of the post-embryonic eye are generated and organized. GO:0048048]

**estrous cycle** [A type of ovulation cycle, which occurs in most mammalian therian females, where the endometrium is resorbed if pregnancy does not occur. GO:0044849]

**fructosamine catabolic process** [The chemical reactions and pathways resulting in the breakdown of fructosamine, a fructose molecule containing an amino group in place of a hydroxyl group. GO:0030392]

**fructose catabolic process** [The chemical reactions and pathways resulting in the breakdown of fructose, the ketohexose arabino-2-hexulose. GO:0006001]

**gamma-aminobutyric acid biosynthetic process** [The chemical reactions and pathways resulting in the formation of gamma-aminobutyric acid (GABA, 4-aminobutyrate), an amino acid which acts as a neurotransmitter in some organisms.|See also the biological process term ‘neurotransmitter biosynthetic process ; GO:0042136’. GO:0009449]

**kidney development** [The process whose specific outcome is the progression of the kidney over time, from its formation to the mature structure. The kidney is an organ that filters the blood and/or excretes the end products of body metabolism in the form of urine. GO:0001822]

**liver development** [The process whose specific outcome is the progression of the liver over time, from its formation to the mature structure. The liver is an exocrine gland which secretes bile and functions in metabolism of protein and carbohydrate and fat, synthesizes substances involved in the clotting of the blood, synthesizes vitamin A, detoxifies poisonous substances, stores glycogen, and breaks down worn-out erythrocytes. GO:0001889]

**maintenance of lens transparency** [A homeostatic process in which the lens is maintained in a highly refractive, transparent state to allow for optimal focusing of light on the retina. GO:0036438]

**midgut development** [The process whose specific outcome is the progression of the midgut over time, from its formation to the mature structure. The midgut is the middle part of the alimentary canal from the stomach, or entrance of the bile duct, to, or including, the large intestine. GO:0007494]

**negative regulation of cold-induced thermogenesis** [Any process that stops, prevents, or reduces the rate of cold-induced thermogenesis. GO:0120163]

**optic cup morphogenesis involved in camera-type eye development** [The invagination of the optic vesicle to form two-walled indentations, the optic cups, that will go on to form the retina. This process begins with the optic vesicle becoming a two-walled structure and its subsequent shape changes. It does not include the fate commitment of cells to become the pigmented retina and the neural retina. An example of this process is found in Mus musculus. GO:0002072]

**positive regulation of apoptotic process** [Any process that activates or increases the frequency, rate or extent of cell death by apoptotic process.|This term should only be used when it is not possible to determine which phase or subtype of the apoptotic process is positively regulated by a gene product. Whenever detailed information is available, the more granular children terms should be used. GO:0043065]

**response to estradiol** [Any process that results in a change in state or activity of a cell or an organism (in terms of movement, secretion, enzyme production, gene expression, etc.) as a result of stimulus by estradiol, a C18 steroid hormone hydroxylated at C3 and C17 that acts as a potent estrogen. GO:0032355]

**response to ethanol** [Any process that results in a change in state or activity of a cell or an organism (in terms of movement, secretion, enzyme production, gene expression, etc.) as a result of an ethanol stimulus. GO:0045471]

**response to organic cyclic compound** [Any process that results in a change in state or activity of a cell or an organism (in terms of movement, secretion, enzyme production, gene expression, etc.) as a result of an organic cyclic compound stimulus. GO:0014070]

**response to oxidative stress** [Any process that results in a change in state or activity of a cell or an organism (in terms of movement, secretion, enzyme production, gene expression, etc.) as a result of oxidative stress, a state often resulting from exposure to high levels of reactive oxygen species, e.g. superoxide anions, hydrogen peroxide (H2O2), and hydroxyl radicals. GO:0006979]

**response to retinoic acid** [Any process that results in a change in state or activity of a cell or an organism (in terms of movement, secretion, enzyme production, gene expression, etc.) as a result of a retinoic acid stimulus. GO:0032526]

**response to xenobiotic stimulus** [Any process that results in a change in state or activity of a cell or an organism (in terms of movement, secretion, enzyme production, gene expression, etc.) as a result of a stimulus from a xenobiotic, a compound foreign to the organim exposed to it. It may be synthesized by another organism (like ampicilin) or it can be a synthetic chemical. GO:0009410]

**retinoic acid biosynthetic process** [The chemical reactions and pathways resulting in the biosynthesis of retinoic acid, one of the three components that makes up vitamin A. GO:0002138]

**retinoic acid metabolic process** [The chemical reactions and pathways involving retinoic acid, one of the three components that makes up vitamin A. GO:0042573]

**retinoid metabolic process** [The chemical reactions and pathways involving retinoids, any member of a class of isoprenoids that contain or are derived from four prenyl groups linked head-to-tail. Retinoids include retinol and retinal and structurally similar natural derivatives or synthetic compounds, but need not have vitamin A activity. GO:0001523]

**retinol metabolic process** [The chemical reactions and pathways involving retinol, one of the three compounds that makes up vitamin A. GO:0042572]

## MSigDB Signatures:

**KEGG\_RETINOL\_METABOLISM**: Retinol metabolism [[https://www.gsea-msigdb.org/gsea/msigdb/human/geneset/KEGG\_RETINOL\_METABOLISM.html]](https://www.gsea-msigdb.org/gsea/msigdb/human/geneset/KEGG_RETINOL_METABOLISM.html)

**WP\_AMINO\_ACID\_METABOLISM**: Amino acid metabolism [[https://www.gsea-msigdb.org/gsea/msigdb/human/geneset/WP\_AMINO\_ACID\_METABOLISM.html]](https://www.gsea-msigdb.org/gsea/msigdb/human/geneset/WP_AMINO_ACID_METABOLISM.html)

**REACTOME\_ETHANOL\_OXIDATION**: Ethanol oxidation [[https://www.gsea-msigdb.org/gsea/msigdb/human/geneset/REACTOME\_ETHANOL\_OXIDATION.html]](https://www.gsea-msigdb.org/gsea/msigdb/human/geneset/REACTOME_ETHANOL_OXIDATION.html)

**REACTOME\_METABOLISM\_OF\_CARBOHYDRATES**: Metabolism of carbohydrates [[https://www.gsea-msigdb.org/gsea/msigdb/human/geneset/REACTOME\_METABOLISM\_OF\_CARBOHYDRATES.html]](https://www.gsea-msigdb.org/gsea/msigdb/human/geneset/REACTOME_METABOLISM_OF_CARBOHYDRATES.html)

**WP\_DISORDERS\_OF\_FRUCTOSE\_METABOLISM**: Disorders of fructose metabolism [[https://www.gsea-msigdb.org/gsea/msigdb/human/geneset/WP\_DISORDERS\_OF\_FRUCTOSE\_METABOLISM.html]](https://www.gsea-msigdb.org/gsea/msigdb/human/geneset/WP_DISORDERS_OF_FRUCTOSE_METABOLISM.html)

**REACTOME\_FRUCTOSE\_METABOLISM**: Fructose metabolism [[https://www.gsea-msigdb.org/gsea/msigdb/human/geneset/REACTOME\_FRUCTOSE\_METABOLISM.html]](https://www.gsea-msigdb.org/gsea/msigdb/human/geneset/REACTOME_FRUCTOSE_METABOLISM.html)

**WP\_FATTY\_ACID\_OMEGA\_OXIDATION**: Fatty acid omega oxidation [[https://www.gsea-msigdb.org/gsea/msigdb/human/geneset/WP\_FATTY\_ACID\_OMEGA\_OXIDATION.html]](https://www.gsea-msigdb.org/gsea/msigdb/human/geneset/WP_FATTY_ACID_OMEGA_OXIDATION.html)

**WP\_FOLATE\_ALCOHOL\_AND\_CANCER\_PATHWAY\_HYPOTHESES**: Folate alcohol and cancer pathway hypotheses [[https://www.gsea-msigdb.org/gsea/msigdb/human/geneset/WP\_FOLATE\_ALCOHOL\_AND\_CANCER\_PATHWAY\_HYPOTHESES.html]](https://www.gsea-msigdb.org/gsea/msigdb/human/geneset/WP_FOLATE_ALCOHOL_AND_CANCER_PATHWAY_HYPOTHESES.html)

**REACTOME\_BIOLOGICAL\_OXIDATIONS**: Biological oxidations [[https://www.gsea-msigdb.org/gsea/msigdb/human/geneset/REACTOME\_BIOLOGICAL\_OXIDATIONS.html]](https://www.gsea-msigdb.org/gsea/msigdb/human/geneset/REACTOME_BIOLOGICAL_OXIDATIONS.html)

**WP\_VITAMIN\_A\_AND\_CAROTENOID\_METABOLISM**: Vitamin A and carotenoid metabolism [[https://www.gsea-msigdb.org/gsea/msigdb/human/geneset/WP\_VITAMIN\_A\_AND\_CAROTENOID\_METABOLISM.html]](https://www.gsea-msigdb.org/gsea/msigdb/human/geneset/WP_VITAMIN_A_AND_CAROTENOID_METABOLISM.html)

**WP\_ETHANOL\_EFFECTS\_ON\_HISTONE\_MODIFICATIONS**: Ethanol effects on histone modifications [[https://www.gsea-msigdb.org/gsea/msigdb/human/geneset/WP\_ETHANOL\_EFFECTS\_ON\_HISTONE\_MODIFICATIONS.html]](https://www.gsea-msigdb.org/gsea/msigdb/human/geneset/WP_ETHANOL_EFFECTS_ON_HISTONE_MODIFICATIONS.html)

**REACTOME\_RA\_BIOSYNTHESIS\_PATHWAY**: RA biosynthesis pathway [[https://www.gsea-msigdb.org/gsea/msigdb/human/geneset/REACTOME\_RA\_BIOSYNTHESIS\_PATHWAY.html]](https://www.gsea-msigdb.org/gsea/msigdb/human/geneset/REACTOME_RA_BIOSYNTHESIS_PATHWAY.html)

**REACTOME\_FRUCTOSE\_CATABOLISM**: Fructose catabolism [[https://www.gsea-msigdb.org/gsea/msigdb/human/geneset/REACTOME\_FRUCTOSE\_CATABOLISM.html]](https://www.gsea-msigdb.org/gsea/msigdb/human/geneset/REACTOME_FRUCTOSE_CATABOLISM.html)

**REACTOME\_SIGNALING\_BY\_NUCLEAR\_RECEPTORS**: Signaling by Nuclear Receptors [[https://www.gsea-msigdb.org/gsea/msigdb/human/geneset/REACTOME\_SIGNALING\_BY\_NUCLEAR\_RECEPTORS.html]](https://www.gsea-msigdb.org/gsea/msigdb/human/geneset/REACTOME_SIGNALING_BY_NUCLEAR_RECEPTORS.html)

**REACTOME\_SIGNALING\_BY\_RETINOIC\_ACID**: Signaling by Retinoic Acid [[https://www.gsea-msigdb.org/gsea/msigdb/human/geneset/REACTOME\_SIGNALING\_BY\_RETINOIC\_ACID.html]](https://www.gsea-msigdb.org/gsea/msigdb/human/geneset/REACTOME_SIGNALING_BY_RETINOIC_ACID.html)

**REACTOME\_PHASE\_I\_FUNCTIONALIZATION\_OF\_COMPOUNDS**: Phase I - Functionalization of compounds [[https://www.gsea-msigdb.org/gsea/msigdb/human/geneset/REACTOME\_PHASE\_I\_FUNCTIONALIZATION\_OF\_COMPOUNDS.html]](https://www.gsea-msigdb.org/gsea/msigdb/human/geneset/REACTOME_PHASE_I_FUNCTIONALIZATION_OF_COMPOUNDS.html)

# 7. Gene Descriptions

**NCBI Gene Summary**: The protein encoded by this gene belongs to the aldehyde dehydrogenase family. Aldehyde dehydrogenase is the next enzyme after alcohol dehydrogenase in the major pathway of alcohol metabolism. There are two major aldehyde dehydrogenase isozymes in the liver, cytosolic and mitochondrial, which are encoded by distinct genes, and can be distinguished by their electrophoretic mobility, kinetic properties, and subcellular localization. This gene encodes the cytosolic isozyme. Studies in mice show that through its role in retinol metabolism, this gene may also be involved in the regulation of the metabolic responses to high-fat diet.

**GeneCards Summary**: ALDH1A1 (Aldehyde Dehydrogenase 1 Family Member A1) is a Protein Coding gene. Diseases associated with ALDH1A1 include Alcohol Dependence and Androgen Insensitivity Syndrome. Among its related pathways are Oxidation by cytochrome P450 and Fructose metabolism. Gene Ontology (GO) annotations related to this gene include oxidoreductase activity and acyl-CoA dehydrogenase activity. An important paralog of this gene is ALDH1A2.

**UniProtKB/Swiss-Prot Summary**: Cytosolic dehydrogenase that catalyzes the irreversible oxidation of a wide range of aldehydes to their corresponding carboxylic acid [PMID: 19296407, PMID: 12941160, PMID: 15623782, PMID: 17175089, PMID: 26373694, PMID: 25450233]. Functions downstream of retinol dehydrogenases and catalyzes the oxidation of retinaldehyde into retinoic acid, the second step in the oxidation of retinol/vitamin A into retinoic acid. This pathway is crucial to control the levels of retinol and retinoic acid, two important molecules which excess can be teratogenic and cytotoxic. Also oxidizes aldehydes resulting from lipid peroxidation like (E)-4-hydroxynon-2-enal/HNE, malonaldehyde and hexanal that form protein adducts and are highly cytotoxic. By participating for instance to the clearance of (E)-4-hydroxynon-2-enal/HNE in the lens epithelium prevents the formation of HNE-protein adducts and lens opacification [PMID: 19296407, PMID: 12941160, PMID: 15623782]. Functions also downstream of fructosamine-3-kinase in the fructosamine degradation pathway by catalyzing the oxidation of 3-deoxyglucosone, the carbohydrate product of fructosamine 3-phosphate decomposition, which is itself a potent glycating agent that may react with lysine and arginine side-chains of proteins [PMID: 17175089]. Has also an aminobutyraldehyde dehydrogenase activity and is probably part of an alternative pathway for the biosynthesis of GABA/4-aminobutanoate in midbrain, thereby playing a role in GABAergic synaptic transmission.

# 8. Cellular Location of Gene Product

Cytoplasmic expression in selected tissues. Localized to the cytosol. Predicted location: Intracellular [<https://www.proteinatlas.org/ENSG00000165092/subcellular>]

# 9. Mechanistic Information

* ALDH1A1 prolongs the stability of transcription factor Gli2 protein in a catalytic-independent manner. The overexpression of ALDH1A1 activated the hedgehog (Hh) signaling pathway and promoted cell growth, migration and invasion in hepatocellular cancer cells. Together, these results illustrate regulatory roles of ALDH1A1 in the activation of the Hh signaling pathway and highlight a novel mechanism for the aberrant activation of the Hh signaling pathway in hepatocellular cancer cells [PMID: 26896768].
* The enzyme euchromatic histone lysine methyltransferase 1 (EHMT1) is expressed in primary and relapse alveolar rhabdomyosarcoma (ARMS) tumors. EHMT1 suppression impaired motility and induced differentiation in ARMS cell lines and reduced tumor progression in a mouse xenograft model in vivo. RNA sequencing of EHMT1-depleted cells revealed downregulation of ALDH1A1 that is associated with cancer stem cells (CSCs) and inhibition of ALDH1A1 expression and activity mimicked EHMT1 depletion phenotypes. Results showed that EHMT1 does not bind to the ALDH1A1 promoter but activates it by stabilizing C/EBP-beta, a known regulator of ALDH1A1 expression [PMID: 34897678].
* Cancer cells form three-dimensional (3D) multicellular aggregates (or spheroids) under non-adherent culture conditions. In ovarian cancer (OC) spheroids, beta-Catenin function and ALDH1A1 expression were increased in OC spheroids vs monolayers and in successive spheroid generations, suggesting that 3D aggregates are enriched in cells with stem cell characteristics. Beta-Catenin knockdown decreased ALDH1A1 expression levels and beta-catenin co-immunoprecipitated with the ALDH1A1 promoter, suggesting that ALDH1A1 is a direct beta-catenin target. An ALDH1A1 small-molecule enzymatic inhibitor also disrupted OC spheroid formation and cell viability [PMID: 24954508].
* Elevated FUBP1 was positively correlated with colorectal cancer (CRC) lymph node metastasis and clinical stage, and negatively associated with overall survival. Overexpression of FUBP1 significantly enhanced CRC cell migration, invasion, tumor sphere formation, and CD133 and ALDH1 expression in vitro, and tumorigenicity in vivo. Mechanistically, FUBP1 promoted the initiation of CSCs by activating Wnt/beta-catenin signaling via directly binding to the promoter of DVL1, a potent activator of beta-catenin [PMID: 34288405].
* In breast cancer cells, loss of lncRNA HULC (Highly Upregulated in Liver Cancer) suppressed the expression of IGF1R and the activation of its downstream PI3K/AKT pathway, while HULC overexpression activated the axis in breast cancer cells. Results showed that HULC functioned as a nuclear lncRNA and epigenetically activated IGF1R by directly binding to the intragenic regulatory elements of the gene, orchestrating intrachromosomal interactions, and promoting histone H3K9 acetylation. The activated HULC-IGF1R/PI3K/AKT pathway mediated tumor resistance to cisplatin through the increased expression of cancer stemness markers, including NANOG, SOX2, OCT4, CD44 and ALDH1A1 [PMID: 35981570].
* Human head and neck squamous cell carcinoma (HNSCC)-derived HSC-3 cells contain a subpopulation of cancer stem cells (CSCs) characterized by high levels of CD44v3 and aldehyde dehydrogenase-1 (ALDH1) expression. These tumor cells also express several stem cell markers (the transcription factors Oct4, Sox2, and Nanog) and display the hallmark CSC properties of self-renewal/clonal formation and the ability to generate heterogeneous cell populations [PMID: 22847005].
* Evaluation of oral cancers for heterogeneity of cancer stem cells (CSC) identified two compartments within the CSC pool. One compartment was detected using a reporter for expression of the H3K4me3 demethylase JARID1B to isolate a JARID1B(high) fraction of cells with stem cell-like function. JARID1B(high) cells expressed oral CSC markers including CD44 and ALDH1 and showed increased PI3K pathway activation. They were distinguished from a fraction in a G0-like cell-cycle state characterized by low reactive oxygen species and suppressed PI3K/AKT signaling [PMID: 27488530].

## Summary

Aldh1a1, as a cytosolic aldehyde dehydrogenase, plays a crucial role in detoxifying aldehydes produced during lipid peroxidation, a process often exacerbated in liver diseases and toxicities [CS: 8]. In conditions like nonalcoholic fatty liver disease (NAFLD), where lipid accumulation and subsequent peroxidation are prominent, Aldh1a1 expression is upregulated as a protective response [CS: 7]. This upregulation aids in reducing the accumulation of cytotoxic aldehydes like (E)-4-hydroxynon-2-enal/HNE, malonaldehyde, and hexanal, which can form harmful protein adducts [CS: 8]. By catalyzing the oxidation of these aldehydes to less harmful carboxylic acids, Aldh1a1 mitigates cellular damage and prevents further disease progression [CS: 9].

In the context of liver toxicity, such as that induced by alpha-naphthyl isothiocyanate (ANIT), the decreased expression of Aldh1a1 contributes to the exacerbation of toxic effects due to reduced detoxification capacity [CS: 7]. Normally, Aldh1a1 would oxidize toxic aldehydes to carboxylic acids, thus neutralizing their harmful effects [CS: 9]. However, in the reduced presence of Aldh1a1, there’s an accumulation of these aldehydes, leading to increased cellular damage and toxicity [CS: 7]. This mechanism also explains the observed upregulation of Aldh1a1 in conditions of acute inflammation, as seen in lipopolysaccharide-treated rats, where its increased expression in periportal macrophages might be an adaptive response to counteract the elevated production of cytotoxic aldehydes during inflammatory processes [CS: 6].

# 10. Upstream Regulators

* AURKA phosphorylates ALDH1A1 at three critical residues which exert a multifaceted regulation over its level, enzymatic activity, and quaternary structure. While all three phosphorylation sites contribute to its increased stability, T267 phosphorylation primarily regulates ALDH1A1 activity. AURKA-mediated phosphorylation rapidly dissociates tetrameric ALDH1A1 into a highly active monomeric species. ALDH1A1 also reciprocates and prevents AURKA degradation, thereby triggering a positive feedback activation loop which drives highly aggressive phenotypes in cancer. Phospho-resistant ALDH1A1 fully reverses EMT and cancer stem cell phenotypes, thus serving as dominant negative, which underscores the clinical significance of the AURKA-ALDH1A1 signaling axis in pancreatic cancer. Increased phosphorylation of ALDH1A1 at the T267 site is observed in human cancers and healthy liver tissues where ALDH1A1 is highly expressed and active, indicating that this regulation is likely crucial both in normal and diseased states [PMID: 28193222].
* Overexpression of c-Jun results in induction of luciferase activity, suggesting that c-Jun transactivates the Aldh1a1 promoter as a homodimer and not as a c-Jun/c-Fos heterodimer at an AP-1 site of the mouse Aldh1a1 gene promoter [PMID: 22740640].
* The RARalpha and C/EBPbeta activate the ALDH1 gene promoter through the RARE and C/EBP response elements, and in Hepa-1 cells, high levels of retinoic acid inhibit this activation by decreasing cellular levels of C/EBPbeta [PMID: 10995752].
* Aldehyde dehydrogenase 1 (ALDH1) is a marker of breast cancer stem cells (BCSCs), and its enzymatic activity regulates tumor stemness. Results show that KK-LC-1 determines the stemness of triple negative breast cancer (TNBC) ALDH+ cells via binding with FAT1 and subsequently promoting its ubiquitination and degradation. This compromises the Hippo pathway and leads to nuclear translocation of YAP1 and ALDH1A1 transcription [PMID: 37147285].
* Hypoxic culture conditions induce and/or upregulate ALDH1 expression in established and primary human glioblastoma multiforme cells [PMID: 24197864].

# 11. Tissues/Cell Type Where Genes are Overexpressed

**Tissue type enchanced**: liver (tissue enhanced) [<https://www.proteinatlas.org/ENSG00000165092/tissue>]

**Cell type enchanced**: hepatocytes, leydig cells (cell type enhanced) [<https://www.proteinatlas.org/ENSG00000165092/single+cell+type>]

# 12. Role of Gene in Other Tissues

* ALDH1A1 is an independent prognostic factor in patients with head and neck squamous cell carcinoma (HNSCC), and the expression level of PDL-1 may be involved in ALDH1A1-mediated poor prognosis in patients with HNSCC [PMID: 32825960].
* High ALDH1A1 mRNA expression was associated with improved disease-free and overall survival for patients with triple-negative breast cancer independent of age at diagnosis, TNM stage and treatment[PMID: 26462023]. However, conflicting results were reported, where in breast carcinoma tissues, the expression of ALDH1 as detected by immunostaining correlated with poor prognosis [PMID: 18371393].
* Results indicate that ALDH1 plays an important role in the metabolism of acetaldehyde in human blood [PMID: 8003124].
* Cytoplasmic and stromal expression of ALDH1A1 is not significantly associated with prognosis either in colon or in rectal cancer. Furthermore, cytoplasmic expression of ALDH1A1 does not predict response to palliative chemotherapy in patients with metastatic diseases. Immunohistochemical expression analysis of ALDH1A1 in colon cancer is useful for the detection of nuclear expression in a small subpopulation of patients and is associated with shorter survival. Cytoplasmic expression fails to be of clinical relevance as prognostic or predictive marker in colorectal cancer [PMID: 22878609].
* High levels of ALDH1A1 mRNA expression is associated with features of poor prognosis, including a poorly differentiated histology and ‘right-sidedness’ of the primary colorectal tumor, and with shorter overall survival. ALDH1A1 is also highly expressed in therapy-surviving tumors and in liver metastases [PMID: 30308036].
* Stratification of the TCGA cohorts by the mutational subtypes of melanoma specifically revealed that gene expression ALDH1A1 correlated with better prognosis in BRAF wild-type melanoma [PMID: 31580832].
* A significant increase in the mRNA expression of ALDH1A1 was detected in cervical intraepithelial neoplasia (CIN) II-III and cervical squamous cell carcinoma (SCC) tissue specimens compared with healthy subjects. ALDH1A1 may serve as biomarkers for the early detection of cervical carcinoma [PMID: 30344706].

# 13. Chemicals Known to Elicit Transcriptional Response of Biomarker in Tissue of Interest

## Compounds that increase expression of the gene:

* 2,2’,4,4’,5,5’-hexachlorobiphenyl [PMID: 20959002]
* 2,2’,4,4’-Tetrabromodiphenyl ether [PMID: 31826744, PMID: 32679240]
* 2-acetamidofluorene [PMID: 22584684]
* 3H-1,2-dithiole-3-thione [PMID: 19162173]
* Aroclor 1254 [PMID: 17851650]
* Echimidine [PMID: 32419051]
* Heliotrine [PMID: 32419051]
* Lasiocarpine [PMID: 32419051]
* Muraglitazar [PMID: 21515302]
* N-nitrosodiethylamine [PMID: 17602206]
* N-nitrosodimethylamine [PMID: 17072980]
* N-nitrosomorpholine [PMID: 19716841]
* Senkirkine [PMID: 32419051]
* Tesaglitazar [PMID: 21515302]
* aflatoxin B1 [PMID: 23385219, PMID: 23630614, PMID: 25378103]
* alpha-hexachlorocyclohexane [PMID: 17785943]
* anthracen-2-amine [PMID: 23038007]
* benzophenone [PMID: 22584684]
* bifenthrin [PMID: 26071804]
* bis(2-ethylhexyl) phthalate [PMID: 22584684]
* cisplatin [PMID: 22023808]
* clofibrate [PMID: 17585979, PMID: 22496397]
* clofibric acid [PMID: 17602206]
* cyhalothrin [PMID: 29727961]
* dichloroacetic acid [PMID: 28962523]
* epoxiconazole [PMID: 25182419]
* erythromycin estolate [PMID: 24412560]
* fipronil [PMID: 23962444]
* monuron [PMID: 22584684]
* nefazodone [PMID: 24136188]
* nimesulide [PMID: 24136188]
* oltipraz [PMID: 22496397]
* p-toluidine [PMID: 27638505]
* pentachlorophenol [PMID: 23892564]
* perfluorooctane-1-sulfonic acid [PMID: 19162173]
* perfluorooctanoic acid [PMID: 19162173]
* pirinixic acid [PMID: 19162173]
* pregnenolone 16alpha-carbonitrile [PMID: 19162173, PMID: 22496397]
* prochloraz [PMID: 25182419]
* propiconazole [PMID: 21278054]
* pyrogallol [PMID: 20362636]
* senecionine [PMID: 32419051]
* streptozocin [PMID: 25905778]
* thioacetamide [PMID: 23411599, PMID: 34492290]
* trichloroethene [PMID: 19254012]
* troglitazone [PMID: 21515302]
* valdecoxib [PMID: 24136188]

## Compounds that decrease expression of the gene:

* 17beta-estradiol [PMID: 32145629]
* cyclosporin A [PMID: 27989131]
* flutamide [PMID: 24793618]
* glafenine [PMID: 24136188]
* methapyrilene [PMID: 30467583]
* octadecanoic acid [PMID: 26739624]
* oleic acid [PMID: 26739624]
* sulfasalazine [PMID: 31830553]
* tetrachloromethane [PMID: 31919559, PMID: 31150632]
* tetracycline [PMID: 24489787]
* trovafloxacin [PMID: 24136188]

# 14. DisGeNet Biomarker Associations to Disease in Organ of Interest

Most relevant biomarkers with lower score or lower probability of association with disease or organ of interest:

* melanoma [PMID: 26391478, PMID: 28106104, PMID: 30377573, PMID: 31580832]
* Neoplasms [PMID: 18082256, PMID: 19329942, PMID: 21957977, PMID: 22012766, PMID: 22938492]
* Non-Small Cell Lung Carcinoma [PMID: 19025616, PMID: 21118965, PMID: 25881507, PMID: 25955300, PMID: 26366059]
* ovarian neoplasm [PMID: 19329942, PMID: 23762304, PMID: 29753392, PMID: 30965686]
* Tumor Cell Invasion [PMID: 19385968, PMID: 22012766, PMID: 23585015, PMID: 24402192, PMID: 25253129]