[Frontiers in Bioscience, Landmark, 20, 335-376, January 1, 2015]

Single amino acid polymorphism in Aldehyde dehydrogenase gene superfamily

J. Priyadharshini Christy[1], C. George Priya Doss[1]

[1]Medical Biotechnology Division, School of Biosciences and Technology, VIT University, Vellore, Tamil Nadu 632014, India

TABLE OF CONTENTS

1. ABSTRACT

The aldehyde dehydrogenase gene superfamily comprises of 19 genes and 3 pseudogenes. These superfamily genes play a vital role in the formation of molecules that are involved in life processes, and detoxification of endogenous and exogenous aldehydes. ALDH superfamily genes associated mutations are implicated in various diseases, such as pyridoxine-dependent seizures, gamma-hydroxybutyric aciduria, type II Hyperprolinemia, Sjogren-Larsson syndrome including cancer and Alzheimer's disease. Accumulation of large DNA variations data especially Single Amino acid Polymorphisms (SAPs) in public databases related to ALDH superfamily genes insisted us to conduct a survey on the disease associated mutations and predict their functional impact on protein structure and function. Overall this study provides an update and highlights the importance of pathogenic mutations in associated diseases. Using KD4v and Project HOPE a computational based platform, we summarized all the deleterious properties of SAPs in ALDH superfamily genes by the providing valuable insight into structural alteration rendered due to mutation. We hope this review might provide a way to define the deleteriousness of a SAP and helps to understand the molecular basis of the associated disease and also permits precise diagnosis and treatment in the near future.

2. INTRODUCTION

Aldehyde dehydrogenases (EC 1.2.1.3.) group of enzymes are mainly involved in the maintenance of cell division, cell differentiation, cell proliferation, cellular homeostasis (redox balance) and cell response to oxidative stress (1, 2). ALDHs oxidize aldehydes that are highly reactive aliphatic or aromatic molecules which participate in various physiological, pathological and pharmacological

processes (3). They utilize NAD or NADP as a co-enzyme to oxidize aldehydes to the carboxylic acid (2, 4). Different classes of aldehydes are meant to perform various functions that are both harmful as well as beneficial to human health. ALDHs detoxify cytotoxic and cytostatic aldehydes into less reactive forms (3). ALDHs exhibit diverse functions in catalytic ester hydrolysis and nitrate reductase activity (3, 5). In addition to catalytic properties, non-catalytic interactions with endogenous compounds such as androgen binding protein in human genital fibroblasts, cholesterol-binding protein in bovine lens epithelium, cytosolic thyroid hormone-binding protein in Xenopus, flavopiridol-binding protein in non-small cell lung carcinomas and daunorubicin binding protein in rat liver by ALDH1A1 (2, 6). Mutations in ALDHs are susceptible to several diseases, which includes various forms of cancers, Alzheimer diseases, γ-hydroxyl butyric aciduria, Pyridoxine-dependent Seizures, Sjogren-Larson Syndrome, type II- hyperprolinemia, etc. (7). Interestingly more than 160 ALDHs cDNAs or genes have been isolated and sequenced from plants, fungi, yeast, bacteria and animals (8) which illustrate their wide distribution in nature. In Eukaryotes, a total of 86 cDNAs or genes have been cloned and segregated into families and sub-families (http://www.aldh.org/superfamily.php) (9). Human ALDH superfamilies possess 19 functional genes clustered into 11 families, and 4 subfamilies with distinct chromosomal location. ALDH1 possess 11 subfamilies; ALDH3 and ALDH5 possess 3 subfamilies, ALDH4, ALDH6, ALDH7, ALDH8 and ALDH9 possess 2 subfamilies; ALDH16 and ALDH18 possess 1 subfamily (9).

3. NOMENCLATURE OF ALDH (ALDEHYDE DEHYDROGENASE) GENE FAMILY

3.1. ALDH1A1

ALDH1A1 is a highly conserved cytosolic homo-tetramer ubiquitously expressed in brain, kidney, lungs, retina, and eye lens and mapped to chromosome 9q21.31 (10). ALDH1A1 catalyzes the oxidation of retinaldehyde to retinoic acid (10, 11). ALDH1A1 has a potential role in supporting cellular homeostasis from oxidative stress and defends ocular surface tissues from reactive oxygen species (6). ALDH1A1 catalyzes 9-cis retinal and all-trans-retinal which serve as ligands for retinoid X receptor (RXR) and the retinoic receptor (RAR) to mediate gene expression in growth and development (12). ALDH1A1 targets the homeodomain transcription factor Pitx3, and expresses itself in dopaminergic neurons to maintain low intra-neuronal levels of DOPAL by catalyzing toxic (3, 4-dihydroxyphenylacetaldehyde) to non-toxic metabolite (3, 4- dihydroxyphenylacetic acid) (DOPAC) (13). ALDH1A1 expresses decreased or lower level of dopaminergic neuron in Parkinson's and schizophrenic patients (12, 14). On the other hand, ALDH1A1 has a prominent role in cancer therapeutics. ALDH1A1 binds to certain anti-cancer drugs (daunorubicin and flavopiridol) and down-regulated in certain carcinomas (11).

3.2. ALDH1A2

ALDH1A2/RALDH2 is a cytosolic homo-tetramer located on the chromosome 15q21.3 (15). It exhibits an expression pattern throughout the early embryonic development (16). ALDH1A2 is highly expressed in embryonic and adult tissues (lungs, liver, kidney, brain, and retina and intestine) (17, 18). ALDH1A2 act as the key regulator of RA (Retinoic acid) synthesis, and induces tissue development, promotes cell differentiation, cell growth, apoptosis, and anticancer activity (19, 20, 21). Like ALDH1A1, ALDH1A2 catalyzes all tans –retinal and 9-cis retinal. ALDH1A2 has a defensive role against ethanol toxicity, which is either through retinoic acid synthesis or the acetaldehyde detoxification (22). A research study has demonstrated that ALDH1A2 has a possible role in prostate cancer, and retinoic acid is involved in the prevention and treatment of prostate cancer (22). In human prostate cancer cell lines, ALDH1A2 expression is highly induced by DNA methylation and over expression of ALDH1A2 (wild type) inhibits cell growth (21).

3.3. ALDH1A3

ALDH1A3 (RALDH3 / ALDH6) is a cytosolic homo-tetramer, located on the chromosome 15q26.3 (15). In humans, ALDH1A3 is highly expressed in fetal nasal mucosa, breast, kidney, stomach and salivary glands (23). Like ALDH1A2, ALDH1A3 is differentially expressed in early embryonic development (head and forebrain) and oxidizes all-trans-retinal and 9-cis-retinal to retinoic acid. ALDH1A3 plays a critical role in cancer and mitigates oxidative stress by detoxifying LPO-derived aldehydes. It has been studied that ALDH1A3 expression is up-regulated in colon cancer cell line and carries a wild- type p^{53} transgene controlled by lactose operon (24), and down-regulated in (MCF -7 cells) human breast cancer cells (25). In glioblastoma cells, ALDH1A3 is highly induced by antitumor agent – IL -13 cytotoxin (26).

3.4. ALDH1B1

ALDH1B1 a NAD^+ dependent mitochondrial homo-tetramer located on the chromosome 9p13.2 (15). In humans, ALDH1B1 is highly expressed in placenta, brain, heart, lungs, liver, skeletal muscle and kidney (27). ALDH1B1 shares 75 % sequence homology with ALDH2. ALDH1B1 metabolizes aliphatic aldehydes and displays a high affinity for acetaldehyde oxidation (27). ALDH1B1 has an increased contribution to ethanol detoxification and up-regulated in response to UV light to protect the cornea from UV exposure (28).

3.5. ALDH1L1

ALDH1L1 is a multi-domain homo-tetramer located on the chromosome 3q21.3 (15), highly expressed in liver pancreas, kidney and liver, and moderately expressed in heart, skeletal muscle, prostate, ovary, and thymus (29). It is a large cytosolic protein of about 98.8 kDa subunits and consists of 2 domains namely; carboxy-terminal domain and amino-terminal formyl transferase domain (30). ALDH1L1 catalyzes NAD $(P)^+$ dependent oxidation of 10-formyl tetrahydrofolate to tetrahydrofolate, and NAD $(P)^+$ independent hydrolysis of 10-formyl tetrahydrofolate to tetrahydrofolate and formate (31, 32). ALDH1L1 is required for purine biosynthesis; inhibition of ALDH1L1 results in the depletion of intracellular 10- formyl tetrahydrofolate, followed by loss of de-novo purine biosynthesis, which results in diminished DNA / RNA availability. This normally induces apoptotic cell death; therefore down-regulation of ALDH1L1 in tumors enhances cell proliferation (29). ALDH1L1 is a potent inducer of methanol toxicity and highly toxic to ocular systems and retinal muller cells (33).

3.6. ALDH1L2

ALDH1L2 is a fusion gene composed of 23 exons and located on the chromosome 12q23.3. It encodes a protein of 923-amino acids, (15) and 72.3 % identical to that of ALDH1L1. ALDH1L2 consists of three domains; a) aldehyde dehydrogenase domain at the carboxy terminal, b) formyl transferase carboxyl-terminal domain in the middle, c) formyl-trans- N-formyl transferase at the amino terminal. The expression pattern of ALDH1L2 is upregulated in *invitro* studies, when breast cancer cells treated with Indometacin (34).

3.7. ALDH2

ALDH2 encodes a mitochondrial enzyme located on the chromosome 12q24.12 (15). It plays an important role in the acetaldehyde detoxification. ALDH2 is highly expressed in brain, lung, liver and kidney (35). ALDH2 has a pivotal role in the bioactivation of nitroglycerin and act as a nitrate reductase, a principle enzyme primarily necessary for the activation of nitroglycerin, (used to treat heart failure and angina) (36). It is associated with hypertension in Japanese populations (37) and myocardial infarction in Korean populations (38). Like ALDH1A1, ALDH2 is implicated in Dopamine metabolism. ALDH2*2 allele has a single-base pair substitution (G/C→A/T) in exon 12 causing E504K substitution. In the mutant state Glutamic acid are replaced by Lysine at position 504. Nearly 40% of the East Asian descents possess a functional polymorphism in ALDH2*2. The inactive form of ALDH2 favors accumulation of acetaldehyde and alcohol-induced flushing reaction (38- 42). Alcoholic individual who possesses ALDH2*2 alleles has an increased affinity for esophageal cancer, head and neck cancer, colorectal cancer, and late- onset Alzheimer disease (43). Among American Jewish, who is alcoholic an additional polymorphic locus in the promoter (G/A) regions of ALDH2 affected the ALDH2 activity, mainly by transcriptional mechanisms (44). Site-directed mutagenesis of Glu 268, Cys 302, Lys 192, and Glu 399 residue, changes the rate-limiting step from deacylation to hydride transfer step. The mutant Lys 192 alters the substrate specificity from aliphatic aldehydes to aromatic aldehydes (45). On the contrary similar to ALDH1A1, ALDH2 displays a similar binding capability with exogenous compounds (e.g. acetaminophen binding protein). Inhibition of ALDH2 results in the accumulation of 3, 4-dihydroxy phenyl acetaldehyde (neurotoxic dopamine metabolites), thereby contributes Parkinson's diseases (46).

3.8. ALDH3A1

ALDH3A1 is a cytosolic homodimer located on the chromosome 17p11.2 (15). Moreover, constitutively expressed in the cornea, lungs, esophagus and stomach (47). ALDH3A1 shares a significant corneal protein called "Crystalline," a structural element in the eye which maintains the corneal transparency and protects the lens crystalline from the hydroxyl radicals; direct absorption of UV light and from metabolism of cytotoxic aldehydes generated due to UV-induced lipid peroxidation (47, 48). ALDH3A1 catalyzes the LPO- derived aldehydes (α, ß hydroxy alkenals). By maintaining the redox balance, ALDH3A1 plays a crucial role in cellular homeostasis and contributes antioxidant capacity of the cell by generating NADH or NADPH (49, 50). Notably the enzymatic activity

of ALDH3A1 generates NADPH, which is associated with the regeneration of reduced glutathione (GSH) from its oxidized form (GSSG) *via* glutathione peroxidase or reductase system (49, 50). Like ALDH1A1, ALDH3A1 defends the ocular surface tissues from reactive oxygen species as well as supports the cellular homeostasis from oxidative stress. It maintains the integrity of the lens and cornea and serves as chaperones to prevent aggregation of misfold proteins. Accumulation of aldehyde results in the deficiency of ALDH3A1, which in turn reduces cellular growth rate. In *invitro* studies, ALDH3A1 was active against toxic agents (Hydrogen peroxide, mitomycin C and etoposide) and protective against DNA damage and apoptosis (51). ALDH3A1 acts as a candidate gene in the pathogenesis of esophageal squamous cell carcinoma (52) and potent diagnostic marker for non-small-cell lung cancer (53). Expression of ALDH3A1 is induced by several xenobiotic, which includes polycyclic hydrocarbons, 3-methylcholanthrene mainly through multiple xenobiotic response elements (XREs) (54, 55).

3.9. ALDH3A2

ALDH3A2 or FALDH is a microsomal NAD^+ dependent homodimer located on the chromosome 17p11.2. (15). It has a high affinity for straight and branched chain aliphatic aldehydes (both saturated and unsaturated aldehydes) (56). ALDH3A2 is highly essential for the oxidation of long-chain aliphatic aldehydes to fatty acids. The long-chain aliphatic aldehyde has a significant functional role in several inborn errors of metabolism. In human, ALDH3A2 is highly expressed in liver, kidney, intestine, stomach, skeletal muscles, skin, lung, pancreas, placenta, heart and brain (57, 58). Deficiency of ALDH3A2 is associated with an autosomal recessive disorder “Sjögren–Larsson syndrome” which is characterized by ichthyosis, mental retardation and spasticity (59). So far, seventy-two mutations have been discovered in SLS patients; which includes amino acid substitutions, insertions, deletions, and splicing errors (60). SLS pathogenesis is mainly due to the accumulation of lipids (in the membrane of the brain and skin), defective eicosanoid metabolism and aldehyde Schiff- base adducts (with amine-containing lipids and proteins) (60). In contrast to “Sjögren–Larsson syndrome”, ALDH3A2 has a potential role in diabetes and oxidative stress-induced complications (61).

3.10. ALDH3B1

ALDH3B1 is a homodimer located on the chromosome 11q13.2 (15) and was first cloned and sequenced by Hsu *et al.* (62). It is highly expressed in kidney and liver and moderately expressed in lung and brain (striatum, cortex, brain stem, cerebellum and hippocampus) (63). ALDH3B1 utilizes NAD^+ or $NADP^+$ as a cofactor and protects the brain from the detoxification of aldehydes produced during oxidative stress (63). ALDH3B1 possesses distinct substrate specificity towards medium and long chain unsaturated and saturated aliphatic aldehydes. A SNP in intron 2 (rs581105 (T/G)) was found to be in association with paranoid schizophrenia and also in the alteration of dopamine metabolism (64, 65). Few studies have described DOPAL (dopamine derived aldehydes) as a poor substrate of metabolizer (63), but ALDH3B1 protects the brain from the detoxification of another aldehyde (Octanol) which may be highly toxic. Based on this study, over-expression of ALDH3B1 *in vitro* tends to protect the cells from highly toxic LPO- derived aldehyde (63).

3.11. ALDH3B2

ALDH3B2 is a mitochondrial homodimer, a pseudogene located on the chromosome 11q13.2 (15). ALDH3B2 contains an in-frame stop codon at 17^{th} codon from methionine. ALDH3B2 shares 83% sequences identity with ALDH3B1. In the human, ALDH3B2 transcripts are found in human salivary gland tissue which indicates that ALDH3B2 possess promoter activity (66). In addition, microarray analysis data indicates that ALDH3B2 is highly expressed in brain, liver, kidney, prostrate and lung (67).

3.12. ALDH4A1

ALDH4A1/Pyroline-5-carboxylase dehydrogenase is a mitochondrial matrix homodimer located on the chromosome 1p36.13 (15) and highly expressed in liver, skeletal muscles and kidney (68). ALDH4A1 is transcriptionally induced by p^{53}, whereby upregulated in response to DNA damage and possessed an important role in DNA repair and cell survival (69). ALDH4A1 has a protective role during oxidative stress; it catalyzes the irreversible conversion of delta-1-pyroline-5-carboxylate (P5C), which is derived either from proline or ornithine to glutathione, which is an important step to connect urea and TCA cycle. ALDH4A1 protects the cell from oxidative stress and responsible for the oxidation of short and medium chain aliphatic LPO-derived aldehydes (70). Mutation in ALDH4A1 leads to an autosomal recessive disorder “Type- II Hyperprolinemia” a neurological manifestation which is characterized by the accumulation of P5C and Proline (71). Patients with type II Hyperprolinemia

have a missense mutation (S352L) that results in abnormal ALDH4A1 activity (72). The phenotypes of type II Hyperprolinemia are associated with P5C mediated deactivation of vitamin B6 derivative (Pyridoxal phosphate) through Knoevenagel –type condensation reaction (73).

3.13. ALDH5A1

ALDH5A1 is often referred to as succinic semi-aldehyde dehydrogenase. It encodes a mitochondrial enzyme which is mainly involved in the catabolism of neurotransmitter (GABA) and located on the chromosome 6p22.3 (15). It is expressed highly in liver, kidney, skeletal muscle and brain (74). Succinic semi-aldehyde is oxidized by ALDH5A1 to succinate. A small fraction is reduced by cytosolic succinic semialdehyde reductase to gamma- hydroxybutyric acid (75), a compound with neurotransmitter and neuro-modulator like propensity found in the central nervous system (76). Several mutations have been identified in the coding, as well as in the non-coding regions. Mutation in ALDH5A1 leads to a rare autosomal recessive metabolic disorder 4-amino butyric acid degradation, which is characterized by the accumulation of GABA and 4-hydroxybutyric acid leading to moderate to severe phenotypic neurological disorder which includes mental retardation, ataxia and seizures (77).

3.14. ALDH6A1

ALDH6A1 is a mitochondrial tetramer also known as CoA - dependent methyl malonate-semialdehyde dehydrogenase located on the chromosome 14q24.3 (15) and expressed in liver, kidney and heart and moderately in muscle and brain (78). Most isozyme forms of ALDH shares the same catalytic mechanism for both esterase and dehydrogenase activity, but ALDH6A1 utilizes CoA as a cofactor rather than NAD (P). ALDH6A1 catalyzes the irreversible oxidative decarboxylation of malonate and methyl malonate semialdehyde to acetyl and propionyl-CoA. ALDH6A1 is mainly involved in pyrimidine and valine catabolism. In lipogenesis, during differentiation of 3T3-L1 fibroblasts into mature adipocytes, ALDH6A1 is upregulated by valine carbon utilizations (79). Mutations in ALDH6A1 lead to a disorder characterized by an increased level of beta-alanine and 3-hydroxypropionic acid, 3-amino and 3-hydroxyisobutyric acids, accompanied by increased degree of psychomotor delay (80).

3.15. ALDH7A1

ALDH7A1 is a homotetramer located on the chromosome 5q23.2 (15) and highly expressed in the eye, ovary, heart and kidney and moderately expressed in liver, spleen, muscle, lung and brain (15). ALDH7A1 has a potential role in pipecolic acid pathway of lysine catabolism, thereby catalyzes the oxidation of AASA (Alpha- aminoadipic semialdehyde) to alpha-aminoadipate (81). ALDH7A1 possess three pseudogenes namely; ALDH7A1P1, ALDH7A1P2, and ALDH7A1P3 located on different chromosomes (15). Among the three pseudogenes, ALDH7A1P1 meets the HSNC criteria for perfect pseudogenes. ALDH7A1P2 and ALDH7A1P3 are in sequence homology with ALDH7A1 (15). ALDH7A1 generates betaine (Osmolytes) and protects the cell against hyperosmotic stress, and metabolize toxic aldehydes which are mainly derived from lipid peroxidation (82). ALDH7A1 mutations are the molecular basis for pyridoxine-dependent epilepsy, an autosomal disorder which is characterized by seizures during infancy and early childhood (83). Increased concentration of AASA and piperidiene-6-carboxylate are seen in patients with ALDH7A1 deficiency which leads to inactivation and reduction of the coenzyme PLP mainly through Knoevenagel adduction reaction (83). PLP has a significant role in neuro-transmitter pathways, which includes GABA, serotonin, and non-adrenalin (84). Reduction of PLP results in subsequent disruption in the neurotransmitter metabolism (83).

3.16. ALDH8A1

ALDH8A1 or RALDH4 is a cytosolic enzyme located on the chromosome 6q23.3 (15) Moreover, highly expressed in the liver and kidney and moderately expressed in brain, spinal cord, mammary gland, thymus, adrenal, prostate, and gastrointestinal tract (85). ALDH8A1 oxidizes retinal and plays an important role in 9-*cis*-retinal (85). Like ALDH1 family, ALDH8A1 participate in the biosynthesis of RA (Retinoic acid) *via* oxidation of retinal and metabolizes the aliphatic aldehydes which include acetaldehyde, decanal, octanal, hexanal, propanal, glutaraldehydes, and SSA (85).

3.17. ALDH9A1

ALDH9A1 is a cytosolic tetramer located on the chromosome 1q24.1 (15) and expressed in liver, skeletal muscle, kidney and brain (86). It was first purified and characterized by Kurys (15). ALDH9A1 catalyzes the dehydrogenation of gamma-aminobutyraldehyde and amino aldehyde from poly chromosome amines and oxidizes gamma-trimethyl amino butyraldehyde and betaine aldehyde (involved carnitine biosynthesis) (87). ALDH9A1 plays a significant role in the metabolism of DOPAL,

a catecholamine-derived aldehyde. High affinity for DOPAL and gamma-amino butyraldehyde affects GABA as well as dopamine pathway and brain development. Interestingly, ALDH9A1 is a candidate gene in human non-alcoholic steatohepatitis (an inflammatory –mediated aggravation) (88).

3.18. ALDH16A1

ALDH16A1 is located on the chromosome 19q13.33 (15) and highly expressed in bone marrow, heart, kidney and lung (15). ALDH16A1 was identified and sequenced from the cDNA library of human uterine, by National institutes of Health Mammalian Gene Collection program (89). The 2627 base pair transcript is composed of 17 exons and encodes 802 amino acid proteins which contains NAD^+ dependent ALDH domain. The orthologues of ALDH16A1 are 98% identical to that of a chimpanzee, and 81-% identity to mouse, rat dog and zebra fish.

3.19. ALDH18A1

ALDH18A1/Pyrolline–5carboxylase–synthetase (P5C) is a bifunctional inner mitochondrial enzyme located on the chromosome 10q24.1 (15) Moreover, expressed highly in pancreas, ovary, testis and kidney and moderately expressed in colon, placenta, small intestine, heart and skeletal muscle (90). ALDH18A1 catalyzes the reduction of L-glutamate to Δ (1)-pyroline-5-carboxylase, a significant step in the biosynthesis of proline, ornithine and arginine (90). Mutation in ALDH18A1 exhibits hyperammonemia, hypoornithenemia, hypocitrullinemia, hypoargininemia, hypoprolinemia associated with cataract formation, neurodegeneration and connective tissue anomalies (91, 92). Missense mutations result in the replacement of leucine (highly conserved) with serine at position 396 (L396S), and arginine to glutamine (conserved residue) at position 84 (R84Q) within the gamma-glutamyl kinase domain (91). Alternative splicing of ALDH18A1 results in the formations of two respective isoforms namely ALDH18A1 _i1 and aldh18a1 _i2, these distinct isoforms differ by 2 amino acid inserts, closer to gamma-glutamyl kinase active sites, mutations in either of the isoform may lead to ALDH18A1 metabolic abnormalities (93).

4. FUNCTIONAL EFFECTS OF SAPS ON PROTEIN STRUCTURE

Technological advances in sequencing technologies have revolutionized the field of biology and medicine. Last decade has witnessed the drastic shift in genomic research towards variants especially Single Nucleotide Polymorphisms (SNPs) and also interlinking them with disease susceptibility. SNPs information are made available in online databases such as HGVBase (Human Genome Variation database) (94), dbSNP (95), JSNP (Japanese Single Nucleotide polymorphism database) (96), Human Genome Mutation database (HGMD) (97), Uniprot (98). SNPs can be classified based on the location in the coding and non-coding region. SNPs are used as markers to identify disease-causing genes and understanding the molecular mechanisms of sequence evolutions and inter-individual variability in drug response. SNPs in the non-coding regions alter the expression patterns of the genes, mainly by affecting the regulatory elements, intronic regions and active splice sites, leading to alternative splicing. SNPs in the coding region alter the amino acid sequence of expressed proteins either through premature termination codon or missense substitutions. Nonsynonymous SNPs (nsSNPs) also known as SAPs and are categorized into deleterious and neutral, but most of the SAPs are not deleterious. Therefore, it is very important to discriminate harmful SAPs from the neutral ones. To predict the impact of SAPs (Single amino acid polymorphism), many computational methods have been developed in predicting the effect of mutation based on the utilization of sequence, structure and sequence and structure information. Structure-based methods classify SAPs as deleterious or neutral based on the physicochemical properties of amino acids and protein structural properties such as solvent accessibility, secondary structure elements and disulphide bridges. They make their prediction on the protein function, protein stability and pathogenecity.

Knowledge of three-dimensional structures is very important in mapping the deleterious SAPs to a known protein structure which provides valuable information on the effect of mutation based on charge, size, dynamics, solvent accessibility, secondary structure information and hydrogen bonding. This information can guide us in understanding the molecular basis of disease and also in drug designing. Large numbers of protein sequences are available in public databases when compared to the solved structures by X-Ray and NMR which is limited by cost and time effectiveness. This makes way for computational methods such as SWISS-MODEL (99), MODELLOR (100), I-TASSER (101), ROBETTA (102), YASARA (103) etc. to predict 3D structure by homology modeling, fold recognition, and *ab-initio* techniques. 3D visualization of

macromolecules, including protein and protein-ligand complexes can be performed using PyMOL (104), Rasmol (105), and SWISSPDB VIEWER (106). These are molecular graphic visualization program intended to visualize protein structure, nucleic acids and small molecules. Molecular dynamics helps to understand the molecular assemblies in terms of structure, function and its interactions between them. Visual molecular dynamics (VMD) is a visualizing molecular graphics programs designed to analyze and visualize the molecular assemblies of a particular proteins and nucleic acids (107). VMD provides a complete set of a graphical interface and provides a command script and has been designed to animate MD simulation trajectories. VMD makes use of NAMD (107) and MDCOMM (107) software for visualization and simulation program (107). CASP (108) makes use of realistic modeling servers namely LEE- SERVER, YASARA, and ROBETTA. YASARA runs the molecular simulation using atom force field (derived from Amber); to help MODELLER and to create perfect alignment LEE- SERVER uses conformational space annealing (108). Till date many computational methods have been put forward in understanding the effects of SNPs and SAPs based on 3D structure of the protein such as SNP@D (109), TopoSNP (110), ModSNP (111), LS- SNP (112), StSNP (113), ColiSNP (114), etc. Mod SNP database contains information's necessary for the SAPs *via* through homology modeling. It provides valuable information's regarding the structural analysis of SAPs (111). SAAP pipeline (115) helps to analysis the mutations based on the structural analysis. Initial considerations were focused on the mapping of mutations on to the protein structure (116). SNP@3D provides the information of SAPs within the protein domain. It collects all the SAPs information is from the dbSNP (109); all the information's tends to match the SCOP and Pfam domain sequences, which are assigned to Ensemble. Both the 2D and 3D maps within the domains are provided. TopoSNP facilitates the structural mapping analysis of both the disease and non-disease association of SAPs; it allows easy prediction with structural entropy and structural characterization of SAPs (110). Conservation and entropy are calculated using HMM (Hidden Markov Model), comparison and homology modeling are favored by PSSM (Position-specific scoring matrices). LS-SNP initiates the human SAPs based on the evolutionary conservation, binding interactions and structural features (112), it allows accession query and selects data's from HUGO gene, dbSNP, UniProt, and PDB ID. StSNP database collects protein structure information from PDB, SNP form dbSNP and pathway information's from Kyoto Encyclopedia of Genes and Genome (116). StSNP provides the users to compare the distribution of SAPs in human proteins, protein- ligand interactions and protein- protein interactions. The multiple structure viewer Friend (117) provides mapping of SNPs and SAPs from the desired protein structure, mapping on to the protein structure provides homology modeling. StSNP allows the user to visualize and collects possible information's about the desired variations in the protein structure. ColiSNP a database server maps SAPs on to the protein 3D structures (114). Using Jmol or Rasmol it enables visualization of 3D protein structures and based on the stability and function of the protein it evaluates the effects of mutations. KD4V (Comprehensive knowledge Discovery System for Missense Variations) server tends to characterize and predict the phenotypic effects (deleterious or neutral) of missense variants (118). The server provides a set of rules learned by Induction Logic Programming (ILP) on a set of missense variants described by conservation, physiochemical, functional and 3D structure predicates. Hope Project (Have your protein explained) are used to study and analyze the insight structural features of native and mutant model, and provides 3-D structure visualization of the mutated proteins (119).

4.1. KD4v

KD4v provides the users to discover and exploit the knowledge between the human disease phenotypes and computed aided mutations. KD4v makes use of IPL methods (120) to map the 3D structure, as a set involved in human diseases. SAPs of all human proteins and the 3D structure mappings are annotated by MSV3d pipeline (Missense Variant mapped to 3D structure) (121). ILP rules of KD4v are based on sixteen predicates, which are based on structure/sequence /evolutionary analysis. Being robust and efficient, it helps to characterize between deleterious SAPs from neutral (122,123). For compiling, a dataset KD4V makes use of PolyPhen-2 training set which are extracted from SwissVar (124). SM2PH helps users to predict the provided information's related to phenotypic descriptions, tissue expressions and protein-protein interactions (125), and provides various parameters which include physicochemical changes, status of mutant residues with their functional effects and the conserved pattern of the mutant residue. KD4v makes use of SCOP to characterize the 3D protein model (126); nearly 63,000 nsSNPs for known 10713 proteins with 3D structures are currently available.

MACSIMS (127) are used to annotate the multiple alignments, importantly it makes use of UniProt and PDB sequences with provided information includes functional descriptions, taxonomic data, and potential disordered regions, domains with known 3D structure and conservation patterns of residues and domains. MODELLOR (128) and CSU (129) and I- Mutant (130) are used to construct the 3D structure model.

4.2. HOPE

HOPE is an automatic protein mutant analyzer and designed to understand the molecular basis of diseases and predict the disease-related phenotypes mainly due to mutations. HOPE makes use of FASTA or BLAST (as input sequences) (131), UniProt database (132), and PDB (133). To obtain DAS prediction, UniProt database identifies the entry identity and the accession code of the respective proteins. HOPE makes use of PDB- file with its respective residues to be mutated. For a proper homology modeling, BLAST or FASTA search is used against the PDB. Using twinset version of YASARA a perfect homology modeling is constructed. YASARA contains an automatic homology modeling script and makes use of FASTA or BLAST as an input. Automatically the script performs loop building, side chain modeling, sequence alignment, and energy minimization. For sequence alignment, ClustalW is used (134). WHAT IF Web server makes use of DSSP to calculate the secondary prediction, the protein structure of interest, either PDB- file or a homology modeling (135), and favors to calculate the hydrogen bonds, accessibility, salt bridges, ligands or ion interactions, variability and mutability (136). UniProt database is used to retrieve the features that can be mapped onto the sequences (137). Conservation score of the respective mutated residue is calculated using HSSP (138). DAS server are used to predict the accessibility (using PHDacc) (139), transmembrane regions (using Phobius) (140), secondary structure by PHDsec (139), phosphorylation sites by NetPhos (141). The whole decision is implemented by means of Groovy (a dynamic language) which tends to run on JAVA virtual machine.

5. METABOLIC DISEASES ASSOCIATED WITH ALDH SUPER GENE FAMILIES

In this study, we had analyzed the impact of experimentally proved mutations in ALDH superfamily genes (Table 1) and mapped the deleterious SAPs from the desired protein 3D structures using KD4v and Project HOPE. KD4v leads a way to summarize the charge, size, hydrophobicity, polarity, solvent accessibility, secondary structure elements, disulphide bridges, protein stability, hydrogen bond formation and pathogenecity, etc. The interpretations of SAPs for all ALDH supergene families by KD4v are tabulated in (Table 2). Project HOPE was used to analyze and visualize the 3 D structural prediction, and helped to characterize the physical- chemical properties of both mutant and wild-type residue of ALDH supergene families. The effects of mutations and properties of ALDH supergene families by HOPE analysis are tabulated in (Table 3). From our survey analysis, we have proved that the below mentioned amino acid variants for each ALDH genes are in good correlation with the clinically proved results which are obtained from *in- vitro/in- vivo results.*

5.1. Human Congenital heart disease

Congenital heart disease (CHD) (OMIM*603687) is a multifactorial polygenic disease (142). The familial forms of CHD are hypertrophic cardiomyopathy, (TOF) Tetralogy of Fallot, (ASD) Atrial Septal defects and (VSD) Ventricular Septal defects, these forms, are well explained by haploinsufficiency of genes (with their major roles in cardiac development) (143, 144, 145). The instability of ALDH1A2 is quite critical and permissive throughout the cardiac development and weakens the cardiac morphogenesis (146). RA (Retinoic acid) is the modifier of CHD; Marilène Pavan *et al. (*2009) used ALDH1A2 to understand the relationship between RA and CHD, the genetic screening was utilized to detect the mutations and family- based genetic association study was used to predict the genetic variation at the ALDH1A2 locus. In Tetralogy of Fallot, Marilène Pavan *et al. (*2009) described that A151S and I157T mutations at exon 4 favored the change of non-polar residue to polar residue (146), and from molecular mechanism simulation A151S and I157T tend to hinder tetramerization. A151S (non- conservative mutation) was mapped to the same cistron, whereas the I157T was found to be located 6 citron apart (146).

5.1.1. A151S

Alanine (wild type) and serine (mutant) differ in size; the mutation will cause an empty space in the core of the protein. The mutant residue is more hydrophobic than the wild-type residue. The hydrophobicity of the wild-type and mutant residue differs. The mutation will cause loss of hydrogen bond and, as a result, it may disturb the protein folding.

Table 1. Summary of experimentally proved SAPs in ALDH gene superfamily

Gene	Variants/Reference seq ID	Mutation position	Disease	Amino acid length	References
ALDH1A2	rs115875978 rs115913750	A151S I157T	Congenital Heart disease	518	146
ALDH1A3	VAR_069322/rs58142816	R89C	Recessive Anophthalmia and Microphthalmia	512	151
ALDH1B1	VAR_002257/rs2228093	A86V	Alcohol induced Hypersentivity	517	152,153
ALDH1L1	VAR_052297/rs1127717	D793G	Hodgkin's Lymphoma	902	162
ALDH2	VAR_002248/rs671	E504K	Esophageal cancer Diabetic Cardiomyopathy Cardiac dysfunction Alzheimer's disease Colorectal cancer	517	164,165,166 167,168,169 172 179 182,183
ALDH3A2	VAR_017519/rs72547569	K266N	Sjogren-Larsson syndrome	485	187
ALDH4A1	VAR_002260/rs137852937	S352L	Type II Hyperprolinemia	563	190
ALDH5A1	VAR_026204	K301E	Semi-aldehyde dehydrogenase deficiency – Gamma- hydroxyl butyric aciduria	535	195
ALDH6A1	rs183066442 VAR_010244/rs72552258	R535C G466R	Dysmyelination and transient methylmalonic aciduria	535	201 80
ALDH7A1	VAR_069189 VAR_031719/rs121912707 VAR_069186/rs121912711	L455P E427Q N301I	Pyridoxine dependent Epilepsy and Folic acid responsive seizures	539	211
ALDH16A1	rs201033569	P527R	Gout and Mast syndrome	802	216
ALDH18A1	VAR_038482/ rs121434582	R84Q	Urea Cycle defects - hyperprolinemia, hypoornithinemia, hypocitrullinemia, hypoargininemia and hyperammonemia	795	217

Table 2. Interpretation of SAPs in ALDH gene superfamily by KD4v

Gene name	Reference ID	A. a position	Size	Charge	Polarity	Hydrophobicity	Disulphide bond	Accessibility	
								Mutant	Wild
ALDH1A1	rs151269980	E313K	Increase	Opposite	Unchanged	Unchanged	Unchanged	27.1.9 (Intermediate)	41.4.7 (Accessible)
	rs150537821	P76R	Increase	Increase	Increase	Unchanged	Unchanged	32.3.7 (Accessible)	46.5.3 (Accessible)
	rs143464638	R143C	Decrease	Decrease	Decrease	Increase	Unchanged	42.6.0 (Accessible)	27.6.1 (Intermediate)
	rs149236405	R143H	Decrease	Unchanged	Unchanged	Increase	Unchanged	42.6.0 (Accessible)	40.3.7 (Accessible)
	rs142280224	G379A	Unchanged	Unchanged	Unchanged	Increased	Unchanged	5.0.2 (Buried)	4.8.4 (Buried)
	rs11554424	D15Y	Increase	Decrease	Unchanged	Increase	Unchanged	40.6.3 (Accessible)	52.9.0 (Accessible)
ALDH1A2	rs145263250	F171L	Decrease	Unchanged	Unchanged	Unchanged	Unchanged	21.7.3 (Intermediate)	28.8.0 (Intermediate)
	rs139464985	A151S	Decrease	Unchanged	Unchanged	Decrease	Unchanged	26.1.4 (Intermediate)	23.0.1 (Intermediate)

(Contd...)

Table 2. Contd...

Gene name	Reference ID	A. a position	Size	Charge	Polarity	Hydrophobicity	Disulphide bond	Accessibility	
								Mutant	Wild
	rs141245344	R347H	Decrease	Unchanged	Unchanged	Increase	Unchanged	4.0.5 (Buried)	5.1.3 (Buried)
	rs115913750	I157T	Decrease	Unchanged	Decrease	Unchanged	Unchanged	NA	NA
ALDH1A3	rs199537142	R98W	Increase	Increase	Unchanged	Decrease	Unchanged	NA	NA
	rs147752643	A470T	Increase	Unchanged	Unchanged	Unchanged	Unchanged	NA	NA
	rs113661159	R89C	Increase	Unchanged	Increase	Increase	Unchanged	27.6.1 (Intermediate)	14.9.2 (Intermediate)
ALDH1B1	rs148903407	R84Q	Decrease	Decrease	Unchanged	Increase	Unchanged	15.2.6 (Intermediate)	15.4.1 (Intermediate)
	rs145597002	R84W	Unchanged	Decrease	Decrease	Increase	Unchanged	15.2.6 (Intermediate)	19.8.6 (Intermediate)
	rs117820646	T325N	Increase	Unchanged	Unchanged	Decrease	Unchanged	0.3.6 (Buried)	0.4.5 (Buried)
	rs111325536	M191V	Decrease	Unchanged	Unchanged	Unchanged	Unchanged	2.6.7 (Buried)	2.5.5 (Buried)
	rs71504569	P215S	Decrease	Unchanged	Increase	Increase	Unchanged	0.0.0 (Buried)	0.0.0 (Buried)
	rs2228093	A86V	Increase	Unchanged	Decrease	Increase	Unchanged	8.3.7 (Buried)	4.6.0 (Buried)
ALDH1L1	rs150865017	P216L	Increase	Unchanged	Unchanged	Increase	Unchanged	24.7.2 (Intermediate)	27.1.3 (Intermediate)
	rs149080804	P107L	Increase	Unchanged	Unchanged	Increase	Unchanged	6.2.4 (Buried)	10.0.9 (Intermediate)
	rs142194643	P52L	Increase	Unchanged	Unchanged	Increase	Unchanged	23.6.1 (Intermediate)	30.3.9 (Accessible)
ALDH1L2	rs187752137	G496S	Increase	Unchanged	Increase	Unchanged	Unchanged	12.7.0 (Intermediate)	13.0.8 (Intermediate)
	rs149557935	P278L	Increase	Unchanged	Unchanged	Increase	Unchanged	28.5.5 (Intermediate)	17.8.1 (Intermediate)
	rs146335509	I205V	Decrease	Unchanged	Unchanged	Unchanged	Unchanged	2.8.0 (Buried)	5.8.6 (Buried)
	rs146068404	N273H	Increase	Increase	Unchanged	Increase	Unchanged	10.6.1 (Intermediate)	16.7.9 (Intermediate)
	rs144770528	G796E	Increase	Increase	Increase	Decrease	Unchanged	10.1.5 (Intermediate)	6.3.6 (Buried)
	rs143568544	D62G	Decrease	Decrease	Decrease	Increase	Unchanged	NA	NA
	rs142554052	R42C	Decrease	Decrease	Increase	Increase	Unchanged	27.7.4 (Intermediate)	25.7.2 (Intermediate)
	rs140127163	R733Q	Decrease	Decrease	Increase	Unchanged	Unchanged	3.7.3 (Buried)	4.5.2 (Buried)
	rs140102957	V643F	Increase	Unchanged	Unchanged	Unchanged	Unchanged	0.0.0 (Buried)	0.5.4 (Buried)
	rs140095224	A477T	Increase	Unchanged	Increase	Decrease	Unchanged	0.0.9 (Buried)	0.0.0 (Buried)
	rs138108238	G286D	Increase	Increase	Increase	Decrease	Unchanged	28.3.4 (Intermediate)	21.3.0 (Intermediate)
	rs138062395	T623M	Increase	Unchanged	Decrease	Increase	Unchanged	0.0.0 (Buried)	0.0.6 (Buried)
	rs115446211	Y782H	Unchanged	Increase	Unchanged	Decrease	Unchanged	1.8.8 (Buried)	1.8.5 (Buried)

(Contd...)

Table 2. Contd...

Gene name	Reference ID	A. a position	Size	Charge	Polarity	Hydrophobicity	Disulphide bond	Accessibility	
								Mutant	Wild
	rs191240032	P184L	Increase	Unchanged	Unchanged	Increase	Unchanged	8.0.8 (Buried)	7.0.4 (Buried)
	rs150202179	R103W	Unchanged	Decrease	Decrease	Increase	Unchanged	39.3.9 (Accessible)	47.6.5 (Accessible)
	rs149442966	P400S	Decrease	Unchanged	Increase	Increase	Unchanged	2.5.6 (Buried)	0.8.7 (Buried)
	rs148698157	A298D	Increase	Increase	Increase	Decrease	Unchanged	5.6.8 (Buried)	6.0.9 (Buried)
	rs147086207	D110H	Increase	Opposite	Unchanged	Increase	Unchanged	23.0.0 (Intermediate)	32.4.5 (Accessible)
	rs145077856	G287W	Increase	Unchanged	Increase	Increase	Unchanged	3.5.7 (Buried)	14.9.2 (Intermediate)
	rs144598865	E431K	Increase	Opposite	Unchanged	Unchanged	Unchanged	39.1.6 (Accessible)	47.6.1 (Accessible)
	rs142271678	D299N	Unchanged	Decrease	Unchanged	Unchanged	Unchanged	35.1.9 (Accessible)	37.5.6 (Accessible)
	rs141629803	A150T	Increase	Unchanged	Increase	Decrease	Unchanged	4.1.1 (Buried)	6.2.6 (Buried)
	rs141574314	I41V	Decrease	Unchanged	Unchanged	Unchanged	Unchanged	2.1.0 (Buried)	3.6.0 (Buried)
	rs140347209	V80A	Decrease	Unchanged	Unchanged	Unchanged	Unchanged	0.0.0 (Buried)	0.0.0 (Buried)
	rs139133423	S277G	Decrease	Unchanged	Decrease	Unchanged	Unchanged	8.5.6 (Buried)	8.9.2 (Buried)
ALDH3A1	rs200726755	D182Y	Increase	Decrease	Unchanged	Increase	Unchanged	NA	NA
	rs200539114	R426H	Decrease	Unchanged	Unchanged	Increase	Unchanged	NA	NA
	rs188828873	Q242K	Increase	Increase	Unchanged	Decrease	Unchanged	NA	NA
	rs149784212	L185P	Decrease	Unchanged	Unchanged	Decrease	Unchanged	NA	NA
	rs147790087	A59T	Increase	Unchanged	Increase	Decrease	Unchanged	0.0.0 (Buried)	0.0.0 (Buried)
	rs147298045	R10G	Decrease	Decrease	Decrease	Increase	Unchanged	32.4.9 (Accessible)	16.0.3 (Intermediate)
	rs147275012	R231C	Decrease	Decrease	Decrease	Increase	Unchanged	NA	NA
	rs146746671	P139L	Increase	Unchanged	Unchanged	Increase	Unchanged	NA	NA
	rs145697414	H183R	Increase	Unchanged	Unchanged	Decrease	Unchanged	3.2.6 (Buried)	3.0.1 (Buried)
	rs145505711	R231L	Decrease	Decrease	Decrease	Increase	Unchanged	25.6.9 (Intermediate)	19.6.5 (Intermediate)
	rs145465198	I342T	Decrease	Unchanged	Increase	Decrease	Unchanged	NA	NA
	rs143872646	Y282C	Decrease	Unchanged	Unchanged	Decrease	Unchanged	NA	NA
	rs140881083	R314C	Decrease	Decrease	Decrease	Increase	Unchanged	20.8.7 (Intermediate)	8.7.9 (Buried)
	rs11554978	S143I	Increase	Unchanged	Decrease	Increase	Unchanged	NA	NA
	rs11554977	S140P	Decrease	Decrease	Decrease	Increase	Unchanged	NA	NA
ALDH3A2	rs72547569	K266N	Decrease	Decrease	Unchanged	Unchanged	Unchanged	NA	NA
	rs72547564	C214Y	Increase	Unchanged	Increase	Unchanged	Unchanged	0.0.0 (Buried)	0.0.0 (Buried)
	rs148103086	C249S	Unchanged	Unchanged	Increase	Decrease	Unchanged	0.0.0 (Buried)	0.0.0 (Buried)

(Contd...)

Table 2. Contd...

Gene name	Reference ID	A. a position	Size	Charge	Polarity	Hydrophobicity	Disulphide bond	Accessibility	
								Mutant	Wild
	rs147200808	E330G	Decrease	Decrease	Decrease	Increase	Unchanged	19.2.7 (Intermediate)	10.4.2 (Intermediate)
	rs144365591	P358S	Decrease	Unchanged	Increase	Increase	Unchanged	0.9.9 (Buried)	1.2.2 (Buried)
	rs144287421	A110T	Increase	Unchanged	Increase	Decrease	Unchanged	6.1.7 (Buried)	5.6.5 (Buried)
	rs142587791	Y215D	Decrease	Increase	Unchanged	Decrease	Unchanged	0.6.8 (Buried)	4.5.9 (Buried)
	rs138110265	T316I	Increase	Unchanged	Decrease	Increase	Unchanged	0.0.0 (Buried)	0.1.4 (Buried)
	rs112346474	G334R	Increase	Increase	Increase	Decrease	Unchanged	0.0.0 (Buried)	0.0.0 (Buried)
	rs72547576	F419S	Decrease	Unchanged	Increase	Decrease	Unchanged	0.2.1 (Buried)	3.7.3 (Buried)
	rs72547575	N386S	Decrease	Unchanged	Unchanged	Increase	Unchanged	21.0.1 (Intermediate)	18.3.7 (Intermediate)
	rs72547574	S380N	Increase	Unchanged	Unchanged	Decrease	Unchanged	7.9.4 (Buried)	5.6.9 (Buried)
	rs72547573	S365L	Increase	Unchanged	Decrease	Increase	Unchanged	1.1.2 (Buried)	0.9.0 (Buried)
	rs72547571	P315S	Decrease	Unchanged	Increase	Increase	Unchanged	1.2.4 (Buried)	0.3.5 (Buried)
	rs72547568	D245N	Unchanged	Decrease	Unchanged	Unchanged	Unchanged	1.3.4 (Buried)	1.5.6 (Buried)
	rs72547567	C237Y	Increase	Unchanged	Increase	Unchanged	Unchanged	0.0.0 (Buried)	0.0.0 (Buried)
	rs72547566	R228C	Decrease	Decrease	Decrease	Increase	Unchanged	19.5.7 (Intermediate)	6.1.3 (Buried)
	rs72547562	T184M	Increase	Unchanged	Decrease	Increase	Unchanged	5.0.5 (Buried)	4.6.4 (Buried)
	rs72547560	P121L	Increase	Unchanged	Unchanged	Increase	Unchanged	0.2.7 (Buried)	0.1.4 (Buried)
	rs72547559	P114L	Increase	Unchanged	Unchanged	Increase	Unchanged	1.2.6 (Buried)	0.6.0 (Buried)
	rs72547556	V64D	Increase	Increase	Increase	Decrease	Unchanged	0.2.6 (Buried)	0.6.5 (Buried)
	rs72547555	L27P	Decrease	Unchanged	Unchanged	Decrease	Unchanged	0.0.0 (Buried)	2.8.3 (Buried)
ALDH3B1	rs189023710	R236C	Decrease	Decrease	Decrease	Increase	Unchanged	0.3.7 (Buried)	1.2.4 (Buried)
	rs188211650	F57L	Decrease	Unchanged	Unchanged	Unchanged	Unchanged	42.8.0 (Accessible)	42.6.9 (Accessible)
	rs18261043	G301R	Increase	Increase	Increase	Decrease	Unchanged	18.4.3 (Intermediate)	11.2.3 (Intermediate)
	rs115555799	T319P	Unchanged	Unchanged	Decrease	Decrease	Unchanged	0.5.2 (Buried)	0.5.3 (Buried)
ALDH3B2	rs188008035	R221H	Decrease	Unchanged	Unchanged	Increase	Unchanged	50.1.1 (Accessible)	45.6.4 (Accessible)
	rs180859869	G303S	Increase	Unchanged	Increase	Unchanged	Unchanged	5.5.6 (Buried)	3.5.6 (Buried)
	rs150084712	P257L	Increase	Unchanged	Unchanged	Increase	Unchanged	1.3.4 (Buried)	0.0.7 (Buried)
	rs149110344	R212W	Unchanged	Decrease	Decrease	Increase	Unchanged	16.2.8 (Intermediate)	36.7.5 (Accessible)
	rs147492892	L96P	Decrease	Unchanged	Unchanged	Decrease	Unchanged	5.0.6 (Buried)	13.8.1 (Intermediate)
	rs144853999	I235L	Unchanged	Unchanged	Unchanged	Unchanged	Unchanged	2.4.7 (Buried)	1.3.4 (Buried)

(Contd...)

Table 2. Contd...

Gene name	Reference ID	A. a position	Size	Charge	Polarity	Hydrophobicity	Disulphide bond	Accessibility	
								Mutant	Wild
	rs144739802	R221C	Decrease	Decrease	Decrease	Increase	Unchanged	50.1.1 (Accessible)	29.2.2 (Intermediate)
	rs144564033	A148T	Increase	Unchanged	Increase	Decrease	Unchanged	0.0.9 (Buried)	3.3.6 (Buried)
	rs142113776	A236T	Increase	Unchanged	Increase	Decrease	Unchanged	3.8.3 (Buried)	5.2.4 (Buried)
	rs140248913	P36L	Increase	Unchanged	Unchanged	Increase	Unchanged	11.0.2 (Intermediate)	14.5.3 (Intermediate)
	rs13967003	G48R	Increase	Increase	Increase	Decrease	Unchanged	6.0.5 (Buried)	39.8.1 (Accessible)
	rs137924137	C155R	Increase	Increase	Increase	Decrease	Unchanged	1.2.3 (Buried)	0.4.7 (Buried)
	rs113681988	G226S	Increase	Unchanged	Increase	Unchanged	Unchanged	19.7.4 (Intermediate)	18.9.4 (Intermediate)
ALDH4A1	rs137852937	S352L	Increase	Increase	Decrease	Increase	Unchanged	NA	NA
ALDH5A1	rs148733464	T202I	Increase	Unchanged	Decrease	Increase	Unchanged	NA	NA
	rs142482046	D474A	Decrease	Decrease	Decrease	Increase	Unchanged	20.9.9 (Intermediate)	13.7.6 (Intermediate)
	rs139719918	A404G	Unchanged	Unchanged	Unchanged	Decrease	Unchanged	1.8.8 (Buried)	1.8.4 (Buried)
	rs139486423	A161T	Increase	Unchanged	Increase	Decrease	Unchanged	0.0.9 (Buried)	0.0.0 (Buried)
	rs115784602	V321M	Increase	Unchanged	Unchanged	Unchanged	Unchanged	16.6.4 (Intermediate)	18.7.6 (Intermediate)
ALDH6A1	rs72552258	G446R	Unchanged	Decrease	Decrease	Increase	Unchanged	28.5.7 (Intermediate)	31.2.8 (Accessible)
	rs201935817	R140G	Decrease	Unchanged	Unchanged	Increase	Unchanged	Nil	Nil
	rs199699588	K364E	Increase	Unchanged	Unchanged	Unchanged	Unchanged	Nil	Nil
	rs147813764	G352R	Decrease	Unchanged	Unchanged	Unchanged	Unchanged	17.4.1 (Intermediate)	21.0.3 (Intermediate)
	rs147635769	R488P	Unchanged	Decrease	Decrease	Increase	Unchanged	39.3.9 (Accessible)	47.6.5 (Accessible)
ALDH7A1	rs181287032	R404C	Decrease	Decrease	Decrease	Increase	Unchanged	27.0.1 (Intermediate)	19.3.5 (Intermediate)
	rs151107837	P431L	Increase	Unchanged	Unchanged	Increase	Unchanged	0.7.9 (Buried)	0.6.9 (Buried)
	rs14052962	S317L	Increase	Unchanged	Decrease	Increase	Unchanged	3.9.6 (Buried)	2.5.0 (Buried)
	rs145329188	L121S	Decrease	Unchanged	Increase	Decrease	Unchanged	0.0.0 (Buried)	2.1.3 (Buried)
	rs140701322	P244T	Unchanged	Unchanged	Increase	Increase	Unchanged	18.7.3 (Intermediate)	18.8.0 (Intermediate)
	rs202084501	L455P	Decrease	Unchanged	Unchanged	Decrease	Unchanged	0.5.8 (Buried)	0.7.8 (Intermediate)
	rs121912707	E427Q	Decrease	Decrease	Unchanged	Increase	Unchanged	15.8.4 (Intermediate)	14.7.5 (Intermediate)
	rs121912711	N301I	Unchanged	Unchanged	Decrease	Increase	Unchanged	0.2.2 (Buried)	0.1.4 (Buried)

(Contd...)

Table 2. Contd...

Gene name	Reference ID	A. a position	Size	Charge	Polarity	Hydrophobicity	Disulphide bond	Accessibility	
								Mutant	Wild
ALDH18A1	rs121434583	H784Y	Unchanged	Decrease	Unchanged	Increase	Unchanged	22.8.8 (Intermediate)	23.7.5 (Intermediate)
	rs121434582	R84Q	Decrease	Decrease	Unchanged	Increase	Unchanged	Nil	Nil
	rs150980819	Y668C	Decrease	Unchanged	Decrease	Unchanged	Unchanged	29.0.3 (Intermediate)	28.0.1 (Intermediate)
	rs150429857	A461T	Increase	Unchanged	Increase	Decrease	Unchanged	1.2.4 (Buried)	2.6.0 (Buried)
	rs148127786	T520S	Decrease	Unchanged	Unchanged	Unchanged	Unchanged	0.0.0 (Buried)	0.8.6 (Buried)

Table 3. Effect of wild type and mutant type amino acids by HOPE analysis

Gene	Amino acid position	Properties	Wild type	Mutant type	Effects of mutations
ALDH3A2	K266N	Size	Smaller	Bigger	The wild-type residue was buried in the core of the protein. The mutant residue is bigger and probably will not fit
		Charge	NA	NA	NA
		Hydrophobicity	More hydrophobic	Less hydrophobic	The hydrophobicity of the wild-type and mutant residue differs. Due to mutation it may cause loss of hydrophobic interactions in the core of the protein
		Structure	NA	NA	NA
		Domain			The residue is buried in the core of a domain. Mutation might disturb the core structure of this domain
		Conservation	NA	NA	NA
		Salt bridge formation	NA	NA	NA
		Hydrogen bond formation	NA	NA	NA
		Contacts		Is not in direct contact with a ligand.	The mutation can affect the local stability which in turn could affect the ligand-contacts made by one of the neighboring residues.
		Variants			VAR_002250. The variant is annotated with severity: “DISEASE-”Sjogren-Larsson syndrome”
	C214Y	Size	Smaller	Bigger	The wild-type residue was buried in the core of the protein. The mutant residue is bigger and probably will not fit
		Charge	NA	NA	NA

(Contd...)

Table 3. Contd...

Gene	Amino acid position	Properties	Wild type	Mutant type	Effects of mutations
		Hydrophobicity	More hydrophobic	Less hydrophobic	The hydrophobicity of the wild-type and mutant residue differs. Due to mutation it may cause loss of hydrophobic interactions in the core of the protein
		Structure	NA	NA	NA
		Domain			The residue is buried in the core of a domain. Mutation might disturb the core structure of this domain
		Conservation	NA	NA	NA
		Salt bridge formation	NA	NA	NA
		Hydrogen bond formation	NA	NA	NA
		Contacts		Is not in direct contact with a ligand.	The mutation can affect the local stability which in turn could affect the ligand-contacts made by one of the neighboring residues
		Variants			VAR_002250. The variant is annotated with severity: “DISEASE-”Sjogren-Larsson syndrome”
	P114L	Size	Bigger	Smaller	Due to differences in size the mutation leads to loss of external interactions
		Charge	Positive	Neutral	The difference in charge will disturb the ionic interaction made by the original, wild-type residue and may cause loss of interactions with other molecules
		Hydrophobicity	NA	NA	NA
		Structure	In the 3D-structure the wild-type residue is located in an α-helix		The mutation converts the wild-type residue in a residue that does not prefer α-helices as secondary structure
		Domain		Asparagine is located on the surface of a domain with unknown function	Due to this mutation contact with other molecules or a domain might be affected
		Conservation	NA	NA	NA
		Salt bridge formation	Lysine salt bridge with the Glutamic acid on position 271, with the Glutamic acid on position 263		
		Hydrogen bond formation	Lysine forms a hydrogen bond with the Glutamic acid on position 271		The difference in size between Lysine and Asparagine disturbs the hydrogen bond formation.

(Contd...)

Table 3. Contd...

Gene	Amino acid position	Properties	Wild type	Mutant type	Effects of mutations
		Variants			The variant: VAR_017519. The variant is annotated with severity: “DISEASE”. (“Sjogren-Larsson syndrome”
ALDH4A1	S352L	Size	Smaller	Bigger	Leucine is bigger than the Serine Interactions of Serine with neighboring residues may be disturbed
		Charge			
		Hydrophobicity	Less hydrophobic	More hydrophobic	The difference in hydrophobicity will affect hydrogen bond formation
		Structure	NA	NA	NA
		Domain		The mutated residue is located on the surface of a domain with unknown function	Leucine is located on the domain with an unknown function; this property might disturb the original functional activity of the protein
		Conservation	very conserved	The mutant residue is located near a highly conserved position	Based on conservation scores this mutation is probably damaging to the protein
		Salt bridge formation	NA	NA	NA
		Hydrogen bond formation	Serine Forms a hydrogen bond with the Lysine on position 318		Hydrogen bond formation is affected
		Variants			The variant is annotated with severity: “DISEASE”-. Type II Hyperprolinemia.
ALDH7A1	E427Q	Size	NA	NA	NA
		Charge	Negative	Neutral	The charge of the wild-type residue is lost by this mutation. This can cause loss of interactions with other molecules. The difference in charge will disturb the ionic interaction made by the original, wild-type residue
		Hydrophobicity	NA	NA	NA
		Structure	NA	NA	NA
		Domain		The mutated residue is located on the surface of a domain with unknown function	The residue was not found to be in contact with other domains of which the function is known within the used structure. However, contact with other molecules or domains are still possible and might be affected by this mutation
		Conservation	NA	NA	NA

(Contd...)

Table 3. Contd...

Gene	Amino acid position	Properties	Wild type	Mutant type	Effects of mutations
		Salt bridge formation	The wild type residue forms a salt bridge with the Arginine on position 334		
		Hydrogen bond formation	NA	NA	NA
		Contacts			In the 3D-structure the wild-type residue has interactions with a ligand annotated as: NAI. The difference in properties between wild-type and mutation can easily cause loss of interactions with the ligand. Because ligand binding is often important for the protein's function, this function might be disturbed by this mutation. According to the PISA-database, the mutated residue is involved in a multimer contact. The PISA-database contains protein assemblies that are highly likely to be biologically relevant. This residue is involved in an ionic interaction, which might be important for multimerisation. This interaction is lost by this mutation
		Variants			The variant is annotated with severity: "DISEASE". "Pyridoxine-dependent epilepsy (PDE) (MIM: 266100)"
	N301I	Size	Bigger	Smaller	The mutant is smaller than the wild type, hence by mutation will cause an empty space in the core of the protein
		Charge	NA	NA	NA
		Hydrophobicity	NA	NA	The hydrophobicity of wild and mutant differs
		Structure	NA	NA	NA
		Domain			The residue is buried in the core of a domain. The differences between the wild-type and mutant residue might disturb the core structure of this domain
		Conservation	NA	NA	NA
		Salt bridge formation	NA	NA	NA
		Hydrogen bond formation	Asparagine forms a hydrogen bond with the Cysteine on position 330, with Serine on position 456		The mutation will cause loss of hydrogen bond formation, which leads or disrupts the correct protein folding

(Contd...)

Table 3. Contd...

Gene	Amino acid position	Properties	Wild type	Mutant type	Effects of mutations
		Variants			The variant is annotated with severity: "DISEASE". "Pyridoxine-dependent epilepsy (PDE) (MIM: 266100)"
NA: Not available.					

5.1.2. I157T

Loss of interactions with ligands is mainly due to the differences in properties between the wild type (isoleucine) and the mutant type (threonine). From PISA-database, it was predicted that threonine is involved in a multimer contact. The PISA-database contains protein assemblies that are highly likely to be biologically relevant. The mutation introduces a smaller residue at this position, whereby the new residue might be too small to make multimer contacts. A mutation (I157T) converts the alanine in a residue which does not prefer α helices as secondary structure, more likely the contact with other domains or molecules is also affected.

5.2. Recessive Anophthalmia and Microphthalmia

Monogenic Anophthalmia / Microphthalmia (OMIM *600463) is inherited as an autosomal-dominant (AD), autosomal recessive (AR), or X-linked traits (147). True Anophthalmia are the absence of ocular tissue in the orbit, whereas Microphthalmia refers to tiny or no visible remnants of globe in the orbit. Both the anomalies are associated with systemic anomalies in at least 50% of individuals (148, 149, 150). Lucas *et al. (*2013) confirmed the mutation c.265C>T (Arg89Cys) by means of Sanger sequencing, and familial analysis was confirmed by the biparental transmission (151).

5.2.1. R89C

The wild type residue arginine is amphipathic in nature and (positively charged) is bigger than the mutant type residue cysteine (neutral). The size difference between wild-type and mutant residue makes that the new residue is not in the correct position to make the same hydrogen bond as the original wild-type residue did. This will cause a possible loss of external interactions. The charge, size and polarity decreased. This mutation lost the charge of the wild-type residue. This can cause loss of interactions with other molecules. Due to the conversion of arginine to cysteine, the charge differs and, as a result, it can disturb the ionic interactions which are needed for multimerization. The side chains of arginine are buried. Cysteine is frequently involved in disulphide bond whereby a pair of cysteine forms covalent bonds. The hydrophobicity increased, and their disulphide bond remained unchanged. The difference in hydrophobicity can affect the hydrogen bond formation. Any hydrogen bonds that could be made by the wild type residue to other monomers will be lost and whereby affect the multimeric contacts. The accessibility of the wild type residue was 27.6.1 (intermediate), and the mutant residue was 14.9.2 (intermediate). The mutant residue is more hydrophobic than the wild-type residue. The wild type residue forms a salt bridge with the Glutamic acid on position 480, 959, and 1449.

5.3. Alcohol-induced Hypersentivity

Husemoen *et al. (*2008) and Linneberg *et al.* (2010) proposed that based on large population studies ALDH1B1 variant (A86V) was associated with Alcohol-induced hypersensitivity (OMIM *100670), and based on these findings it was strongly proved that ALDH1B1 may have a role in the ethanol detoxification, and therefore any alterations in ALDH1B1 may contribute alcohol- related diseases and alcohol induced hypersensitivity (152, 153).

5.3.1. A86V

Alanine the dullest amino acid is non-polar and non-hydrophobic, whereas valine is more hydrophobic and prefers to be buried in the protein hydrophobic core. Valine is mainly involved in the binding and recognition of hydrophobic ligands. The charge, polarity, hydrophobicity and disulphide bond remained unchanged. The mutant residue (valine) is smaller than the wild type residue (alanine); the difference in size possibly can cause loss of interactions. The new residue might be too small to make multimer contacts. The accessibility of both the mutant and wild type was buried (3.4.2), (3.7.5). Valine was found to be located on the surface of the domain (unknown function), possible contact with other domains or molecules might be affected by this mutation.

5.4. Non - Hodgkin's Lymphoma

Non-Hodgkin lymphoma (OMIM*600249) develops mainly due to chromosomal alterations and altered methylation of oncogenes or tumor suppressor genes (154, 155, 156). Genetic variability of ALDH1L1/FTHFD in DNA synthesis and DNA methylation could influence susceptibility to NHL. The one-carbon metabolic pathway regulates nucleotide synthesis and DNA methylation *via* a complex process involving at least 30 different enzymes (157). To study the relationship between the polymorphisms in one-carbon metabolism genes and NHL risks, many studies have been done (158,159,160,161). The ALDH1L1 Ex 21 + 31A >G (D793G) variant was associated with NHL and located on the catalytic terminal domain (162). ALDH1L1 catalyzes the conversion of 10- formyl-tetrahydrofolate to tetrahydrofolate. The down-regulation of ALDH1L1 in tumors may increase the proliferation of tumor cells (163). Increased ALDH1L1 expression may impair cell growth by depleting the supply of 10-formyl-THF for purine biosynthesis, while low ALDH1L1 expression facilitates the incorporation of one-carbon units into purine (163).

5.4.1. D793G

The mutant residue (glycine) is smaller than the wild-type residue (aspartic acid). The residue is located on the surface of the protein; mutation of this residue can disturb interactions with other molecules or other parts of the protein. The torsion angles for this residue are unusual. The charge and polarity are decreased; the hydrophobicity increased, but the disulphide bond remains unchanged. There is a difference in charge between the wild-type and mutant amino acid. The mutation introduces a charge at this position; this can cause repulsion between the mutant residue and neighboring residues. Mutation to glycine can induce the function of flexibility. Due to the conversion of aspartic acid to glycine the possibilities of contacts with other domains or molecules can be affected.

5.5. Alcohol-induced disease

Esophageal cancer (OMIM *100650) – Strong convincing evidence have provided through epidemiological studies that alcohol drinking is a strong risk factor for esophageal cancer development (164, 165, 166). Despite alcohol being a carcinogen; acetaldehyde (a primary metabolite) proved to be a better carcinogen in the experimental model.

Diabetic Cardiomyopathy (OMIM *100650) - The prevalence of diabetes and associated heart diseases has been found to be steadily increasing (particularly) in Asian countries. Approximately 50% of populations found to possess one copy of the mutant *ALDH2* gene (167, 168, and 169). Evidences have revealed that ALDH2 polymorphism was closely associated with an increased risk of diabetes (170), followed by experimental findings showed that reduced expression and activities of ALDH2 are associated with oxidative stress and cardiac dysfunction in diabetes (171).

Cardiac dysfunction (OMIM *100650) - Deficiency of ALDH2 was found to aggravate ER (Endoplasmic reticulum) stress- induced cardiac dysfunction (172), which is possible through NADPH – mediated cell death (173). Bing fang *et al. (*2013) (174) and his co- workers from their research showed that ALDH2 has an essential role in the regulation of cardiac homeostasis in diabetes, ischemia-reperfusion injury and alcoholism (175, 176,177).

Alcohol dependence (OMIM*100650) is a complex disorder related to alcohol use. Wolff (1972) investigated the effectiveness of inactive ALDH2 and observed the racial differences in alcoholic responses (alcoholic induced facial flushing) (178). The polymorphism E504K was confined to North-East Asian descents.

Alzheimer's disease (OMIM *100650) – A neurodegenerative disease is the serious cause of Dementia. It was postulated that accumulation of acetaldehyde may have an important role in the Alzheimer's disease. The metabolite acetaldehyde acts as a neurotoxic (179).

Colorectal cancer (OMIM *100650) occurs mainly due to the impact of accumulated alcohol on the colon. One hypothesis involves the interference with one-carbon metabolism by alcohol or its metabolite, acetaldehyde. Folate the main micronutrient in the one-carbon metabolism, are considered as the potential protective factors for CRC. According to Giovannucci (2004) many studies have suggested the clinical interactions between folate and alcohol consumption (180). Secondly acetaldehyde acts as an inducer for colorectal cancer development (both direct and indirect genotoxic effect) (181). Regarding ALDH2 Glu487Lys polymorphism, the 487Lys allele encodes a catalytically inactive subunit (182). It is indicated that individuals with the ALDH2 Glu/Lys genotype possess 6.2.5% normal ALDH2 487Glu protein; whereby indicating a dominant effect of ALDH2 487Lys (183), the genetic polymorphisms

which modify the drinking habit (183, 184, 185) are expected to affect CRC risk.

5.5.1. E504K

Lysine (mutant) is bigger than glutamic acid (wild type). Glutamic acid was found to be located on the surface of ALDH2; mutation or conversion of glutamic acid to lysine can disturb the interactions made with other domains or molecules or other parts of the molecules. This mutation introduces an opposite charge at this position; this might disrupt the contacts with other molecules. Lysine is located on the domain, which is essential for the central activity of the protein. Glutamic acid is not conserved at this position. Quite often a new residue was observed in the homologous sequences, which means that other homologous proteins exist with another residue type than with the glutamic acid in the protein sequence. From the conservation score information's, it was depicted that mutation does not damage or affect the protein functional activity. The accessibility of the wild type was 23.3.1 (intermediate), and the mutant type was 31.7.1 (accessible).

5.6. Sjögren-Larsson syndrome

Sjögren–Larsson syndrome (SLS) (OMIM*609523) is a rare autosomal recessive disorder and an inherited neurocutaneous disorder characterized by mental retardation, spastic diplegia or tetraplegia, and congenital ichthyosis (186). The less common accompanied features are retinal changes, short stature, kyphoscoliosis, pre-term birth, photophobia, reduction of visual acuity, seizures, and delayed speech (186). SLS was mainly due to mutation in the gene ALDH3A2 located (187). Alternative splicing of ALDH3A2 results in a minor transcript which includes an additional sequence (exon 90) which in turn gives rise to a larger protein of 508-amino acids with a variant on the carboxy-terminal sequence (188). The carboxy- terminal domain of ALDH3A2 is necessary for anchoring the enzyme (ALDH3A2) to the microsomal membrane (189). According to William *et al.* (1999) (187), 22-missense mutations were detected. The amino acid substitutions were found to be scattered throughout the protein, but none was found to be involved in the amino- or carboxy-terminal amino acids. 5 of the observed missense mutations were transitions that occurred at CpG dinucleotides, whereas 2 different mutations (551CrG and 551CrT) involved the same nucleotide and predicted to cause substitution of Thr184 by Arg or Met, respectively (187). To determine whether these missense mutations were destructive to enzyme catalytic activity, chinese hamster ovary cells (genetically deficient in ALDH3A2 activity) were used to express the 20 different mutations, among them 19 mutant proteins had no or little detectable ALDH3A2 activity, but surprisingly the protein encoded by 798GrC (K266N) possessed considerable residual activity (55% of normal) (187).

5.6.1. K266N

The wild-type (lysine) and mutant (asparagine) amino acids differ in size. This mutation can cause a possible loss of external interactions. Lysine is positively charged, whereas asparagine is neutral. This mutation NIL the charge of the wild-type residue, and its ionic interaction is affected or disturbed, this can cause loss of interactions with other molecules. Lysine (wild type) forms a salt bridge with the glutamic acid on position 263 and forms a hydrogen bond with the glutamic acid on position 271. In the 3D-structure, lysine is located in α-helix; this mutation converts the lysine to asparagine that does not prefer α-helices as secondary structure, and was found to be located on the surface of the domain of unknown function. Contact with other domains or molecules are still possible and can be affected by this mutation. The variant: Variant _ 017519 is annotated with severity: "DISEASE" "Sjögren–Larsson syndrome". The superimposed structure of K266N is shown in Figure 1a.

5.6.2. C214Y

Tyrosine (mutant) is bigger than cysteine (wild type). Cysteine was found to be buried in the core of the protein. Tyrosine is bigger and probably will not fit. Cysteine is more hydrophobic than tyrosine being less hydrophobic. This mutation will cause loss of hydrophobic interactions in the core of the protein. Due to differences in cysteine and tyrosine the core structure of the domain is disturbed. Though tyrosine does not have any direct contact with ligands, this mutation can affect the local stability of the protein, and, as a result, it can alter the ligand contact in the neighboring residues. Variant VAR_017519 is annotated with severity: "DISEASE-""Sjögren–Larsson syndrome"." The superimposed structure of C214Y is shown in Figure 1b.

5.6.3. P114L

Leucine is bigger than proline. Proline was buried in the core of the protein, and lysine is bigger and probably will not fit. Proline is very rigid and induces a special backbone conformation. The mutation can disturb this special conformation. The differences between the wild-type and mutant

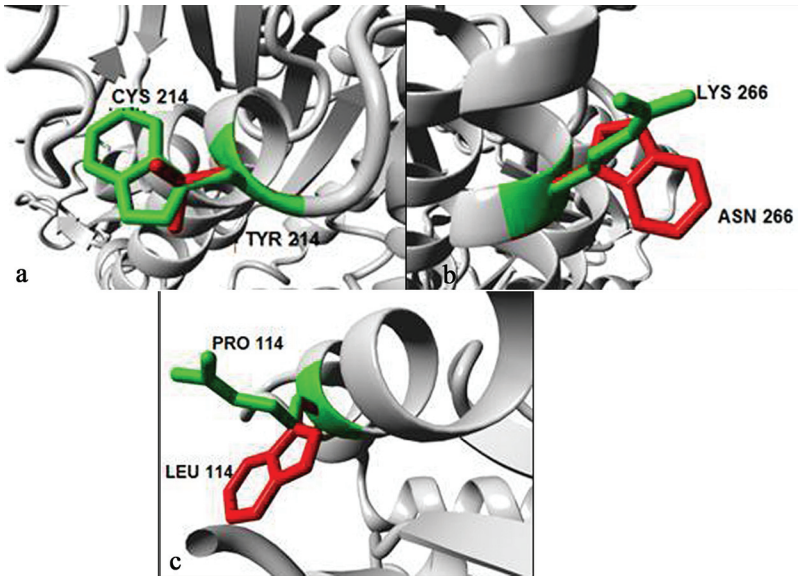

Figure 1. Close up view of mutations at position 266, 214 and 114 in ALDH3A2. The superimposed structures are obtained from HOPE server. The main protein is in grey color orientation, while the wild type and the mutant residue are shown in green and red color. a) C214Y, b) K266N, c) P114L.

residues might disturb the core structure of the domain. Proline forms a salt bridge with the leucine on position 115 and forms a hydrogen bond with the phenylalanine on position 152. The variant: VAR_017512 is annotated with severity: "DISEASE" Sjogren-Larsson syndrome." The superimposed structure of P114L is shown in Figure. 1c.

5.7. Type II Hyperprolinemia

Type II Hyperprolinemia (OMIM*606811) is an inherited autosomal recessive disorder characterized by a P5CDH deficiency, associated with elevated levels of P5C and proline in plasma, urine and cerebrospinal fluid (190). HPII is causally linked to neurologic manifestations which are associated with an increased incidence of seizures and possibly mental retardation, and schizophrenia (191, 192). Elevation of P5C results in the conversion of P5C to glutamate increased level of proline reflects the conversion of some of the accumulated P5C to proline by P5C reductase (192, 193). Human genetics studies have provided insight into the molecular basis of the Type II hyperprolinemia (194). Based on the studies several nsSNPs (non-synonymous single-nucleotide polymorphisms) in ALDH4A1 are associated with Type II hyperprolinemia, among them the known pathological SNP includes S353L, the missense mutation of Ser 353 to Leu, a frame shift mutations that can cause premature termination of translation (190).

5.7.1. S352L

The mutant residue (leucine) is bigger than the wild-type (serine). The residue is located on the surface of the protein; mutation of this residue can disturb the interactions with other molecules of the protein. Serine is less hydrophobic, whereas leucine is more hydrophobic. The difference in hydrophobicity will affect hydrogen bond formation. Leucine was not found to be in contact with other domains of which the function is known. However, contact with other molecules or domain is still possible and might be affected by this mutation. The polarity decreased, and the hydrophobicity increased. Based on the conservation scores this mutation tends to damage the protein functionally activity. Serine forms a hydrogen bond with lysine on position 318. However, due to mutation the hydrogen bond formed by serine is lost and can affect the multimeric contacts. The multimeric interactions are disturbed when this mutation replaces or converts serine to leucine (bigger residue). The variant is annotated with severity: "DISEASE"-. Type II Hyperprolinemia. Mutagenesis experiments have been performed on this position. Mutation of the wild-type residue into alanine has the following effect: "Reduced affinity for NAD and no effect on the enzyme activity." The superimposed structure is shown in Figure 2a.

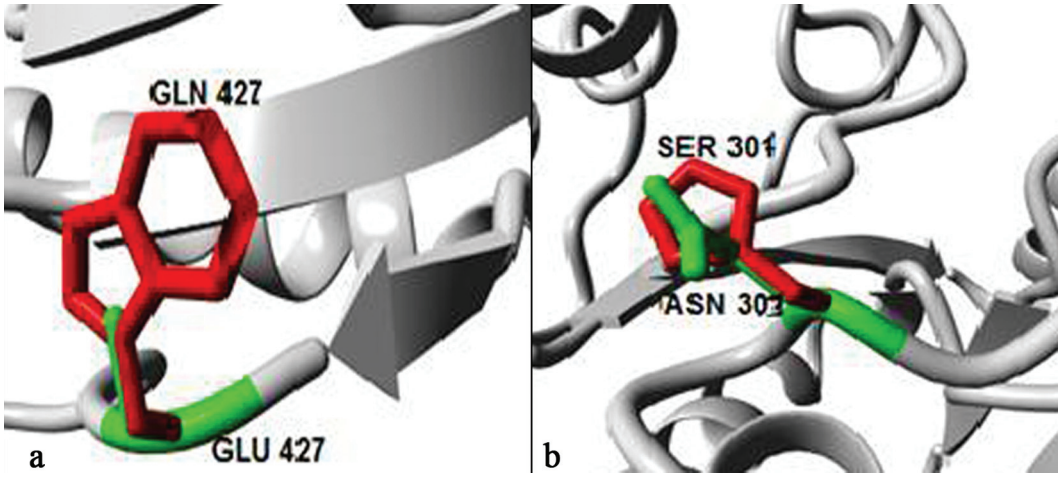

Figure 2. Close up view of mutation at position 352 in ALDH4A1. The superimposed structures are obtained from HOPE server The main protein is in grey color orientation, while the wild type and the mutant residue are shown in green and red color of S352L.

5.8. Semi-aldehyde dehydrogenase deficiency – Gamma- hydroxyl butyric aciduria

SSADH deficiency (g-hydroxybutyric aciduria) (OMIM*610045) is an autosomal recessive inherited neurometabolic disorder of GABA metabolism. It impairs the major oxidative conversion of succinic semialdehyde (SSA) to Succinic acid. The genetic block leads to accumulation of SSA, which is converted to GHB. The genetic basis resides in the SSADH (ALDH5A1) gene which maps to chromosome 6p22. SSADH deficiency is a disorder that manifests predominantly with neurological findings, but has considerable phenotypic heterogeneity (195, 196, and 197). Jakobs *et al.* (1981) identified and described gamma hydroxy– butyric acid (GHB) in their patient (first), and they were identified with striking accumulation of GHB in their body (body fluid) (198). The clinical picture encompasses a wide spectrum of neurological manifestations, and psychiatric dysfunction, the patients usually present with mild to severe developmental delay, predominantly involving expressive language, hypotonia, truncate or appendicular ataxia, and hyporeflexia and patients often develop neuropsychiatric symptoms such as inattention, anxiety, hyperkinesis, sleep disturbances and excessive daytime somnolence (199, 200). The variant responsible for SSADH was proved to be K301E (195).

5.8.1. K301E

Glutamic acid (mutant) is smaller than Lysine (wild-type residue). The residue is located on the surface of the protein; mutation of this residue can disturb interactions with other molecules or other parts of the protein. There is a difference in charge between the lysine and glutamic acid; lysine is positively charged whereas glutamic acid is negatively charged. The mutation introduces the opposite charge at this position. This can disrupt contacts with other molecules. Lysine was located on the surface of the domain and was found to be non- conserved; a new residue (glutamic acid) was observed in this position, but was found in different or other homologous sequences, which means that another homologous protein exists with another residue in the protein sequence. Based on the scores obtained from the conservation information the mutation does not affect the protein.

5.9. Dysmyelination, transient methyl malonic aciduria and 3- hydroxy isobutyric acid aciduria

Methylmalonate Semialdehyde dehydrogenase (MMSDH) is involved in the catabolic breakdown of both thymine and valine, which leads to dysmyelination and transient methyl malonic aciduria and 3- hydroxy isobutyric acid aciduria (OMIM*603178) (201, 202, 203, 204). Metabolism of thymine generates (R)-aminoisobutyric acid (AIBA), which deaminates to the (R)-methyl malonic semialdehyde, whereby valine produces

the intermediate (S)-3-hydroxyisobutyric acid (HIBA), which is oxidized to (S)-methyl malonic semialdehyde (MMSA) by 3-hydroxyisobutyrate dehydrogenase. The above mentioned (201) enantiomers of MMSA are substrates for MMSDH, which catalyzes the oxidative decarboxylation to propionyl-CoA. Pollitt *et al. (*1985) described an asymptomatic child ascertained due to the increased level of methionine in newborn screening, and was found to possess 3-hydroxyisobutyric aciduria, and ultimately possessed c.1336G > A, a homozygous missense mutation (205). Julian *et al.* (2013) reported in their study, a child with development delay followed by abnormal myelination on brain, transient/variable elevations in lactate, methyl malonic acid, 3-hydroxyisobutyric and 3-aminoisobutyric acids (201). Heterozygous mutations were then identified by means of exon sequencing. Using Sanger sequencing (within exon 6 c.514 T > C; p.T172H, and exon 12 c.1603C > T; p.R535C) the results were confirmed (201). Chambliss *et al. (*2000) in their study reported 3 patients with MMSDH, using molecular analysis they identified that the first patients possessed alterations in the coding region of MMSDH and revealed homozygosity for transversion of 1336G >A. The following transversion leads to the substitution of Glycine to Arginine at position 446 (G466R) (80).

5.9.1. R535C

The wild type residue arginine (positively charged) is bigger than the mutant type residue cysteine (neutral). The size difference between arginine and cysteine makes that the cysteine is not in the correct position to make the same hydrogen bond as the original wild-type residue did. This can cause a possible loss of external interactions. The charge, size and polarity decreased. The charge of the wild-type residue (arginine) was lost by this mutation and can disturb the ionic interaction made by the original wild-type residue. Arginine is amphipathic in nature and their side chains are buried. Cysteine is frequently involved in disulphide bond formation, whereby a pair of cysteine forms covalent bonds. The hydrophobicity increased, and their disulphide bond remained unchanged. Cysteine is more hydrophobic than arginine; the difference in hydrophobicity will affect the hydrogen bond formation. The accessibility of the mutant residue was 42.6.0 (Accessible), and the wild-type residue was 27.6.1 (Intermediate). According to PDB-file and the PISA-assembly, the wild type (arginine) is involved in multimer contact. This mutation introduces a small residue cysteine and is too small to make multimer contacts. Moreover, the hydrogen bond made by the arginine is lost, thereby affects the multimeric contacts. The ionic interaction which is important for multimerization is also affected by this mutation.

5.9.2. G466R

Arginine (mutant) is bigger than glycine (wild type), in this mutation the difference in size may disturb the interaction made with a metal ion “K.” Arginine is positively charged whereas glycine is neutral. Due to this mutation a positive charge is introduced near the metal ion “K” (positively charged) this may cause destabilization and repulsion of the domain or the neighboring residues. Arginine being less hydrophobic disturbs the multimeric interactions or contacts. Glycine being more flexible is important for the protein’s function, but due to this mutation it can abolish the flexibility of the protein.

5.10. Pyridoxine dependent Epilepsy and Folic acid responsive seizures

Pyridoxine-dependent epilepsy (PDE) (OMIM*107323) is an autosomal recessive disorder is characterized by recurrent seizures mainly found in prenatal, neonatal or the infantile period and death occurs quite often in untreated patients (206, 207). Affected neonates are usually presented with encephalopathy, hypothermia, poor feeding and onset of seizures (206). Linkage analysis of 5 PDE families associated with multiple (affected) members was located on the chromosome 5q31 (208). Using sequencing analysis, ALDH7A1 was found to be encoded with antiquitin and in further it revealed 7 different mutations in 13 patients but was found among 8 unrelated families (209, 210). The novel L455P a pathogenic variant in exon 15 of ALDH7A1 gene induces PDE (211). Barbara *et al. (*2007) in their cohort studies isolated a marked mutation E399G, accounting for 33% within their cohort, i, e 12 out of 36 alleles (212).

5.10.1. L455P

Leucine (wild type) is bigger than proline (mutant) and is very rigid. Due to differences in size it may cause the space in the core of the protein. Leucine was found to be buried in the core of the domain, but due to mutation proline may disturb the core structure of the domain. This mutation can abolish the flexibility of the protein.

5.10.2. E399G

Glutamic acid (wild type) is bigger and is the polar amino acid which has a carboxylic acid group and is negatively charged, whereas glycine

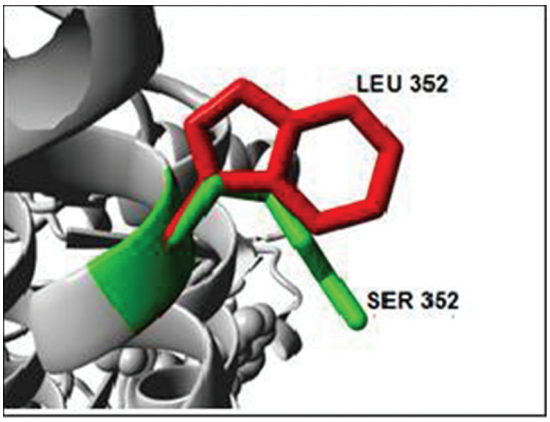

Figure 3. Close up view of mutations at position 427 and 301 in ALDH7A1. The superimposed structures are obtained from HOPE server. The main protein is in grey color orientation, while the wild type and the mutant residue are shown in green and red color. a) E427Q, b) N301I.

the smallest amino acid possess only hydrogen atom as its substitution, especially for this reason it has an increased ability to fit tight spaces of molecules. The charge differs; glutamic acid is negatively charged, and glycine is neutral, this mutation can cause loss of interactions with other molecules.

5.10.3. E427Q

Glutamic acid is negatively charged, and glutamine is neutral. In the 3D-structure, the wild-type residue has interactions with a ligand annotated as NAI. There is a difference in charge between the wild-type and mutant amino acids. This mutation can cause loss of interactions with other molecules. Glutamic acid forms a salt bridge with the arginine on position 334. The difference in properties between wild-type and mutants can easily cause loss of interactions with the ligand because ligand binding is often important for the protein's function, this function might be disturbed by this mutation. PISA database showed that glutamine is located in a multimeric contact. Due to mutation the ionic interactions of glutamic acid is lost, which is very important for multimerization. The variant is annotated with severity: "DISEASE- Pyridoxine dependent Epilepsy." The superimposed structure is shown in Figure 3a.

5.10.4. N301I

The mutant residue (isoleucine) is smaller than the wild-type residue (asparagine). This mutation can cause a space in the core of the protein and loss of hydrogen bonds can disturb the proper protein folding. The difference in hydrophobicity can affect the hydrogen bond formation. Any hydrogen bonds made by asparagine to other molecules are lost by this mutation and can affect the multimeric contacts. According to PISA-database, the mutation introduces isoleucine and is too small to make multimer contact. The variant is annotated with severity: "DISEASE – Pyridoxine dependent Epilepsy." The superimposed structure is shown in Figure 3b.

5.11. Gout and Mast syndrome

Gout (OMIM *613358) a painful inflammatory arthritis more commonly predisposed to men (213) was found to be mediated by the extracellular deposition of urate crystals in joints. The major risk factor for gout is hyperuricemia. A recent study from Iceland identified a rare missense SNP in the ALDH16A1 gene (ALDH16A1⁄2) (214). As of 2009, reports on ALDH16A1 and its interaction with Maspardin when truncated were found to be associated with Mast syndrome (OMIM*613358) (215). The ALDH16A1⁄2 allele carrying the missense SNP was located on the exon 12 (c.1427C>G) of the 16- exon transcript and exon 13 (c.1580C>G) of the 17- exon transcript, importantly this leads to substitution of proline to arginine amino acid at position 527 (P527R) (216).

5.11.1. P527R

Arginine (mutant) is bigger than proline (wild type). There is a difference in charge between proline and arginine, this mutation introduces a charge at this position; this can cause repulsion between the mutant residue and neighboring residues. The size, charge and polarity increased whereas hydrophobicity and disulphide bond remained unchanged. Proline is known to be very rigid and induces a special back-bone conformation; this mutation can disturb the special conformations. Arginine is amphipathic, and the side chain is buried. Proline is located on the surface of the protein; mutation of this residue can disturb interactions with other molecules or other parts of the protein. Salt bridge formation is less frequent in proline, whereas more frequent in arginine. The accessibility of mutant residue (arginine) was 32.3.7 (accessible), and wild type (proline) was 46.5.3 (accessible). This mutation can cause loss of hydrophobic interactions with other molecules, and arginine was found to be located in the domain which is necessary for the protein functional activity, contacts with other domains can affect the function of the protein.

5.12. Urea Cycle defects

Louise *et al. (*2008) in their study predicted a new neuro-cutaneous disorder which was described in a consanguineous Algerian family with two affected sibs presenting with an inborn error of metabolism suggestive of a urea cycle defect (OMIM*138250) (hyperprolinemia, hypoornithinemia, hypocitrullinemia, hypoargininemia and hyperammonemia (217). The connective abnormalities included the skin hyperelasticity, joint laxity in conjunction with mental retardation. From the affected individuals, the metabolic abnormalities were observed and suggested that ALDH18A1 as a candidate gene encoding D 1-the pyrroline- 5-carboxylase synthase. Subsequently both the affected sibs were homozygous for a missense mutation, and the amino acid substitution was predicted to be R84Q (217).

5.12.1. R84Q

Arginine (positively charged) is bigger than glutamate (neutral). Due to the difference in size and charge it can cause loss of external interactions and can disturb the ionic interactions made by the arginine. The residue is located on the surface of the protein; mutation of this residue can disturb the interactions with other molecules or other parts of the protein. Arginine is quite frequent in protein active binding sites. Arginine forms a salt bridge with the glutamic acid on position 85, 254, and 256.

5.12.2. H784Y

Tyrosine (mutant) is bigger than histidine (wild type). This mutation causes loss of hydrophobic interactions with other molecules. Histidine is the common amino acid in protein active or binding sites. The ease with which protons are transferred on and off of histidine makes ideal for charge relay systems. In the 3D-structure, histidine is located in the alpha helix, but due to mutation it converts the histidine in a residue which does not prefer alpha helix as secondary structure. Histidine is very much conserved and in addition other residues have been observed, but the tyrosine is located near the highly conserved region or position.

6. CONCLUSION

Aldehydes are quite ubiquitous in nature and are produced widely during biotransformation and through physiological and environmental processes, as well. ALDH superfamily exhibits catalytic and non- catalytic functions and in addition, they have a well-defined role in drug bioactivation, ROS scavenging and absorption of UV light and detoxification of aldehydes and methanol metabolites. Multiple disease states were associated with ALDH malformations or dysfunctions, which includes metabolic diseases, neurological abnormalities and cancers. On the other hand, ALDH genotypes play a major role in drug resistance whereby affects the efficacy of drug treatment for various diseases and disorder. In this study, we focused on the functional characterization and structural properties of ALDH mutations that are experimentally determined. Many computational methods have been developed successfully in designing a protein. The major problem in structural biology is the relative defective activity of the protein. Structural based calculations of protein play an important role in calculating or predicting the relative activity of the protein. The defective protein structure may lead to a complete loss of protein function and subsequently favors a diseased state. In this study, using KD4V and Project HOPE we summarized all the deleterious properties of the SAPs and our results was correlated with the experimentally proved mutations. We conclude that the protein structure prediction methods may provide insights into functional and structural characterization of the proteins. It provides a framework of high-thoroughput design of experiments such as site-directed mutagenesis, structure based design, rational drug design, analogous fold recognition, sequence-structure alignment, interactions with other macromolecules and studies based on disease-related mutations or phenotypical effects of mutations. This can eventually leads to a better understanding and makes disease diagnosis very simple and can help biomedical researchers and clinicians to conclude the pathogenecity of the mutation and can help cure patients.

7. ACKNOWLEDGEMENT

We gratefully acknowledge the management of VIT University for providing the facilities to carry out this work.

8. REFERENCES

1. G Muzio, M Maggiora, E Paiuzzi, M Oraldi, RA Canuto: Aldehyde dehydrogenases and cell proliferation. *Free Radical Biology and Medicine* 52 (4), 735-46 (2012) DOI: 10.1016/j.freeradbiomed.2011.11.033
2. V Vasiliou, A Pappa, DR Petersen: Role of aldehyde dehydrogenases in

endogenous and xenobiotic metabolism. *Chem Chemico-Biological Interactions* 129, 1 – 19 (2000)
DOI: 10.1016/S0009-2797(00)00211-8

3. NE Sladek: Human aldehyde dehydrogenases: potential pathological, pharmacological and toxicological impact. *Journal of Biochemical and Molecular Toxicology* 23 (17),7-23 (2003)
DOI: 10.1002/jbt.10057

4. B Jackson, C Brocker, C Thompson, W Black, K Vasiliou, DW Nebert, V Vasiliou: Update on the aldehyde dehydrogenase gene (ALDH) super family. *Human Genomics* 5, 283-303 (2011)
DOI: 10.1186/1479-7364-5-4-283

5. K Sydow, A Daiber, M Oelze, Z Chen, M August, M Wendt, V Ullrich, A Mulsch, E Schulz, JF Keaney, JS Stamler, T Munzel : Central role of mitochondrial aldehyde dehydrogenase and reactive oxygen species in nitroglycerin tolerance and cross-tolerance. *Journal of Clinical Investigation* 113(3),482- 489 (2004)
DOI: 10.1172/JCI19267

6. V Vasiliou, A Pappa, T Estey: Role of human aldehyde dehydrogenases in endobiotic and xenobiotic metabolism. *Drug Metabolism Reviews* 36, 279-299 (2004)
DOI: 10.1081/DMR-120034001

7. V Vasiliou, A Pappa: Polymorphisms of human aldehyde dehydrogenases. Consequences for drug metabolism and disease. *Pharmacology* 61, 192-198 (2000)
DOI: 10.1159/000028400

8. J Perozich, H Nicholas, BC Wang, R Lindahl, J Hempel: Relationships within the aldehyde dehydrogenase extended family. *Protein Science* 8, 137-146 (1999)
DOI: 10.1110/ps.8.1.137

9. V Vasiliou, A Bairoch, K Tipton, DW Nebert: Eukaryotic aldehyde dehydrogenase (ALDH) genes: Human polymorphisms and recommended nomenclature based on divergent evolution and chromosomal mapping. *Pharmacogenetics* 9, 421-434 (1999)
[DOI NOT FOUND]

10. D Zhao, P Mc Caffery, KJ Ivins, RL Neve, P Hogan, WW Chin, UC Drager: Molecular identification of a major retinoic-acid-synthesizing enzyme, a retinaldehyde-specific dehydrogenase. *European Journal of Biochemistry* 240, 15-22 (1996)
DOI: 10.1111/j.1432-1033.1996.0015h.x

11. G Duester: Families of retinoid dehydrogenases regulating vitamin A function: Production of visual pigment and retinoic acid. *European Journal of Biochemistry* 267, 4315-4324 (2000)
DOI: 10.1046/j.1432-1327.2000.01497.x

12. K Niederreither, V Fraulob, JM Garnier, P Chambon, P Dolle: Differential expression of retinoic acid-synthesizing (RALDH) enzymes during fetal development and organ differentiation in the mouse. *Mechanisms of Development* 110, 165-171 (2002)
DOI: 10.1016/S0925-4773(01)00561-5

13. D Galter, S Buervenich, A Carmine, M Anvert, L Olson : ALDH1 mRNA: presence in human dopamine neurons and decreases in substantia nigra in Parkinson's disease and in the ventral tegmental area in schizophrenia. *Neurobiology of Disease* 14, 637-647 (2003)
DOI: 10.1016/j.nbd.2003.09.001

14. S Mandel, E Grunblatt, P Riederer, N Amariglio, JJ Hirsch, G Rechavi, MB Youdim: Gene expression profiling of sporadic Parkinson's disease substantia nigra pars compacta reveals impairment of ubiquitin-proteasome subunits, SKP1A, aldehyde dehydrogenase, and chaperone HSC-70. *Annals of New York Academic of Science* 1053, 356-375 (2005)
DOI: 10.1111/j.1749-6632.2005.tb00044.x

15. SA Marchitti, C Brocker, D Stagos, V Vasiliou: Non-P450 aldehyde oxidizing enzymes: the aldehyde dehydrogenase

superfamily. *Expert Opinion on Drug Metabolism and Toxicology* 4(6), 697-720 (2008)
[DOI NOT FOUND]

16. JR Haselbeck, I Hoffmann, G Duester: Distinct functions for Aldh1 and Raldh2 in the control of ligand production for embryonic retinoid signaling pathways. *Developmental Genetics* 25(4), 353-364 (1999)
DOI: 10.1002/(SICI)1520-6408(1999)25:4<353::AID-DVG9>3.3.CO;2-7

17. D Zhao, P McCajery, KJ Ivins, RL Neve, P Hogan, WW Chin, VC Drager : Molecular identification of a major retinoic acid-synthesizing enzyme, a retinaldehyde-specific dehydrogenase. *European Journal of Biochemistry* 240, 15-22 (1996)
[DOI NOT FOUND]

18. LC Hsu, WC Chang, A Yoshida: Mouse type-2 retinaldehyde dehydrogenase (RALDH2): genomic organization, tissue-dependent expression, chromosome assignment and comparison to other types. *Biochimica et Biophysica Acta* 1492, 289-293 (2000)
DOI: 10.1016/S0167-4781(00)00108-1

19. M Maden: Heads or tails? Retinoic acid will decide. *Bioassays* 21, 809-812 (1999)
DOI: 10.1002/(SICI)1521-1878(199910)21:10<809::AID-BIES2>3.0.CO;2-0

20. LM De Luca: Retinoids and their receptors in differentiation, embryogenesis, and neoplasia. *FASEB Journal* 5, 2924-33 (1991)
[DOI NOT FOUND]

21. H Kim, J Lapointe, G Kaygusuz, DE Ong, C Li, M van de Rijn, JD Brooks, JR Pollack : The retinoic acid synthesis gene ALDH1a2 is a candidate tumor suppressor in prostate cancer. *Cancer Research* 65, 8118-24 (2005)
DOI: 10.1158/0008-5472.CAN-04-4562

22. K Niederreither, J Vermot, IL Roux, B Schuhbaur, P Chambon, P Dolle: The regional pattern of retinoic acid synthesis by RALDH2 is essential for the development of posterior pharyngeal arches and the enteric nervous system. *Development* 130, 2525-34(2003)
DOI: 10.1158/0008-5472.CAN-04-4562

23. CL Hsu, WC Chang, L Hiraoka, CL Hsieh: Molecular Cloning, Genomic Organization, and Chromosomal Localization of an Additional Human Aldehyde Dehydrogenase Gene, ALDH6. *Genomics* 24, 333-341 (1994)
DOI: 10.1006/geno.1994.1624

24. S Okamura, CC Ng, K Koyama, Y Takei, H Arakawa, M Monden, Y Nakamura: Identification of seven genes regulated by wild-type p53 in a colon cancer cell line carrying a well-controlled wild-type p53 expression system. *Oncology Research* 11, 281-5 (1999)
[DOI NOT FOUND]

25. BN Rexer, WL Zheng, DE Ong: Retinoic acid biosynthesis by normal human breast epithelium is via aldehyde dehydrogenase 6, absent in MCF-7 cells. *Cancer Research* 61, 7065-7070 (2001)
[DOI NOT FOUND]

26. J Han, L Yang, RK Puri: Analysis of target genes induced by IL-13 cytotoxin in human glioblastoma cells. *Journal of Neurooncology* 72, 35-46 (2005)
DOI: 10.1007/s11060-004-3119-7

27. MJ Stewart, K Malek, Q Xiao, KM Dipple, DW Crab: The novel aldehyde dehydrogenase gene, ALDH5, encodes an active aldehyde dehydrogenase enzyme. *Biochemical and Biophysical Research Communications* 211, 144-51 (1995)
DOI: 10.1006/bbrc.1995.1789

28. EM Algar, B Cheung, J Hayes, RS Holmes, IR Beacham: Bovine corneal aldehyde dehydrogenases: evidence for multiple gene products (ALDH3 and ALDHX). *Advances in Experimental Medicine and Biology* 328, 153-157 (1993)
DOI: 10.1007/978-1-4615-2904-0_17

29. SA Krupenko, NV Oleinik: 10-formyltetrahydrofolate dehydrogenase,

one of the major folate enzymes, is down-regulated in tumor tissues and possesses suppressor effects on cancer cells. *Cell Growth & Differentiation* 13, 227-36 (2007) [DOI NOT FOUND]

30. SA Krupenko, DA Horstman, C Wagner, RJ Cook: Baculovirus expression and purification of rat 10-formyltetrahydrofolate dehydrogenase. *Protein Expression and Purification* 6, 457-464 (1995) DOI: 10.1006/prep.1995.1061
31. C Kutzbach, EL Stokstad: Partial purification of a 10-formyl-tetrahydrofolate: NADP oxidoreductase from mammalian liver. *Biochemical and Biophysical Research Communication* 30 (2), 111-7 (1968) DOI: 10.1016/0006-291X(68)90456-7
32. C Kutzbach, EL Stokstad: Mammalian methylene tetrahydrofolate reductase. Partial purification, properties, and inhibition by S-adenosylmethionine. *Biochimica et Biophysica Acta* 250, 454-477 (1971) DOI: 10.1016/0005-2744(71)90247-6
33. MK Martinasevic, MD Green, J Baron, TR Tephly: Folate and 10-formyltetrahydrofolate dehydrogenase in human and rat retina: relation to methanol toxicity. *Toxicology and Applied Pharmacology* 141, 373-381 (1996) DOI: 10.1006/taap.1996.0302
34. E Ackerstaff, B Gim, D Artemov, ZM Bhujwalla: Anti-inflammatory agent indomethacin reduce invasion and alters metabolism in a human breast cancer cell line. *Neoplasia* 9, 222-235 (2007) DOI: 10.1593/neo.06673
35. HW Goedde, DP Agarwal: Pharmacogenetics of aldehyde dehydrogenase (ALDH). *Pharmacology & Therapeutics* 45, 345-371 (1990) DOI: 10.1016/0163-7258(90)90071-9
36. Y Li, D Zhang, W Jin, C Shao, P Yan, C Xu, H Sheng, Y Liu, J Yu, Y Xie, Y Zhao, D Lu, DW Nebert, DC Harrison, W Huang, L Jin: Mitochondrial aldehyde dehydrogenase-2 (ALDH2) Glu504Lys polymorphism contributes to the variation in efficacy of sublingual nitroglycerin. *Journal of Clinical Investigation* 116, 506-511 (2006) DOI: 10.1172/JCI26564
37. P Hui, T Nakayama, A Morita, N Sato, M Hishiki, K Saito, Y Yoshikawa, M Tamura, I Sato, T Takahashi, M Soma, Y Izumi, Y Ozawa, Z Cheng : Common single nucleotide polymorphisms in Japanese patients with essential hypertension: aldehyde dehydrogenase 2 gene as a risk factor independent of alcohol consumption. *Hypertension Research* 30, 585-592 (2007) DOI: 10.1291/hypres.30.585
38. SA Jo, EK Kim, MH Park, C Han, HY Park, Y Jang, BJ Song, I Jo: A Glu487Lys polymorphism in the gene for mitochondrial aldehyde dehydrogenase 2 is associated with myocardial infarction in elderly Korean men. *Clinica Chimica Acta* 382, 43-47 (2007) DOI: 10.1016/j.cca.2007.03.016
39. YC Chao, SR Liou, SF Tsai, SJ Yin: Dominance of the mutant ALDH2; allele in the expression of human stomach aldehyde dehydrogenase -2 activity. *Proceedings of the National Science Council, Republic of China* 17, 98-102 (1993) [DOI NOT FOUND]
40. HW Goedde, DP Agarwal, G Fritze, DM Tackmann, S Singh, G Beckmann, K Bhatia, LZ Chen, B Fang, R Lisker : Distribution of ADH2 and ALDH2 genotypes in different populations. *Human Genetics* 88, 344-6 (1992) DOI: 10.1007/BF00197271
41. TL Wall, ML Johnson, SM Horn, LG Carr, TL Smith, MA Schuckit : Evaluation of the self-rating of the effects of alcohol form in Asian Americans with aldehyde dehydrogenase polymorphisms. *Journal of Studies on Alcohol* 60, 784-789 (1999) [DOI NOT FOUND]

42. GS Peng, YC Chen, TP Tsao, MF Wang, SJ Yin: Pharmacokinetic and pharmacodynamic basis for partial protection against alcoholism in Asians, heterozygous for the variant ALDH2*2 gene allele. *Pharmacogenetics and Genomics* 17, 845-855 (2007) DOI: 10.1097/FPC.0b013e3282609e67

43. M Muto, Y Hitomi, A Ohtsu, S Ebihara, S Yoshida, H Esumi: Association of aldehyde dehydrogenase 2 gene polymorphism with multiple esophageal dysplasia's in head and neck cancer patients. *Gut* 47, 256-61 (2000) DOI: 10.1136/gut.47.2.256

44. M Fischer, LF Wetherill, LG Carr, M You, DW Crabb : Association of the aldehyde dehydrogenase 2 promoter polymorphism with alcohol consumption and reactions in an American Jewish population. *Alcoholism: Clinical and Experimental Research* 31, 1654-9 (2007) DOI: 10.1111/j.1530-0277.2007.00471.x

45. L Ni, S Sheikh, H Weiner: Involvement of glutamate 399 and lysine 192 in the mechanism of human liver mitochondrial aldehyde dehydrogenase. *Journal of Biological Chemistry* 272, 18823 -18826 (1997) DOI: 10.1074/jbc.272.30.18823

46. WJ Burke, SW Li, EA Williams, R Nonneman, DS Zahm: 3, 4-Dihydroxyphenylacetaldehyde is the toxic dopamine metabolite in vivo: implications for Parkinson's disease pathogenesis. *Brain Research* 989, 205-213 (2003) DOI: 10.1016/S0006-8993(03)03354-7

47. T Estey, J Piatigorsky, N Lassen, V Vasiliou: ALDH3A1: a corneal crystalline with diverse functions. *Experimental Eye Research* 84, 3-12 (2007) DOI: 10.1016/j.exer.2006.04.010

48. T Estey, M Cantore, PA Weston, JF Carpenter, JM Petrash, V Vasiliou: Mechanisms involved in the protection of UV-induced protein inactivation by the corneal crystalline ALDH3A1. *Journal of Biological Chemistry* 282, 4382-92 (2007) DOI: 10.1074/jbc.M607546200

49. M Kirsch, GH De: NAD (P) H, a directly operating antioxidant?. *FASEB Journal* 15, 1569-1574 (2001) DOI: 10.1096/fj.00-0823hyp

50. A Pappa, D Brown, Y Koutalos, J DeGregori, C White, V Vasiliou: Human Aldehyde Dehydrogenase 3A1 inhibits proliferation and promotes survival of Human corneal epithelial cells. *Journal of Biological Chemistry* 280 (30), 7998-8006 (2005) DOI: 10.1074/jbc.M503698200

51. N Lassen, A Pappa, WJ Black, JV Jester, BJ Day, E Min, V Vasiliou: Antioxidant functions of corneal ALDH3A1 in cultured stromal fibroblasts. *Free Radical Biology and Medicine* 41, 1459-69 (2006) DOI: 10.1016/j.freeradbiomed.2006.08.009

52. J Huang, N Hu, AM Goldstein, MR Emmert – Buck, ZZ Roth, QH Wang, SM Dawsey, XY Han, T Ding, G Li, C Giffen, P R Taylor: High frequency allelic loss on chromosome 17p13.3.-p11.1 in esophageal squamous cell carcinomas from a high incidence area in northern China. *Carcinogenesis* 21, 2019-26 (2000) DOI: 10.1093/carcin/21.11.2019

53. B Kim, HJ Lee, HY Choi, Y Shin, S Nam, G Seo, DS Son, J Jo, J Kim, J Lee, J Kim, K Kim, S Lee: Clinical validity of the lung cancer biomarkers identified by bioinformatics analysis of public expression data. *Cancer Research* 67, 7431-8 (2007) DOI: 10.1158/0008-5472.CAN-07-0003

54. V Vasiliou, SF Reuter, CA Kozak, DW Nebert: Mouse dioxin-inducible cytosolic aldehyde dehydrogenase-3: AHD4 cDNA sequence, genetic mapping, and differences in mRNA levels. *Pharmacogenetics* 3, 281-90 (1993) DOI: 10.1097/00008571-199312000-00002

55. R Reisdorph, R Lindahl: Constitutive and 3-methylcholanthrene-induced rat

ALDH3A1 expression is mediated by multiple xenobiotic response elements. *Drug Metabolism and Disposition* 35, 386-93 (2007)
DOI: 10.1124/dmd.106.012393

56. TL Kelson, JR Secor McVoy, WB Rizzo: Human liver fatty aldehyde dehydrogenase: microsomal localization, purification, and biochemical characterization. *Biochimica et Biophysica Acta* 1335 (1-2), 99-110 (1997)
DOI: 10.1016/S0304-4165(96)00126-2

57. C Chang, A Yoshida: Human fatty aldehyde dehydrogenase gene (ALDH10): organization and tissue-dependent expression. *Genomics* 40, 80-5 (1997)
DOI: 10.1006/geno.1996.4547

58. GR Rogers, NG Markova, V De Laurenzi, WB Rizzo, JG Compton: Genomic organization and expression of the human fattyaldehyde dehydrogenase gene (FALDH). *Genomics* 39, 127-35 (1997)
DOI: 10.1006/geno.1996.4501

59. WB Rizzo, G Carney: Sjogren-Larsson syndrome: diversity of mutations and polymorphisms in the fatty aldehyde dehydrogenase gene (ALDH3A2). *Human Mutation* 26, 1-10 (2005)
DOI: 10.1002/humu.20181

60. WB Rizzo: Sjogren-Larsson syndrome: molecular genetics and biochemical pathogenesis of fatty aldehyde dehydrogenase deficiency. *Molecular Genetics Metabolism* 90, 1-9 (2007)
DOI: 10.1016/j.ymgme.2006.08.006

61. D Demozay, S Rocchi, JC Mas: Fatty aldehyde dehydrogenase: potential role in oxidative stress protection and regulation of its gene expression by insulin. *Journal of Biological Chemistry* 279, 6261-70 (2004)
DOI: 10.1074/jbc.M312062200

62. LC Hsu, WC Chang, A Yoshida: Cloning of a cDNAs encoding human ALDH7, a new member of the aldehyde dehydrogenase family. *Gene* 15, 285-9 (1994)
DOI: 10.1016/0378-1119(94)90672-6

63. SA Marchitti, DJ Orlicky, V Vasiliou: Expression and initial characterization of human ALDH3B1. *Biochemical and Biophysical Research Communications* 356, 792-8 (2007)
DOI: 10.1016/j.bbrc.2007.03.046

64. Q Xu, YB Jia, BY Zhang, K Zou, YB Tao, YP Wang, BQ Qiang, GY Wu, Y Shen, HK Ji, Y Huang, X Q Sun, L Ji, YD Li, YB Yuan, L Shu, X Yu, YC Shen, YQ Yu, GZ Ju: Association study of an SNP combination pattern in the dopaminergic pathway in paranoid schizophrenia: A novel strategy for complex disorders. *Molecular Psychiatry* 9, 510-21 (2004)
DOI: 10.1038/sj.mp.4001472

65. X Sun, Y Jia, X Zhang, Q Xu, Y Shen, Y Li: Multi-locus association study of schizophrenia susceptibility genes with a posterior probability method. *Science China Life Sciences* 48, 263-9 (2005)
DOI: 10.1360/04yc0100

66. LC Hsu, WC Chang, A Yoshida: Human aldehyde dehydrogenase genes, ALDH7 and ALDH8: Genomic organization and gene structure comparison. *Gene* 189, 89-94 (1997)
DOI: 10.1016/S0378-1119(96)00839-6

67. O Shmueli, S Horn-Saban, V Chalifa-Caspi: Gene Note: whole genome expression profiles in normal human tissues. *Competes Rendus Biologies* 326, 1067-72 (2003)
DOI: 10.1016/j.crvi.2003.09.012

68. CA Hu, WW Lin, D Valle: Cloning, characterization, and expression of cDNAs encoding human delta 1-pyrroline-5-carboxylate dehydrogenase. *Journal of Biological Chemistry* 271, 9795-800 (1996)
DOI: 10.1074/jbc.271.16.9795

69. KA Yoon, Y Nakamura, H Arakawa: Identification of ALDH4 as a p53-inducible gene and its protective role in cellular stresses. *Journal of Human Genetics* 49, 134-40 (2004)

DOI: 10.1007/s10038-003-0122-3

70. J Farres, P Julia, X Pares: Aldehyde oxidation in human placenta. Purification and properties of 1-pyrroline-5-carboxylate dehydrogenase. *Biochemical Journal* 256, 461-7 (1998)
[DOI NOT FOUND]

71. RD Farrant, V Walker, GA Mills, JM Mellor, GJ Langley: Pyridoxal phosphate de-activation by pyrroline-5-carboxylic acid. Increased risk of vitamin B6 deficiency and seizures in hyperprolinemia type II. *Journal of Biological Chemistry* 276, 15107-16 (2001)
DOI: 10.1074/jbc.M010860200

72. MT Geraghty, D Vaughn, AJ Nicholson, WW Lin, G Jimenez-Sanchez, C Obie, M P Flynn, D Valle, CA Hu: Mutations in the Delta1-pyrroline 5-carboxylate dehydrogenase gene cause type II hyperprolinemia. *Human Molecular Genetics* 7, 1411-5 (1998)
DOI: 10.1093/hmg/7.9.1411

73. E Inagaki, N Ohshima, H Takahashi, C Kuroishi, S Yokoyama, TH Tahirov : Crystal structure of Thermus thermophilus Δ1-pyrroline-5-carboxylate dehydrogenase. *Journal of Molecular Biol*ogy 362, 490-501(2006)
DOI: 10.1074/jbc.M010860200

74. JH Kang, YB Park, TL Huh, WH Lee, MS Choi, OS Kwon : High-level expression and characterization of the recombinant enzyme, and tissue distribution of human succinic semialdehyde dehydrogenase. *Protein Expression and Purification* 44, 16-22 (2005)
DOI: 10.1016/j.pep.2005.03.019

75. BI Gold, RH Roth : Kinetics of in vivo conversion of gamma- (3H) amino butyric acid to gamma- (3H) hydroxybutyric acid by rat brain. *Journal of Neurochemistry* 28, 1069-73 (1977)
DOI: 10.1111/j.1471-4159.1977.tb10670.x

76. M Maitre, C Andriamampandry, V Kemmel, C Schmidt, Y Hode, V Hechler, S Gobaille: Gamma-hydroxybutyric acid as a signaling molecule in brain. *Alcohol* 20, 277-83 (2000)
DOI: 10.1016/S0741-8329(99)00092-0

77. S Akaboshi, BM Hogema, A Novelletto, P Malaspina, GS Salomons, GD Maropoulos, C Jakobs, M Grompe, KM Gibson : Mutational spectrum of the succinate semialdehyde dehydrogenase (ALDH5A1) gene and functional analysis of 27 novel disease-causing mutations in patients with SSADH deficiency. *Human Mutation* 22, 442-50 (2003)
DOI: 10.1002/humu.10288

78. NY Kedishvili, KM Popov, PM Rougraff, Y Zhao, DW Crabb, RA Harris: CoA-dependent methylmalonate-semialdehyde dehydrogenase, a unique member of the aldehyde dehydrogenase superfamily. cDNA cloning, evolutionary relationships, and tissue distribution. *Journal of Biological Chemistry* 267, 19724-9 (1992)
[DOI NOT FOUND]

79. NY Kedishvili, KM Popov, JA Jaskiewicz, RA Harris: Coordinated expression of valine catabolic enzymes during adipogenesis: analysis of activity, mRNA, protein levels, and metabolic consequences. *Archives of Biochemistry and Biophysics* 315, 317-22 (1994)
DOI: 10.1006/abbi.1994.1506

80. KL Chambliss, RG Gray, G Rylance, RJ Pollitt, KM Gibson : Molecular characterization of methylmalonate semialdehyde dehydrogenase deficiency. *Journal of Inherited Metabolic Disease* 23, 497-504 (2000)
DOI: 10.1023/A:1005616315087

81. YF Chang, P Ghosh, VV Rao: L-pipecolic acid metabolism in human liver: L-alpha-aminoadipatedelta-semialdehyde oxidoreductase. *Biochimica et Biophysica Acta* 1038, 300-5 (1990)
DOI: 10.1016/0167-4838(90)90241-7

82. C Brocker, N Lassen, T Estey, A Pappa, M Cantore, VV Orlova, T Chavakis, KL Kavanagh, U Oppermann, V Vasiliou :

Aldehyde dehydrogenase7A (ALDH7A1) is a novel enzyme involved in cellular defense against hyper osmotic stress. *Journal of Biological Chemistry* 285, 18452-18463 (2010)
DOI: 10.1074/jbc.M109.077925

83. PB Mills, E Struys, C Jakobs, B Plecko, P Baxter, M Baumgartner, MA Willemsen, H Omran, U Tacke, B Uhlenberg, B Weschke, PT Clayton: Mutations in antiquitin in individuals with pyridoxine-dependent seizures. *Nature Medicine* 12, 307-9 (2006)
DOI: 10.1038/nm1366

84. P Lheureux, A Penaloza, M Gris: Pyridoxine in clinical toxicology: a review. *European Journal of Emergency Medicine* 12, 78-85 (2005)
DOI: 10.1097/00063110-200504000-00007

85. M Lin, JL Napoli: cDNA cloning and expression of a human aldehyde dehydrogenase (ALDH) active with 9-cis-retinal and identification of a rat ortholog, ALDH12. *Journal of Biological Chemistry* 275, 40106-12 (2000)
DOI: 10.1074/jbc.M008027200

86. G Izaguirre, A Kikonyogo, R Pietruszko: Tissue Distribution of Human Aldehyde Dehydrogenase E3 (ALDH9): Comparison of Enzyme Activity with E3 Protein and mRNA Distribution. *Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology* 118, 59-64(1997)
DOI: 10.1006/geno.1996.0300

87. SW Lin, JC Chen, LC Hsu, CL Hsieh, A Yoshida: Human gamma-aminobutyraldehyde dehydrogenase (ALDH9): cDNA sequence, genomic organization, polymorphism, chromosomal localization, and tissue expression. *Genomics* 15, 376-80 (1996)
DOI: 10.1074/jbc.275.10.7390

88. W Sato, Y Horie, E Kataoka, S Ohshima, T Dohmen, M Iizuka, J Sasaki, T Sasaki, K Hamada, H Kishimoto, A Suzuki, S Watanabe: Hepatic gene expression in hepatocyte-specific Pten deficient mice showing steatohepatitis without ethanol challenge. *Hepatology Research* 34, 256-65 (2006)
DOI: 10.1016/j.hepres.2006.01.003

89. RL Strausberg, EA Feingold, LH Grouse: Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences. *Proceedings of the National Academy of Sciences* 99, 16899-03 (2002)
DOI: 10.1073/pnas.242603899

90. CA Hu, WW Lin, C Obie, D Valle: Molecular enzymology of mammalian Delta1-pyrroline-5- carboxylate synthase. Alternative splice donor utilization generates isoform with different sensitivity to ornithine inhibition. *Journal of Biological Chemistry* 274, 6754-62 (1999)
DOI: 10.1074/jbc.274.10.6754

91. P Kamoun, B Aral, JM Saudubray: A new inherited metabolic disease: delta1-pyrroline 5-carboxylate synthetase deficiency. *Bulletin de l'Académie Nationale de Medecine* 182, 131-7 (1998)
[DOI NOT FOUND]

92. MR Baumgartner, D Rabier, MC Nassogne, JL Dufier, JP Padovani, P Kamoun, D Valle, JM Saudubray: Delta1-pyrroline-5-carboxylate synthase deficiency: Neurodegeneration, cataracts and connective tissue manifestations combined with hyperammonemia and reduced ornithine, citrulline, arginine and proline. *European Journal of Pediatrics* 164, 31-6 (2005)
DOI: 10.1007/s00431-004-1545-3

93. H Mitsubuchi, K Nakamura, S Matsumoto, F Endo: Inborn Errors of Proline Metabolism. *Journal of Nutrition* 138, 2016S-2020S (2008)
[DOI NOT FOUND]

94. D Fredman, M Siegfried, YP Yuan, P Bork, HH Lehvaslaiho, AJ Brookes: HGVBase: A human sequence variation database emphasizing data quality and a broad spectrum of data sources. *Nucleic Acids*

Research 30, 387-391 (2002)
DOI: 10.1093/nar/30.1.387

95. M Elizabeth, Smigielski, K Sirotkin, M Ward, ST Sherry: dbSNP: A data base of Single nucleotide polymorphisms. *Nucleic acid research* 28, 352-355 (2000)
DOI: 10.1093/nar/28.1.352

96. MT Hirakawa, T Tanaka, Y Hashimoto, M Kuroda, T Takagi, Y Nakamura: JSNP: A database of common gene variations in the Japanese population. *Nucleic Acids Research* 30, 158-162 (2002)
DOI: 10.1093/nar/30.1.158

97. DP Stenson, M Mort, EV Ball, K Shaw, A Phillips, DN Cooper: The Human Gene Mutation Data base: Building a comprehensive mutation repository for clinical and molecular genetics, diagnostic and personalized medicine. *Human Genetics* 133, 1-9 (2014)
DOI: 10.1007/s00439-013-1358-4

98. The Uniprot Consortium: Activities at the Universal Protein Resource (UniProt). *Nucleic acid Research* 42, D191-D198 Data base issue (2014)
[DOI NOT FOUND]

99. T Schwed, J Kopp, N Guex, CM Peitsch: SWISS-MODEL: An automated protein homology- Modeling server. *Nucleic Acids Research* 31, 3381-3385 (2003)
DOI: 10.1093/nar/gkg520

100. MA Marti-Renom, AC Stuart, A Fiser, R Sanchez, F Melo, A Sali: Comparative protein structure modeling of genes and genomes. *Annual Review of Biophysics and Biomolecular Structure* 29, 291-325 (2000)
DOI: 10.1146/annurev.biophys.29.1.291

101. Y Zhang: I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics* 9, 40 (2008)
DOI: 10.1186/1471-2105-9-40

102. ED Kim, D Chivian, D Baker: Protein structure prediction and analysis using the Robetta server. *Nucleic Acids Research* 32, Web Server issue (2004)
[DOI NOT FOUND]

103. E Krieger, G Koraimann, G Vriend: Increasing the precision of comparative models with YASARA NOVA - a self-parameterizing force field. *Proteins* 47, 393-402 (2002)
DOI: 10.1002/prot.10104

104. H Evan, Baugh, S Lysko, D Brian, Weitzner, J Jeffrey, Gray: Real-Time PyMOL Visualization for Rosetta and PyRosetta. *PLoS ONE* 6 (8), e21931 (2011)
DOI: 10.1371/journal.pone.0021931

105. JT Pembroke: Bio-molecular modeling utilizing Rasmol and PDB resources: A tutorial with HEW lysozyme. *Biochemistry and Molecular Biology Education* 28, 297-300 (2000)
DOI: 10.1111/j.1539-3429.2000.tb00177.x

106. N Guex, MC Peitsch: SWISS- MODEL and the Swiss- Pdb Viewer: an environment for comparative protein modeling. *Electrophoresis* 15, 2714-23 (1997)
DOI: 10.1002/elps.1150181505

107. W Humphrey, A Dalke, K Schulten: VMD: Visual Molecular Dynamics. *Journal of Molecular Graphics* 14, 33-38 (1996)
DOI: 10.1016/0263-7855(96)00018-5

108. E Krieger, K Joo, J Lee, J Lee, S Raman, J Thompson, M Tyka, D Baker, K Karplus: Improving physical realism, stereochemistry and side-chain accuracy in homology modeling: four approaches that performed well in CASP8. *Proteins* 77 (Suppl 9), 114-122 (2009)
DOI: 10.1002/prot.22570

109. A Han, HJ Kang, Y Cho, S Lee, YJ Kim, S Gong: SNP@Domain: a web resource of single nucleotide polymorphisms (SNPs) within protein domain structures and sequences. *Nucleic Acids Research* 34, W642-W644 (2006)
DOI: 10.1093/nar/gkl323

110. ON Stitziel, TA Binkowski, YY Tseng, S Kasif, J Liang: TopoSNP: a topographic database of non-synonymous single nucleotide polymorphisms with and without known disease association.

Nucleic Acids Research 32, D520 - D522 (2004)
DOI: 10.1093/nar/gkh104

111. YL Yip, H Scheib, AV Diemand, A Gattiker, LM Famiglietti, E Gasteiger, A Bairoch: The Swiss-Prot variant page and the ModSNP database: a resource for sequence and structure information on human protein variants. *Human Mutation* 23, 464-470 (2004)
DOI: 10.1002/humu.20021

112. R Karchin, M Diekhans, L Kelly, DJ Thomas, U Pieper, N Eswar, D Haussler, A Sali: LS-SNP: large-scale annotation of coding non-synonymous SNPs based on multiple information sources. *Bioinformatics* 12, 2814-2820 (2005)
DOI: 10.1093/bioinformatics/bti442

113. A Uzun, M Chesley, Leslin, A Abyzov, V Ilyin: Structure SNP (StSNP): a web server for mapping and modeling nsSNPs on protein structures with linkage to metabolic pathways. *Nucleic Acids Research* 35, Web Server issue (2007)
[DOI NOT FOUND]

114. H Kono, T Yuasa, S Nishiue, K Yura: ColiSNP database server mapping nsSNPs on protein structures. *Nucleic Acids Research* 36, D409-D413 (2008)
DOI: 10.1093/nar/gkm801

115. SN Al-Numair, AC Martin: The SAAP pipeline and database: tools to analyze the impact and predict the pathogenecity of mutations. *BMC Genomics* 14 (Suppl 3), S4 (2013)
[DOI NOT FOUND]

116. M Kanehisa: The KEGG database. *Novartis Found Symposium* 247, 91-101 (2002)
DOI: 10.1002/0470857897.ch8

117. A Abyzov, C Leslin, VA Ilyin: An integrated analytical front-end application for bioinformatics. *Friend.* CSBi Conference at MIT (2003)
[DOI NOT FOUND]

118. TD Luu, A Rusu, V Walter, B Linard, L Poideven, R Ripp, L Moulinier, J Muller, W Raffelsberger, N Wicker, O Lecompte, JD Thompson, O Poch, H Nguyen : KD4v: comprehensible knowledge discovery system for missense variant. *Nucleic acid Research* 40, W71-W75 Web server issue (2012)
[DOI NOT FOUND]

119. H Veneselaar, TA te Beek, RP Kuipers, ML Hekkelman, G Vriend: Protein structure analysis of mutations causing inheritable diseases. An e-Science approach with life scientist friendly interfaces. *BMC Bioinformatics* 11, 548 (2010)
DOI: 10.1186/1471-2105-11-548

120. S Muggleton: Inductive logic programming. *New Generation Computing* 8, 295-318 (1991)
DOI: 10.1007/BF03037089

121. TD Luu, AM Rusu, V Walter, R Ripp, L Moulinier, J Muller, T Toursel, JD Thompson, O Poch, H Nguyen: MSV3d: database of human Missense variants mapped to 3D protein structure. Database: *The Journal of Biological Databases and Curation* (2012)
[DOI NOT FOUND]

122. IA Adzhubei, S Schmidt, L Peshkin, VE Ramensky, A Gerasimova, P Bork, AS Kondrashov, SR Sunyaeu: A method and server for predicting damaging missense mutations. *Nature Methods* 7, 248-249 (2010)
DOI: 10.1038/nmeth0410-248

123. F Plewniak, L Bianchetti, Y Brelivet, A Carles, F Chalmel, O Lecompte, T Mochel, L Moulinier, A Muller, J Muller, V Prigent, R Ripp, JC Therry, JD Thompson, N Wicker, O Poch: Pipe Align: A new toolkit for protein family analysis. *Nucleic Acids Research* 31, 3829-3832 (2003)
DOI: 10.1093/nar/gkg518

124. A Mottaz, FP David, AL Veuthey, YL Yip: Easy retrieval of single amino-acid polymorphisms and phenotype information using SwissVar. *Bioinformatics* 26, 851-852 (2010)
DOI: 10.1093/bioinformatics/btq028

125. A Friedrich, N Garnier, N Gagniere, H Nguyen, LP Albou, V Biancalana, E Bettler, G Deleage, O Lecompte, J Muller, D Moras, JL Mandel, T Toursel, L Moulinier, O Poch : SM2PH-db: an interactive system for the integrated analysis of phenotypic consequences of missense mutations in proteins involved in human genetic diseases. *Human Mutation* 31, 127-135 (2010)
DOI: 10.1002/humu.21155

126. A Andreeva, D Howorth, JM Chandonia, SE Brenner, TJ Hubbard, C Chothia, AG Murzin: Data growth and its impact on the SCOP database: new developments. *Nucleic Acids Research* 36, D419–D425 (2008)
DOI: 10.1093/nar/gkm993

127. JD Thompson, A Muller, A Waterhouse, J Procter, GJ Barton, F Plewniak, O Poch: MACSIMS: multiple alignment of complete sequences information management system. *BMC Bioinformatics* 7, 318 (2006)
DOI: 10.1186/1471-2105-7-318

128. N Eswar, D Eramian, B Webb, MY Shen, A Sali: Protein structure modeling with MODELLER. *Methods in molecular biology* 426, 145-159 (2008)
DOI: 10.1007/978-1-60327-058-8_8

129. V Sobolev, A Sorokine, J Prilusky, EE Abola, M Edelman: Automated analysis of inter atomic contacts in proteins. *Bioinformatics* 15, 327-332 (1999)
DOI: 10.1093/bioinformatics/15.4.327

130. E Capriotti, P Fariselli, R Casadio: I-Mutant 2.0: predicting stability changes upon mutation from the protein sequence or structure. *Nucleic Acids Research* 33, W306-W310 (2005)
DOI: 10.1093/nar/gki375

131. SF Altschul, W Gish, W Miller, EW Myers, DJ Lipman: Basic local alignment search tool. *Journal of Molecular Biology* 215 (3), 403-410 (1990)
DOI: 10.1016/S0022-2836(05)80360-2

132. The Universal Protein Resource (UniProt) in 2010: *Nucleic Acids Research* 38, D142-148 (2010)
DOI: 10.1093/nar/gkp846

133. H Berman, K Henrick, H Nakamura: Announcing the worldwide Protein Data Bank. *Nature Structural Biology* 10 (12), 980 (2003)
DOI: 10.1038/nsb1203-980

134. R Chenna, H Sugawara, T Koike, R Lopez, TJ Gibson, DG Higgins, JD Thompson: Multiple sequence alignment with the Clustal series of programs. *Nucleic Acids Research* 31 (13), 3497-3500 (2003)
DOI: 10.1093/nar/gkg500

135. ML Hekkelman, TA Te Beek, SR Pettifer, D Thorne, TK Attwood, G Vriend: WIWS: a protein structure bioinformatics Web service collection. *Nucleic Acids Research* 38 (Suppl), W719-23 (2010)
DOI: 10.1093/nar/gkq453

136. G Vriend: WHAT IF: a molecular modeling and drug design program. *Journal of Molecular Graphics* 8 (1), 52-56 (1980)
DOI: 10.1016/0263-7855(90)80070-V

137. E Jain, AS Bairoch, I Phan, N Redaschi, BE Suzek, MJ Martin, P McGarvey, E Gasteiger: Infrastructure for the life sciences: design and implementation of the UniProt website. *BMC Bioinformatics* 10, 136 (2009)
DOI: 10.1186/1471-2105-10-136

138. C Dodge, R Schneider, C Sander: The HSSP database of protein structure sequence alignments and family profiles. *Nucleic Acids Research* 26 (1), 313-315 (1998)
DOI: 10.1093/nar/26.1.313

139. B Rost: PHD: predicting one-dimensional protein structure by profile based neural networks. *Methods in Enzymology* 266, 525-539 (1996)
DOI: 10.1016/S0076-6879(96)66033-9

140. L Kall, A Krogh, EL Sonnhammer: A combined transmembrane topology and signal peptide prediction method. *Journal of Molecular Biology* 338 (5), 1027-10369 (2004)
DOI: 10.1016/j.jmb.2004.03.016

141. N Blom, S Gammeltoft, S Brunak: Sequence and structure-based prediction of eukaryotic protein phosphorylation sites. *Journal of Molecular Biology* 294 (5), 1351-1362 (1999)
DOI: 10.1006/jmbi.1999.3310

142. JJ Nora: Multifactorial inheritance hypothesis for the etiology of congenital heart diseases. The genetic-environmental interaction. *Circulation* 38 (3), 604-617 (1968)
DOI: 10.1161/01.CIR.38.3.604

143. V Garg, IS Kathiriya, R Barnes, MK Schluterman, IN King, CA Butler, CR Rothrock, RS Eapen, KH Yamada, K Joo, R Matsuoka, JC Cohen, D Srivastava: GATA4 mutations cause human congenital heart defects and reveal an interaction with TBX5. *Nature* 424 (6947), 443-447 (2003)
DOI: 10.1038/nature01827

144. N Rosenthal, RP Harvey: Single allele mutations at the heart of congenital disease. *Journal of Clinical Investigation* 104 (11), 1483-1484 (1999)
DOI: 10.1172/JCI8825

145. G Bonne, L Carrier, P Richard, B Hainque, K Schwartz: Familial hypertrophic cardiomyopathy: from mutations to functional defects. *Circulation Research* 83 (6), 580-593 (1998)
DOI: 10.1161/01.RES.83.6.580

146. M Pavan, FV Ruiz, AF Silva, JT Sobreir, MR Cravo, M Vasconcelos, PL Marques, MF Mesquita, JE Krieger, AA Lopes, PS Oliveira, AC Pereira, J Xavier-Neto: ALDH1A2 (RALDH2) genetic variation in human congenital heart disease. *BMC Medical Genetics* 10, 113 (2009)
DOI: 10.1186/1471-2350-10-113

147. AS Verma, DR Fitzpatrick: Anophthalmia and Microphthalmia. *Orphanet Journal of Rare Diseases* 2, 47 (2007)
DOI: 10.1186/1750-1172-2-47

148. S Tucker, B Jones, R Collin: Systemic anomalies in 77 patients with congenital Anophthalmos or Microphthalmos. *Eye* (Lond.) 10, 310-314 (1996)
DOI: 10.1038/eye.1996.65

149. TM Bardakjian, A Schneider: The genetics of Anophthalmia and Microphthalmia. *Current Opinion in Ophthalmology* 22, 309-313 (2011)
DOI: 10.1097/ICU.0b013e328349b004

150. LF Taie, S Gerber, N Chassaing, JC Smith, S Hanein, E Silva, M Serey, V Serre, X Gerard, C Baumann, G Plessis, B Demeer, L Bretillon, C Bole, P Nitschke, A Munnich, S Lyonnet, P Calvas, J Kaplan, N Ragge, JM Rozet: ALDH1A3 Mutations Cause Recessive Anophthalmia and Microphthalmia. *The American Journal of Human Genetics* 92, 265-270 (2013)
DOI: 10.1016/j.ajhg.2012.12.003

151. AM Slavotinek: Eye development genes and known syndromes. *Molecular Genetics and Metabolism* 104, 448-456 (2011)
DOI: 10.1016/j.ymgme.2011.09.029

152. LL Husemoen, M Fenger, N Friedrich, JS Tolstrup, SB Fredriksen, A Linneberg: The association of ADH and ALDH gene variants with alcohol drinking habits and cardiovascular disease risk factors. *Alcoholism: Clinical and Experimental Research* 32, 1984-1991 (2008)
[DOI NOT FOUND]

153. A Linneberg, AG Quintela, C Vidal, T Jorgensen, M Fenger, T Hansen, O Pedersen, LL Husemoen: Genetic determinants of both ethanol and acetaldehyde metabolism influence alcohol hypersensitivity and drinking behavior among Scandinavians. *Clinical & Experimental Allergy* 40, 123-130 (2010)
DOI: 10.1111/j.1365-2222.2009.03398.x

154. D Rossi, D Capello, A Gloghini, S Franceschetti, M Paulli, K Bhatia, G Saglio, U Vitolo, SA Pileri, M Esteller, A Carbone, G Gaidano: Aberrant promoter methylation of multiple genes throughout the clinico-pathologic spectrum of B-cell neoplasia. *Haematologica* 89 (2), 154-164 (2004)

[DOI NOT FOUND]

155. R Kuppers: Mechanisms of B-cell lymphoma pathogenesis. *Nature Review Cancer* 5 (4), 251-262 (2005) DOI: 10.1038/nrc1589

156. YI Kim: 10-Methylenetetrahydrofolate reductase polymorphisms and phamacogenetics: a new role of single nucleotide polymorphisms in the folate metabolic pathway in human health and disease. *Nutrition Reviews* 63 (11), 398-407 (2005) DOI: 10.1111/j.1753-4887.2005.tb00377.x

157. AJ Gonzalez Ordonez, JM Fernandez Carreira, CR Fernandez Alvarez, L Martin, J Sanchez Garcia, JM Medina Rodriguez, MV Alvarez, E Coto: Normal frequencies of the C677T genotypes on the methylene tetrahydrofolate reductase (MTHFR) gene among lymph proliferative disorders but not in multiple myeloma. *Leukemia Lymphoma* 39 (5- 6), 607-612 (2000) DOI: 10.3109/10428190009113391

158. U Lim, SS Wang, P Hartge, W Cozen, LE Kelemen, S Chanock, S Davis, A Blair, M Schenk, N Rothman, Q LAN: Gene-nutrient interactions among determinants of folate and one-carbon metabolism on the risk of non-Hodgkin lymphoma: NCI-SEER Case-Control Study. *Blood* 109 (7), 3050-3059 (2007) [DOI NOT FOUND]

159. S Niclot, Q Pruvot, C Besson, D Savoy, E Macintyre, G Salles, N Brousse, B Varet, P Landais, P Taupin, C Junien, DB Bluteau: Implication of the folate-methionine metabolism pathways in susceptibility to follicular lymphomas. *Blood* 108 (1), 278-285 (2006) DOI: 10.1182/blood-2005-04-1567

160. A Timuragaoglu, S Dizlek, N Uysalgil, O Tosun, K Yamac: Methylene tetrahydrofolate reductase C677T polymorphism in adult patients with lymph proliferative disorders and its eVect on chemotherapy. *Annals of Hematology* 85 (12), 863-868 (2006) DOI: 10.1007/s00277-006-0175-4

161. SN Reuland, AP Vlasov, SA Krupenko: Modular organization of FDH: Exploring the basis of hydrolase catalysis. *Protein Science* 15 (5), 1076-1084 (2006) DOI: 10.1110/ps.052062806

162. SA Krupenko, NV Oleinik: 10-formyltetrahydrofolate dehydrogenase, one of the major folate enzymes, is down-regulated in tumor tissues and possesses suppressor effects on cancer cells. *Cell Growth & Differentiation* 13 (5), 227-236 (2002) [DOI NOT FOUND]

163. MC Anguera, MS Field, C Perry, H Ghandour, EP Chiang, J Selhub, B Shane, PJ Stover: Regulation of folate-mediated one carbon metabolism by 10-formyltetrahydrofolate dehydrogenase. *Journal of Biological Chemistry* 281 (27), 18335-18342 (2006) DOI: 10.1074/jbc.M510623200

164. R Castelletto, X Castellsague, N Munoz, J Iscovich, N Chopita, A Jmelnitsky: Alcohol, tobacco, diet, mates drinking, and esophageal cancer in Argentina. *Cancer Epidemiology Biomarkers Prevention* 3, 557-64 (1994) [DOI NOT FOUND]

165. Y Kinjo, Y Cui, S Akiba, S Watanabe, N Yamaguchi, T Sobue, S Mizuno, V Beral: Mortality risks of esophageal cancer associated with hot tea, alcohol, tobacco and diet in Japan. *Journal of Epidemiology* 8, 235-43 (1998) DOI: 10.2188/jea.8.235

166. LM Brown, SS Devesa: Epidemiologic trends in esophageal and gastric cancer in the United States. *Surgical Oncology Clinics of North America* 11, 235-56 (2002) DOI: 10.1016/S1055-3207(02)00002-9

167. Y Zhang, J Ren: ALDH2 in alcoholic heart diseases: Molecular mechanism and clinical implications. *Pharmacology & Therapeutics* 132, 86-95 (2011)

DOI: 10.1016/j.pharmthera.2011.05.008

168. AL Bui, TB Horwich, GC Fonarow: Epidemiology and risk profile of heart failure. *Nature Reviews Cardiology* 8, 30-41 (2011)
DOI: 10.1038/nrcardio.2010.165

169. F Xu, Y Chen, R Lu, H Zhang, H Tian, Y Bian, J Feng, Y Sun, R Li, R Wang, Y Zhang: ALDH2 genetic polymorphism and the risk of type II diabetes mellitus in CAD patients. *Hypertension Research* 33, 49-55 (2010)
DOI: 10.1038/hr.2009.178

170. H Ma, J Li, F Gao, J Ren: Aldehyde dehydrogenase 2 ameliorates acute cardiac toxicity of ethanol: role of protein phosphatase and fork head transcription factor. *Journal of the American College of Cardiology* 54, 2187-2196 (2009)
DOI: 10.1016/j.jacc.2009.04.100

171. J Liao, A Sun, Y Xie, T Isse, T Kawamoto, T Zou, J Ge: Aldehyde dehydrogenase-2 deficiency aggravates cardiac dysfunction elicited by endoplasmic reticulum stress induction. *Molecular Medicine* 18, 785-793 (2012)
DOI: 10.2119/molmed.2012.00205

172. H Ma, R Guo, L Yu, Y Zhang, J Ren: Aldehyde dehydrogenase 2 (ALDH2) rescues Myocardia Ischemia/reperfusion injury: role of autophagy paradox and toxic aldehyde. *European Heart Journal* 32, 1025-1038 (2011)
DOI: 10.1093/eurheartj/ehq253

173. GR Budas, MH Disatnik, D Mochly-Rosen: Aldehyde dehydrogenase 2 in cardiac protection: a new therapeutic target?. *Trends in Cardiovascular Medicine* 19, 158-164 (2009)
DOI: 10.1016/j.tcm.2009.09.003

174. BF Zhang, Y Zhang, H Karissa, L Cour, KL Richmond, X-Ming Wang, J Ren: Mitochondrial aldehyde dehydrogenase obliterates endoplasmic reticulum stress-induced cardiac contractile dysfunction via correction of autophagy. *Biochimica et Biophysica Acta* 1832, 574-584 (2013)
DOI: 10.1016/j.bbadis.2013.01.013

175. CH Chen, GR Budas, EN Churchill, MH Disatnik, TD Hurley, D Mochly-Rosen: Activation of aldehyde dehydrogenase-2 reduces ischemic damage to the heart. *Science* 321, 1493-1495 (2008)
DOI: 10.1126/science.1158554

176. CH Chen, L Sun, D Mochly- Rosen: Mitochondrial aldehyde dehydrogenase and cardiac diseases. *Cardiovascular Research* 88, 51-57 (2010)
DOI: 10.1093/cvr/cvq192

177. J Wang, H Wang, P Hao, L Xue, S Wei, Y Zhang, Y Chen: Inhibition of aldehyde dehydrogenase 2 by oxidative stress is associated with cardiac dysfunction in diabetic rats. *Molecular Medicine* 17, 172-179 (2011)
DOI: 10.2119/molmed.2010.00114

178. PH Wolff: Ethnic differences in alcohol sensitivity. *Science* 175, 449-450 (1972)
DOI: 10.1126/science.175.4020.449

179. B Wang, J Wang, S Zhou, S Tan, X He, Z Yang, YC Xie, S Li, C Zheng, X Ma: The association of mitochondrial aldehyde dehydrogenase gene (ALDH2) polymorphism with susceptibility to late-onset Alzheimer's disease in Chinese. *Journal of the Neurological Sciences* 268, 172-175 (2008)
DOI: 10.1016/j.jns.2007.12.006

180. E Giovannucci: Alcohol, one-carbon metabolism, and colorectal cancer: recent insights from molecular studies. *Journal of Nutrition* 134, 2475S-2481S (2004)
[DOI NOT FOUND]

181. NP Singh, A Khan: Acetaldehyde: genotoxicity and cytotoxicity in human lymphocytes. *Mutation Research* 337, 9-17 (1995)
DOI: 10.1016/0921-8777(95)00006-6

182. A Yokoyama, T Muramatsu, T Omori, T Yokoyama, S Matsushita, S Higuchi, K Maruyama, H Ishii: Alcohol and aldehyde dehydrogenase gene polymorphisms and oropharyngolaryngeal, esophageal and stomach cancers in Japanese alcoholics.

Carcinogenesis 22, 433-439 (2001)
DOI: 10.1093/carcin/22.3.433

183. A Yoshida, LC Hsu, MM Yasunami: Genetics of human alcohol-metabolizing enzymes. *Progress in Nucleic Acid Research and Molecular Biology* 40, 255-28 (1991)
DOI: 10.1016/S0079-6603(08)60844-2

184. DW Crabb, HJ Edenberg, WF Bosron, TK Li: Genotypes for aldehyde dehydrogenase deficiency and alcohol sensitivity. The inactive ALDH2 (2) allele is dominant. *Journal of Clinical Investigation* 83, 314-316 (1989)
DOI: 10.1172/JCI113875

185. HR Thomason, DW Crabb, HJ Edenberg, TK Li: Alcohol and aldehyde dehydrogenase polymorphisms and alcoholism. *Behavior Genetics* 23, 131-136 (1993)
DOI: 10.1007/BF01067417

186. T Sjo¨gren, T Larsson: Oligophrenia in combination with congenital ichthyosis and spastic disorders. *Acta Psychiatrica et Neurologica Scandinavica* 32, 1-113 (1957)
[DOI NOT FOUND]

187. BW Rizzo, G Carney, Z Lin: The Molecular Basis of Sjo¨gren-Larsson Syndrome: Mutation Analysis of the Fatty Aldehyde Dehydrogenase Gene. *American Journal of Human Genetics* 65, 1547-1560 (1999)
DOI: 10.1086/302681

188. GR Rogers, NG Markova, V De Laurenzi, WB Rizzo, JG Compton: Genomic organization and expression of the human fatty aldehyde dehydrogenase gene (FALDH). *Genomics* 39, 127-135 (1997)
DOI: 10.1006/geno.1996.4501

189. R Masaki, A Yamamoto, Y Tashiro: Microsomal aldehyde dehydrogenase is localized to the endoplasmic reticulum via its carboxyl-terminal 35 amino acids. *Journal of Cell Biology* 126, 1407-1420 (1994)
DOI: 10.1083/jcb.126.6.1407

190. D Srivastava, RK Singh, MA Moxley, MT Henzel, DF Becker, JJ Tanner: The Three-Dimensional Structural Basis of Type II Hyperprolinemia. *Journal of Molecular Biology* 420, 176-189 (2012)
DOI: 10.1016/j.jmb.2012.04.010

191. H Jacquet, C Demily, E Houy, B Hecketsweiler, J Bou, G Raux, J Lerond, G Allio, S Haouzir, A Tillaux, C Bellegou, G Fouldrin, P Delamillieure, JF Menard, S Dollfus, TD Amato, M Petit, F Thibaut, T Frebourg, D Campion: Hyperprolinemia is a risk factor for schizoaffective disorder. *Molecular Psychiatry* 10, 479-485 (2005)
DOI: 10.1038/sj.mp.4001597

192. CL Clelland, LL Read, AN Baraldi, CP Bart, CA Pappas, LJ Panek, RH Nadrich, Clelland: Evidence for association of Hyperprolinemia with schizophrenia and a measure of clinical outcome. *Schizophrenia Research* 131, 139-145 (2011)
DOI: 10.1016/j.schres.2011.05.006

193. AT Wyse, CA Netto: Behavioral and neurochemical effects of proline. *Metabolic Brain Disease* 26, 159-172 (2011)
DOI: 10.1007/s11011-011-9246-x

194. F He, PJ DiMario: Drosophila delta-1-pyrroline-5-carboxylate dehydrogenase (P5CDh) is required for proline breakdown and mitochondrial integrity—establishing a fly model for human type II hyperprolinemia. *Mitochondrion* 11, 397-404 (2011)
DOI: 10.1016/j.mito.2010.12.001

195. I Knerr, PL Pearl, T Bottiglieri, OC Snead, C Jakobs, KM Gibson: Therapeutic concepts in succinate semialdehydes dehydrogenase (SSADH; ALDH5a1) deficiency (g-hydroxybutyric aciduria). Hypotheses evolved from 25 years of patient evaluation, studies in Aldh5a1 mice and characterization of g-hydroxybutyric acid pharmacology. *Journal of Inherited Metabolic Disease* 30, 279-294 (2007)
DOI: 10.1007/s10545-007-0574-2

196. KM Gibson, E Christensen, C Jakobs,

B Fowler, MA Clarke, G Hammersen, K Raab, J Kobori, A Moosa, B Vollmer, E Rossier, AK Iafolla, D Matern, OF Brouwer, J Finklestein, F Aksu, HP Weber, JM Bakkeren, JM Gabreels, D Bluestone, D Bluestone, TF Barron, P Beauvais, D Rabier, C Santos, R Umansky, W Lehnert: The clinical phenotype of succinic semialdehyde dehydrogenase deficiency (4-hydroxybutyric aciduria): case reports of 23 new patients. *Pediatrics* 99, 567-574 (1997)
DOI: 10.1542/peds.99.4.567

197. PL Pearl, KM Gibson, MT Acosta, LG Vezina, WH Theodore, MA Rogawski, EJ Novotny, A Gropman, JA Conry, GT Berry, M Tuchma: Clinical spectrum of succinic semialdehyde dehydrogenase deficiency. *Neurology 60*, 1413-1417 (2003)
DOI: 10.1212/01.WNL.0000059549.70717.80

198. C Jakobs, M Bojasch, E Monch, D Rating, H Siemes, F Hanefeld: Urinary excretion of gamma-hydroxybutyric acid in a patient with neurological abnormalities. The probability of a new inborn error of metabolism. *Clinica Chimica Acta* 111, 169-178 (1981)
DOI: 10.1016/0009-8981(81)90184-4

199. BM Hogema, M Gupta, H Senephansiri, TG Burlingame, M Taylor, C Jakobs, RB H Schutgens, W Froest, OC Snead, R Diaz-Arrastia, T Bottiglieri, M Grompe, KM Gibson: Pharmacologic rescue of lethal seizures in mice deficient in Succinate semialdehyde dehydrogenase. *Nature Genetics* 29, 212-216 (2001)
DOI: 10.1038/ng727

200. L Puttmann, H Stehr, M Garshasbi, K Hu, K Kahrizi, B Lipkowitz, P Jamali, A Tzschach, H Najmabadi, HH Ropers, L Musante, AW Kuss: A novel ALDH5A1 mutation is associated with succinic semialdehyde dehydrogenase deficiency and severe intellectual disability in an Iranian family. *American Journal of Medical Genetics* 161, 1925-1922 (2013)
[DOI NOT FOUND]

201. LJ Marcadier, AM Smith, MT Geraghty: Mutations in ALDH6A1 encoding Methylmalonate semialdehyde dehydrogenase are associated with Dysmyelination and transient methylmalonic aciduria. *Orphanet Journal of Rare Diseases* 8, 98 (2013)
DOI: 10.1186/1750-1172-8-98

202. PJ Congdon, D Haigh, R Smith, A Green, RJ Pollitt: Hypermethioninaemia and 3-hydroxyisobutyric aciduria in an apparently healthy baby. *Journal of Inherited Metabolic Disease* 4, 79-80 (1981)
DOI: 10.1007/BF02263600

203. LJ Mienie, E Erasmus: Biochemical studies on a patient with a possible 3-hydroxyisobutyrate dehydrogenase deficiency. *Fifth International Congress on Inborn Errors of Metabolism* (1990)
[DOI NOT FOUND]

204. FJ KO, WL Nyhan, J Wolf, B Barshop, L Sweetman: 3-Hydroxyisobutyric aciduria: an inborn error of valine metabolism. *Pediatric Research* 30, 322-326 (1991)
DOI: 10.1203/00006450-199110000-00006

205. RJ Pollitt, A Green, R Smith: Excessive excretion of beta-alanine and of 3-hydroxypropionic, R- and S-3-aminoisobutyric, R- and S-3-hydroxyisobutyric and S-2- (hydroxymethyl) butyric acids probably due to a defect in the metabolism of the corresponding malonicsemialdehydes. *Journal of Inherited Metabolic Disease* 8, 75-79 (1985)
DOI: 10.1007/BF01801669

206. P Baxter: Pyridoxine dependent and pyridoxine responsive seizures. *Developmental Medicine & Child Neurology* 43, 416-420 (2001)
DOI: 10.1017/S0012162201000779

207. SM Jr Gospe: Pyridoxine-dependent seizures: new genetic and biochemical clues to help with diagnosis and treatment. *Current Opinion in Neurology* 19, 148-153 (2006)
DOI: 10.1097/01.wco.0000218230.81301.12

208. P Baxter: Pyridoxine-dependent seizures: a clinical and biochemical conundrum. *Biochimica et Biophysica Acta* 1647, 36-41 (2003)
DOI: 10.1016/S1570-9639(03)00045-1

209. VC Daire, N Dagoneau, R Nabbout, L Burglen, C Penet, C Soufflet, I Desguerre, A Munnich, O Dulac: A gene for pyridoxine dependent epilepsy maps to chromosome 5q31. *American Journal of Human Genetics* 67, 991- 993 (2000)
DOI: 10.1086/303087

210. PB Mills, E Struys, C Jakobs, B Plecko, P Baxter, M Baumgartner, MA Willemsen, H Omran, U Tacke: Mutations in antiquitin in individuals with pyridoxine dependent seizures. *Nature Medicine* 12, 307-309 (2006)
DOI: 10.1038/nm1366

211. A Tlili, NH Hentati, A Gargouri, F Fakhfakh: Identification of a novel missense mutation in the ALDH7A1 gene in two unrelated Tunisian families with pyridoxine-dependent epilepsy. *Molecular Biology Reports* 40, 487 - 490 (2013)
DOI: 10.1007/s11033-012-2084-z

212. B Plecko, K Paul, E Paschke, SS Ipsiroglu, E Struys, C Jakobs, H Hartmann, T Luecke, M di Capua, C Korenke, C Hikel, E Reutershahn, M Freilinger, F Baumeister, F Bosch, W Erwa : Biochemical and Molecular Characterization of 18 Patients With Pyridoxine-Dependent Epilepsy and Mutations of the Antiquitin (ALDH7A1) Gene. *Human mutation* 28 (1), 19-26 (2007)
DOI: 10.1002/humu.20433

213. KT Rott, CA Agudelo: Gout. *The Journal of American Medical Association* 289, 2857-2860 (2003)
DOI: 10.1001/jama.289.21.2857

214. P Sulem, DF Gudbjartsson, G Walters, HT Helgadottir, A Helgason, SA Gudjonsson, C Zanon, S Besenbacher, G Bjornsdottir, OT Magnusson, G Magnusson, E Hjartarson, J Saemundsdottir, A Gylfason, A Jonasdottir, H Holm, A Karason, T Rafnar, H Stefansson, OA Andreassen, JH Pedersen, AI Pack, MC H de Visser, LA Kiemeney, AJ Geirsson, GI Eyjolfsson, I Olafsson, A Kong, G Masson, H Jonsson, U Thorsteinsdottir, I Jonsdottir, K Stefansson : Identification of low-frequency variants associated with gout and serum uric acid levels. *Nature Genetics* 43, 1127-1130 (2011)
DOI: 10.1038/ng.972

215. MC Hanna, C Blackstone: Interaction of the SPG21 protein ACP33/maspardin with the aldehyde dehydrogenase ALDH16A1. *Neurogenetics* 10, 217-228 (2009)
DOI: 10.1007/s10048-009-0172-6

216. V Vasiliou, M Sandoval, DS Backos, BC Jackson, Y Chen, P Reigan, MA Lanaspa, RJ Johnson, V Koppaka, DC Thompson: ALDH16A1 is a novel non-catalytic enzyme that may be involved in the etiology of gout via protein–protein interactions with HPRT1. *Chemico-Biological Interactions* 202, 22-31 (2013)
DOI: 10.1016/j.cbi.2012.12.018

217. LS Bicknell, J Pitt, S Aftimos, R Ramadas, MA Maw, SP Robertson: A missense mutation in ALDH18A1, encoding D1-pyrroline-5-carboxylate synthase (P5CS), causes an autosomal recessive neurocutaneous syndrome. *European Journal of Human Genetics* 16, 1176-1186 (2008)
DOI: 10.1038/ejhg.2008.91

Key Words: ALDH, SNPs, SAPs, KD4v, HOPE, Review

Send correspondence to: George Priya Doss. C, Medical Biotechnology Division, School of Biosciences and Technology, VIT University, Vellore 632014, Tamil Nadu, India, Tel: 91-9500040439, Fax: 0416-2243092, E-mail: georgepriyadoss@vit.ac.in