Identification of mutations in Malaysian patients with argininosuccinate lyase (ASL) deficiency

Ernie Zuraida Ali,⁎ Yusnita Yakob, Lock Hock Ngu,⁎

⁎ Inborn Error of Metabolism and Genetic Unit, Nutrition, Metabolism and Cardiovascular Research Centre, Institute for Medical Research, National Institutes of Health, Ministry of Health Malaysia, Section U13 Seria Alam, 40170 Shah Alam, Selangor, Malaysia

b Molecular Diagnostics and Protein Unit, Specialized Diagnostics Centre, Institute for Medical Research, National Institutes of Health, Ministry of Health Malaysia, Jalan Pahang, 50588 Kuala Lumpur, Malaysia

c Medical Genetics Department, Kuala Lumpur Hospital, Ministry of Health Malaysia, Jalan Pahang, 50588 Kuala Lumpur, Malaysia

ARTICLE INFO

Keywords:
Argininosuccinate lyase deficiency
ASLD
Autosomal recessive
Mutation
Hyperammonemia

ABSTRACT

Argininosuccinate lyase (ASL) deficiency impairs the function of the urea cycle that detoxifies blood ammonia in the body. Mutation that occurs in the ASL gene is the cause of occurrence of ASL deficiency (ASLD). This deficiency causes hyperammonemia, hepatopathy and neurodevelopmental delay in patients. In this study, the clinical characteristics and molecular analysis of 10 ASLD patients were presented. 8 patients were associated with severe neonatal onset, while the other 2 were associated with late onset. Molecular analysis of ASL gene identified four new missense variants, which were c.778C>T, p.(Leu260Arg), c.1340G>C, p.(Ser447Thr), c.436C>G, p.(Arg146Gly) and c.595C>G, p.(Leu199Val) and four reported missense variants, which were c.638G>A, p.(Arg213Gln); c.556C>T, p.(Arg186Trp), c.578G>A, p.(Arg193Gln) and c.436C>G, p.(Arg146Trp).

Insilico servers predicted all new and reported variants as disease-causing. Structural examination exhibited that all pathogenic variants affected the stability of the tetrameric ASL structure by disturbing the bonding pattern with the neighboring residues.

Conclusion: This study revealed the genetic heterogeneity among Malaysian ASL patients. This study has also expanded the mutational spectrum of the ASL.

1. Introduction

Argininosuccinate lyase deficiency (ASLD) is an autosomal recessive urea cycle disorder causing argininosuccinic aciduria. Argininosuccinate lyase (ASL; MIM 608310) is one of the six enzymes in the urea cycle that converts waste nitrogen into urea. ASL involves in the fourth step of the urea cycle by catalyzing argininosuccinic acid (ASA) to produce arginine and fumarate [1,2]. Mutations in the ASL gene result in defective cleavage of ASA leading to an accumulation of ASA in cells and an excessive excretion of ASA in urine. ASLD can manifest a wide clinical spectrum from asymptomatic to severe hyperammonemic neonatal onset life-threatening courses. The estimated incidence of ASLD is approximately one in 218,750 live births making it the second most common disorder in the urea cycle [3]. The diagnosis of ASLD is established based on elevation of plasma citrulline together with elevated argininosuccinic acid in the plasma or urine and can be confirmed by assay of ASL enzyme activity or molecular genetic testing of ASL gene, the latter of which offers additional advantages for genetic counselling of the family or prenatal testing in later pregnancies [4].

The human ASL gene is located on chromosome 7q11.21 [5], spans approximately 17 kb and comprises of 16 coding exons encoding 464 amino acids. ASL is a homotetramer of ~52kDa subunit and is considered as the mature enzyme carrying four identical active sites which provide binding pockets for ASA [6]. The enzyme is located in the cytosol and is expressed predominantly in liver [1]. However, it is also detected in many other tissues, including kidney [7], small intestine [8] and brain [9].

To date, there are 167 variants of the ASL gene reported by the Human Gene Mutation Database (HGMD) (http://www.hgmd.cf.ac.uk) [10], a database that provides information on DNA variants with current listing of 107 variants in its public site and 160 variants in the professional version 2018. Majority of variants are missense, but almost all types of other variants are also found. Even though > 160 ASL variants have been described, however the genotype-phenotype correlations are still not completely clear as standard biochemical testing is not sensitive enough to measure residual activity. Other possible factors
such as ASL tissue-specific expression [11], methylation status [12] and frequent alternative splicing variants [13] may influence the clinical phenotype. In this study, we present the molecular characterization of 10 patients affected with ASLD by analyzing the ASL gene and structure.

2. Materials and methods

2.1. Patients

This study includes samples from 10 patients from a total of 10 different families diagnosed with ASLD at Kuala Lumpur Hospital (HKL). All patients were identified to have ASLD according to clinical observations and metabolic findings. To confirm the diagnosis, molecular genetic testing was carried out on all patient samples, at Molecular Diagnostics and Protein Unit (UMDP), Institute for Medical Research. Once mutation was identified in the patients, clinicians would call parents to test the inheritance cases. Only parent samples from 5 patients were available and all parents were healthy. The control group was also included in this study which consisted of 50 normal healthy individuals. Informed consent was obtained from all patients. Peripheral blood (approximately 5 to 10 mL) was collected inside a standard EDTA tube. This study was performed according to the Declaration of Helsinki.

2.2. DNA extraction, PCR and sequencing

The genomic DNA from all blood samples were extracted by using the DNA Blood Mini Kit (QIAGEN, Hilden, Germany). The ASL gene including intron-exon boundaries [14] was amplified using the touch-down polymerase chain reaction (PCR) method. The amplified PCR product was run on 1.5% agarose gel under standard electrophoresis. The PCR product was purified by using the QiAquick PCR purification kit (QIAGEN, Hilden, Germany) and cycle sequencing was then carried out by using the Big Dye Terminator cycle sequencing version 3.1 (Applied Biosystems, Foster City, CA). The sequencing reaction was then electrophoresed by the Applied Biosystems 3500 Genetic Analyzer.

2.3. Mutation confirmation

SeqScape Software version 2.5 (Applied Biosystems, Foster City, CA) was used to identify the DNA variants, in which the sequencing results were aligned to the genomic and coding DNA reference sequence of ASL gene (NC_000007.13 and NM_000048.3). Human genome GRCh37/hg19 was used to retrieve all reference sequences. The single nucleotide variants (SNVs), insertion and deletion were reported by referring to the Human Genome Variation Society (HGVS) (http://www.hgvs.org/mutnomen).

Identified variants were compared against two public databases; HGMD [10] and 1000 Genome Project data (http://www.1000genomes.org) [15] in order to differentiate the one as reported or new variants. Meanwhile, the new variants were then compared with 50 healthy normal individuals (100 alleles) and Genome Aggregation Database (gnomAD) [16] in order to rule out a polymorphism.

2.4. In silico prediction

To validate the degree of conservation, HomoloGene program (http://www.ncbi.nlm.nih.gov/homologene) was used to run the multiple-sequence alignment. Meanwhile, the pathogenicity of missense variants was then predicted using four in silico web servers, FATHMM-XF [17], Mutation Taster2 [18], M-CAP [19] and PROTEAN [20]. Missense variants predicted as damaging by at least three or more servers were considered as disease-causing. In addition, I-Mutant 3.0 server [21], CUPSAT [22] and Site Directed Mutator (SDM) [23] were used to predict the stability/instability of the protein. Missense variants were considered as destabilizing in nature if two or more than two algorithms showed a decrease in stability upon mutation.

2.5. Modeling of native and mutant ASL structures

The effects of missense variants on the protein structure were then observed through computational modeling. The human ASL X-ray crystal structures were retrieved from the Protein Databank (RCSB-PDB) [24]. At the moment, two human ASL X-ray crystal structures were crystalized with resolutions 4.2 Å (PDB ID: 1AOS) [25] and 2.65 Å (PDB ID: 1K62) [26]. However, structure with resolution 4.2 Å consisted missing residues, while structure with resolution 2.65 Å had a mutation at residue 286. Therefore, the three dimensional (3D) ASL native model was built using SWISS-MODEL program [27] by submitting the targeted sequence (native) to the program. The ASL mutant model was generated via the Swiss-PDB viewer [28]. The ‘Mutate’ tool in the Swiss-PDB viewer was used to change the target amino acid in the ASL native model. Once the ASL native and mutant models were generated, energy minimizations were then performed using the same program to remove unwanted contacts. The GROMOS96 force field was used to assign the atomic charges to all the residues. The ASL native and mutant models were then visualized using PyMol software [29].

2.5.1. Native and mutant models validation

The qualities of the ASL native and mutant models were validated using ProCheck [30], ProSA-web [31] and ERRAT [32] programs. The Ramachandran plot generated by the ProCheck program was used to verify the overall stereochemical quality and accuracy of the predicted model. The geometry of the residues from the predicted model was compared with the X-ray crystal or NMR structures. The good quality of predicted model which covers the residues located in the most favorable and additional allowed region should be > 90%. The ProSA-web used molecular mechanics force field [31] to calculate energy profiles (z-score) for the modeled structures. The z-score predicts overall model quality and measures the total energy deviation of the structure using random conformations. The modeled structure is predicted to be erroneous if the z-scores range beyond the characteristic of reference proteins. Meanwhile, for better interpretation of the predicted protein, the z-score plot can be used to check whether the z-score of the predicted protein is within the range of scores typically found for proteins of similar groups. ERRAT was used to inspect the correctness of the overall fold/structure and faults over localized regions. ERRAT score provides the whole quality factor for non-bonded atomic interactions. Score > 50% indicates the higher quality model. The quality of the ASL native and mutant models were also verified by calculating the root mean square deviation (RMSD) using the Swiss-PDB viewer. The RMSD of the native model was calculated by superimposing the native model with the template structure (PDB ID: 1K62). Meanwhile for the mutant models, the RMSDs were calculated by superimposing the mutant models with the native model.

2.5.2. Hydrogen bond analysis

The hydrogen bond interaction between native and neighboring residues was analyzed using Chimera software [33]. The relax hydrogen bond constraint was set to be within 0.4 Å with a maximum angle of 30 degree.

3. Results

3.1. Clinical findings

Ten patients (4 males and 6 females) from ten families were studied. All patients except one (Patient P7) were from Malay ethnicity background. Table 1 presents the main clinical features of these patients. Eight patients have had their first presentation during neonatal period,
Table 1
Summary of clinical and biochemical features of argininosuccinate lyase deficiency (ASLD) patients.

| Patient | Sex | Initial presentation | Blood ammonia, μmol/L, N.R. < 51) | Therapy received | Outcome | Family history |
|---------|-----|----------------------|------------------------------------|------------------|---------|----------------|
| P1      | F   | 25d Poor suck, lethargy, myoclonus seizures, acute change in sensorium | 264                  | SPI, PFN, SB, SPB, L-Arg | Survived | Spastic tetraplegia, severe psychomotor retardation, frequent seizure episodes despite a low-protein diet and treatment with oral ammonia scavengers, fatal encephalopathy at 21 y | CP, one affected elder sibling |
| P2      | F   | 10d Recurrent vomiting, difficulty in breathing, acute encephalopathy | 210                  | SPI, PFN, SB, SPB, L-Arg | Survived | MLD, infrequent decompensation on a low-protein diet and oral ammonia scavengers. Current age is 24y. | NCP, no other affected family member |
| P3      | F   | 5d Vomiting, lethargy, breathing abnormalities | 274                  | Protein restriction, SB, L-Arg | Survived | MLD, no acute decompensation on a low-protein diet and oral ammonia scavengers. Current age is 24y. | NCP, one affected elder sibling died at 9y (undiagnosed) |
| P4      | M   | 5d Feeding intolerance, irritability, altered sensorium | 501                  | SPI, PFN, SB, SPB, L-Arg, DL, MV | Survived | MDD/LD, recurrent decompensation during first 2 years. Current age is 8y | NCP, no other affected family member |
| P5      | M   | 4d Feeding intolerance, irritability, progressing to somnolence | 446                  | SPI, PFN, SB, SPB, L-Arg, MV | Survived | MLD, infrequent decompensation on a low-protein diet and oral ammonia scavengers. Current age is 11y. | NCP, one affected elder sibling died at 3m |
| P6      | M   | 5d Feeding refusal, vomiting, lethargy, breathing abnormalities, progressing to coma | 510                  | SPI, PFN, SB, SPB, L-Arg, DL, MV | Survived | SLD, recurrent decompensation during first 5 years. Current age is 14y | NCP, no other affected family member |
| P7      | M   | 5y Drowsy, seizures, feeding refusal preceded by a febrile illness | ND                   | SPI, PFN | Survived | MLD, no further decompensation on a low-protein diet. Current age is 11y. | NCP, no other affected family member |
| P8      | F   | 3d Abnormal breathing, progressive acute encephalopathy | 305                  | SPI, PFN, SB, SPB, L-Arg, MV | Survived | MLD, infrequent decompensation on a low-protein diet and oral ammonia scavengers. Current age is 12y. | NCP, no other affected family member |
| P9      | F   | 7d Vomiting, lethargy, seizures, progressing to coma | 447                  | SPI, PFN, SB, SPB, L-Arg, MV | Survived | SDD, recurrent decompensation on a low-protein diet and oral ammonia scavengers. Died at 6y. | NCP, no other affected family member |
| P10     | F   | 6y Learning disabilities, hyperactivity | 78                   | Protein restriction, SB, L-Arg | Survived | MLD, no acute decompensation on a low-protein diet and oral ammonia scavengers. Current age is 12y. | NCP, no other affected family member |

M male, F female, y years, m month, d days, SPI Stop protein intake, PFN Parenteral fluid and nutrition, L-Arg L-arginine, SB sodium benzoate, SPB sodium phenylbutyrate, DL dialysis, MV Mechanical ventilation, MLD Mild learning disability, MDD/LD Moderate developmental delay/learning disability, SDD Severe learning disability, SDD Severe developmental delay, ND not done, CP Consanguineous parents, NCP Non-consanguineous parents.
| Patient | Mutation state | Exon/ intron | Mutation | Type of mutation | Gene conservation in eukaryota (%) | Deleterious or damaging prediction | Stability prediction | Screening with 100 normal alleles | Population allele frequency from gnomAD database | Reference |
|---------|----------------|--------------|----------|------------------|------------------------------------|-----------------------------------|---------------------|---------------------------------|-----------------------------------------------|-----------|
|         |                |              | Genome number | Coding DNA number | Protein number | FATHMM-XF | Mutation taster2 | M-CAP |                             |                  |           |
| P1      | HM             | Ex-8         | g.65552356G > A | c.638G > A | p.(Arg213Gln) | M | 100 | D | DC | PP |                     |
| P2      | HM             | Ex-8         | g.65552356G > A | c.638G > A | p.(Arg213Gln) | M | 100 | D | DC | PP |                     |
| P3      | HM             | Ex-8         | g.65552356G > A | c.638G > A | p.(Arg213Gln) | M | 100 | D | DC | PP |                     |
| P4      | HT             | Ex-8         | g.65552356G > A | c.638G > A | p.(Arg213Gln) | M | 100 | D | DC | PP |                     |
| P5      | HM             | Ex-7         | g.65551762C > T | c.556C > T | p.(Arg186Trp) | M | 100 | D | DC | PP |                     |
| P6      | HM             | Ex-7         | g.65551762C > T | c.556C > T | p.(Arg186Trp) | M | 100 | D | DC | PP |                     |
| P7      | CHT            | Ex-7         | g.65551784G > A | c.578G > A | p.(Arg193Gln) | M | 100 | D | DC | PP |                     |
| P8      | CHT            | Ex-10        | g.65553853C > T | c.778C > T | p.(Leu260Phe) | M | 100 | D | DC | PP |                     |
| P9      | HM             | Ex-5         | g.65548151C > T | c.436C > T | p.(Arg146Trp) | M | 90 | D | DC | PP |                     |
| P10     | CHT            | Ex-7         | g.65551801C > G | c.595C > G | p.(Leu199Val) | M | 100 | D | DC | PP |                     |

**Table 2**

Mutations identified, *in-silico* servers prediction and population frequency of the alleles in ASLD patients.

---

*HM* homozygous, *HT* heterozygous, *CHT* compound heterozygous, *M* missense, – not calculable, *D* damaging, *DC* disease causing, *PP* possibly pathogenic, *Del* deleterious, *Inc* increase, *Dec* decrease, *DS* destabilizing, *S* stabilizing, *RS* reduce stabilizing, *IS* increase stabilizing, *A* absence, *P* presence.
with blood ammonia ranging from 210 μmol/L to 510 μmol/L and required emergency treatment. Two patients, P7 and P10 presented at later age, at 5 years and 6 years old respectively. Despite dietary protein restriction and ammonia scavengers therapy, all patients had impaired cognitive function: 6 with mild learning disabilities, 1 with moderate restriction and ammonia scavengers therapy, all patients had impaired laterage, at 5 years and 6 years old respectively. Despite dietary protein required emergency treatment. Two patients, P7 and P10 presented at later age, at 5 years and 6 years old respectively. Despite dietary protein restriction and ammonia scavengers therapy, all patients had impaired cognitive function: 6 with mild learning disabilities, 1 with moderate restriction and ammonia scavengers therapy, all patients had impaired laterage, at 5 years and 6 years old respectively. Despite dietary protein restriction and ammonia scavengers therapy, all patients had impaired cognitive function: 6 with mild learning disabilities, 1 with moderate restriction and ammonia scavengers therapy, all patients had impaired cognitive function: 6 with mild learning disabilities, 1 with moderate

### 3.2. Molecular diagnosis

**Table 2** shows direct sequencing result of ASL gene in the ten patients. Six (60%) patients were in homozygous state, three (30%) patients were in compound heterozygous state and one (10%) patient was in heterozygous state (Table 2). Eight missense variants (c.638G > A; c.556C > T, c.578G > A, c.778C > T, c.436C > T, c.1340G > C, c.436C > G, c.595C > G) were identified in these ten patients. Four missense variants (c.778C > T, c.1340G > C, c.436C > G and c.595C > G) were detected as new variant, in which these variants were not found previously in HGMD and 1000 Genome Project. The remaining four missense variants (c.638G > A; c.556C > T, c.578G > A, c.436C > G) were previously reported in literature. The presence of all missense variants was confirmed in parental DNA. Only parent samples from Patients 4, 5, 7, 8 and 10 were available to test for inheritance cases (Table 1). All parents carried the variant in heterozygous form. In order to define pathogenicity of new missense variants; p.(Leu260Phe), p.(Ser447Thr), p.(Arg146Gly) and p. (Leu199Val), each of the variants was compared with 50 healthy normal individuals and also with gnomAD database. None of them were discovered in the healthy individuals as well as in gnomAD database.

**Table 3** Evaluation of ASL native and mutant models by using PROCHECK, ProSA-Web and ERRAT programs.

| Patient | Exon | Mutation | Ramachandran plot statistics (%) | ProSA | Error (%) | RMSD (Å) |
|---------|------|----------|----------------------------------|------|----------|--------|
|         |      | Coding DNA number | Protein number | Most favoured | Additionally allowed | Generously allowed | Disallowed | z-Score | Overall quality factor Score |
| Native  | Ex-8 | c.638G > A | p.(Arg213Gln) | 94.4 | 5.1 | 0.5 | 0.0 | -8.61 | 95.37 | 0.05 |
| P1      | Ex-5 | c.556C > T | p.(Arg186Trp) | 94.2 | 5.3 | 0.5 | 0.0 | -8.58 | 93.06 | 0.05 |
| P2      | Ex-7 | c.578G > A | p.(Arg193Gln) | 94.4 | 5.1 | 0.5 | 0.0 | -8.69 | 95.37 | 0.05 |
| P3      | Ex-10 | c.778C > T | p.(Leu260Phe) | 94.2 | 5.3 | 0.5 | 0.0 | -8.66 | 93.52 | 0.05 |
| P4      | Ex-15 | c.436C > T | p.(Arg146Gly) | 94.2 | 5.3 | 0.5 | 0.0 | -8.63 | 93.06 | 0.05 |
| P5      | Ex-7 | c.556C > G | p.(Arg186Trp) | 94.2 | 5.3 | 0.5 | 0.0 | -8.66 | 93.06 | 0.05 |
| P6      | Ex-7 | c.578G > T | p.(Leu199Val) | 94.2 | 5.3 | 0.5 | 0.0 | -8.66 | 93.06 | 0.05 |
| P7      | Ex-5 | c.436C > G | p.(Arg146Gly) | 94.2 | 5.3 | 0.5 | 0.0 | -8.66 | 93.06 | 0.05 |
| P8      | Ex-7 | c.556C > T | p.(Arg186Trp) | 94.2 | 5.3 | 0.5 | 0.0 | -8.66 | 93.06 | 0.05 |
| P9      | Ex-7 | c.595C > G | p.(Leu199Val) | 94.2 | 5.3 | 0.5 | 0.0 | -8.66 | 93.06 | 0.05 |

3.4. Structural analysis

3.4.1. Quality of ASL native and mutant models

The effects of variants on the ASL structure were then further analyzed. The ASL native structure was developed from the known structures of ASL (PDB ID: 1K62 and 1AOS) using the program SWISS-MODEL. The mutant models were developed using the Swiss-PDB viewer. Energy minimization was then performed to the ASL native and mutant models. The quality of ASL native and mutant models was verified with PROCHECK. The Ramachandran plot determined by PROCHECK revealed the residues of the ASL native and mutant models in the most favorable regions were > 90%, as shown in Table 3. None of residues was detected in disallowed regions of the Ramachandran plot of the native and mutant models (Supplementary Fig. S1A). Scores produced by the PROCHECK were closed to the 100%, which showed the stereo-chemical quality of the ASL native and mutant models was in a good quality. The ProSA-web and ERRAT programs were also used to further check the quality of the ASL native and mutant models. The z-score generated from ProSA-web exhibited the overall model quality of the ASL native and mutant models was negative, as shown in Table 3 and Supplementary Fig. S1B. For the ASL native model, the value was within the range of z-score found for protein of similar groups, indicating the quality of this model was high. Meanwhile, for the mutant models, the values were within the range of the score of the native protein, suggested that all mutant models were in very good overall configuration. The ERRAT program exhibited the overall quality score for non-bonded atomic interactions of ASL native and mutant models was > 90%, as shown in Table 3. None of the residues were above the 99% cut off of error-value (Supplementary Fig. S1C). The ERRAT score exhibited the whole quality factor for the native and mutant models was greater than accepted range (50%), which indicated all of them were within the range of a high-quality model.

The RMSD calculation was also carried out to verify the quality of ASL native and mutant models. The superimposition of the native model onto the template structure (PDB ID: 1K62) produced RMSD of 0.05 Å. Meanwhile, for the mutant model, Table 3 shows RMSD values of the mutant models when superimposed onto the ASL native model.

3.4.2. Position of mutations and hydrogen bond interactions

**Fig. 1** shows position of the reported and new variants in the ASL tetramer structure. The missense variants were mapped only to subunit A of tetramer ASL structure. All reported; p.(Arg213Gln), p.(Arg186Trp), p.(Arg193Gln) and p.(Arg146Gly) and three new; p. (Leu260Phe), p.(Arg146Gly) and p.(Leu199Val) variants were located in the tetramer structure. The missense variants were mapped only to subunit A of tetramer ASL structure. All reported; p.(Arg213Gln), p.(Arg186Trp), p.(Arg193Gln) and p.(Arg146Gly) and three new; p. (Leu260Phe), p.(Arg146Gly) and p.(Leu199Val) variants were located in the tetramer structure.
at domain 2, while only one variant; p.(Ser447Thr) was located at domain 3.

Table 4 shows all the hydrogen bond interactions between the native and neighboring residues, and between the mutated and neighboring residues. Information in this table includes the hydrogen bond interactions of the reported and new variants. Fig. 2a shows the substitution of arginine to glutamine affected the bonding patterns with the surrounding residues. Arginine (Arg213) in native structure was found to form hydrogen bond interactions with neighboring residues in subunit A, Leu199, Thr226, Arg217, Pro198 and Leu197. This residue was also found to form hydrogen bond interaction with neighboring residue Leu440 in subunit D. However, loss of hydrogen bond interactions with residues Thr226, Pro198 and Leu197 in subunit A, and residue L440 in subunit D was observed when arginine substituted to glycine at position 213.

Loss of hydrogen bond interactions with residue Asp183 in subunit A, and residues Glu241 and Asp237 in subunit D was also observed when the arginine substituted to tryptophan at position 186, as shown in Fig. 2b. The substituted Trp186 was found to form hydrogen bond interactions with residues Arg182, Val190 and Glu189 in subunit A. These residues (Arg182, Val190 and Glu189) were also found to form hydrogen bond interactions with Arg186 in the native structure. Residue Arg193 in native structure formed hydrogen bond interactions with residues Arg122, Glu189, Val190 and Glu241 in subunit A (Fig. 2c). Concurrently, this residue also formed interactions with residues in subunit D (Glu185 and Glu189). However, substitution to glutamine caused loss of interactions with all residues in subunit D, as shown in Fig. 2c. For mutation occurred at position 260, which changed from leucine to phenylalanine, the mutated residue (Phe260) was observed not to affect the bonding pattern with surrounding residues, as shown in Fig. 2d.

Fig. 2e shows arginine at position 146 in the native structure formed interactions with residues in subunit A, Thr142, Glu150, Ile350, Leu353 and Ala149. Nevertheless, the substituted tryptophan at this position affected the interactions with Glu150, Ile350 and Leu353. Meanwhile, the replacement of serine to threonine at position 447 caused loss of interactions with residue Gly441 (Fig. 2f). Instead, the mutated residue Thr447 formed new interaction with residue Ala439. The substituted glycine at position 146 (Gly146) was found to form similar bonding pattern, as presented in Trp146 mutant (Fig. 2g). For mutation occurred at position 199, which changed from leucine to valine, the mutated residue (Val199) was observed not to affect the bonding patterns with the surrounding residues, as shown in Fig. 2h.

4. Discussion

This study presented the first molecular analysis of ASLD in Malaysia population. Ten patients were identified to carry a missense variant in ASL gene with different clinical and biochemical phenotype, indicating the heterogeneous distribution of pathogenic variants in ASL gene. Although the splice site, deletion and insertion (indel) were not found in this study, the potential of cryptic splice sites and RNA decay also contributed for the severity of ASLD, as reported in other studies [34–36]. Currently, about 167 variants have been discovered in the HGMD, with the largest proportion are missense variants [10]. None of the reported and new missense variants obtained in this study were discovered at the active site of ASL protein. The new missense variants (Leu260Phe, Ser447Thr, Arg146Gly and Leu199Val) were detected as very rare variants as all of them were not discovered in our normal control, gnomAD, HGMD and 1000 Genome Project databases. Two new missense variants found in this study, the potential of cryptic splice sites and RNA decay also contributed for the severity of ASLD, as reported in other studies [34–36]. Currently, about 167 variants have been discovered in the HGMD, with the largest proportion are missense variants [10]. None of the reported and new missense variants obtained in this study were discovered at the active site of ASL protein. The new missense variants (Leu260Phe, Ser447Thr, Arg146Gly and Leu199Val) were detected as very rare variants as all of them were not discovered in our normal control, gnomAD, HGMD and 1000 Genome Project databases. Two new missense variants found in this study and located at positions 146 and 447 were basically a substitution of another amino acid at the exact site. The missense variants reported at these positions were found to be pathogenic [34].

In this study, four unrelated patients (Patients 1, 2, 3 and 4) shared a similar genotype, a Leu213Gln. This pathogenic variant was previously reported in patient associated with severe neonatal onset form in United State (US) population [34]. However, there is no detailed information on the functional study for this pathogenic variant. Our analysis showed although one in silico server (I-Mutant 3.0) predicted the Gln213 mutant increased the stability of the ASL structure, however, the other two in silico servers predicted to affect the stability of the ASL structure. Analysis of hydrogen bonding patterns further supported the mutated residue (Gln213) caused loss of four hydrogen bond interactions in the ASL structure, including hydrogen bond interactions with the adjacent subunit. This would suggest that this pathogenic variant is predicted to highly likely affect the stability of the ASL structure. Furthermore, this pathogenic variant occurs in highly conserved region, implying the pathogenicity of this variant. The Arg213Gln mutant is suggested to have loss of protein function as three
of our patients presented in severe neonatal disease phenotype. The same position was also affected with pathogenic variant, Arg213*, which was reported in patient associated with severe neonatal onset form in Italian population [37] and unknown phenotype in Indian and US populations [34].

Although Patient 4 had only 1 variant detected, nevertheless, metabolic/biochemical testing had established the diagnosis of ASLD. Therefore, the second disease-causing variant that could not be identified in this study might be due to undetectable by sequence analysis such as intragenic deletions or duplications. Hence, incorporating methods such as multiplex ligation-dependent probe amplification (MLPA), a gene-targeted microarray, quantitative PCR or long-range PCR could improve diagnostic yield.

Mutation of Arg186Trp was found as homozygous form in two

| Patient | Exon | Mutation | Hydrogen bond interactions |
|---------|------|----------|---------------------------|
| P1 | Ex-8 | c.638G > A p.(Arg213Gln) | R213-A-NE: L199-A-O |
| P2 | Ex-8 | c.638G > A p.(Arg213Gln) | R213-A-NH1: T226-A-O |
| P3 | Ex-8 | c.638G > A p.(Arg213Gln) | R213-A-NH2: L199-A-O |
| P4 | Ex-8 | c.638G > A p.(Arg213Gln) | R217-A-N: R213-A-O |
| P5 | Ex-7 | c.556C > T p.(Arg186Trp) | R186-A-N: R182-A-O |
| P6 | Ex-7 | c.556C > T p.(Arg186Trp) | R186-A-NE: D183-A-OD2 |
| P7 | Ex-7 | c.556C > T p.(Arg186Trp) | R186-A-NH1: E241-A-OE2 |
| P8 | Ex-5 | c.436C > T p.(Arg146Trp) | R146-A-N: T142-A-O |
| P9 | Ex-5 | c.436C > T p.(Arg146Trp) | R146-A-NH1: L353-A-O |
| P10 | Ex-7 | c.556C > T p.(Arg186Trp) | R186-A-NE: E241-A-OE2 |
| Ex-15 | c.778C > T p.(Leu260Phe) | R186-A-NH2: D237-D-OD1 |
| Ex-15 | c.778C > T p.(Leu260Phe) | V190-A-N: V186-A-O |
| Ex-10 | c.778C > T p.(Leu260Phe) | E189-A-N: W186-A-O |
| P11 | Ex-7 | c.595C > G p.(Leu199Val) | L199-A-N: T119-A-OG1 |
| P11 | Ex-7 | c.595C > G p.(Leu199Val) | R213-A-NE: L199-A-O |
| P11 | Ex-7 | c.595C > G p.(Leu199Val) | R213-A-NH2: L199-A-O |

Table 4

Hydrogen bond interactions between native and neighboring residues and between mutant and neighboring residues.

| Hydrogen bond interactions between native and neighboring residue | Hydrogen bond interactions between mutant and neighboring residue |
|---------------------------------------------------------------|---------------------------------------------------------------|
| Residue-chain-Atom                                            | Residue-chain-Atom                                            |
| P1 Ex-8 c.638G > A p.(Arg213Gln) R213-A-NE: L199-A-O           | Q213-A-N: E219-A-O                                           |
| P2 Ex-8 c.638G > A p.(Arg213Gln) R213-A-NH1: T226-A-O          | R217-A-N: Q213-A-O                                           |
| P3 Ex-8 c.638G > A p.(Arg213Gln) R213-A-NH2: L199-A-O          | R217-A-N: Q213-A-O                                           |
| P4 Ex-8 c.638G > A p.(Arg213Gln) R217-A-N: R213-A-O            | R217-A-N: Q213-A-O                                           |
| P5 Ex-7 c.556C > T p.(Arg186Trp) R186-A-N: R182-A-O            | W186-A-N: R182-A-O                                           |
| P6 Ex-7 c.556C > T p.(Arg186Trp) R186-A-NE: D183-A-OD2         | V190-A-N: W186-A-O                                           |
| P7 Ex-7 c.556C > T p.(Arg186Trp) R186-A-NH1: E241-A-OE2        | E189-A-N: W186-A-O                                           |
| P8 Ex-5 c.436C > T p.(Arg146Trp) R146-A-N: T142-A-O            | W146-A-N: T142-A-O                                           |
| P9 Ex-5 c.436C > T p.(Arg146Trp) R146-A-NH1: L353-A-O          | E150-A-N: W146-A-O                                           |
| P10 Ex-7 c.556C > T p.(Arg186Trp) R186-A-NE: E241-A-OE2        | A149-A-N: W146-A-O                                           |
| Ex-10 c.778C > T p.(Leu260Phe) L260-A-N: M256-A-O              | F260-A-N: M256-A-O                                           |
| Ex-10 c.778C > T p.(Leu260Phe) C264-A-N: L260-A-O              | F260-A-N: M256-A-O                                           |
| P9 Ex-5 c.436C > T p.(Arg146Trp) R146-A-NH2: L353-A-O          | F260-A-N: A257-A-O                                           |
| Ex-15 c.1340G > C p.(Ser447Thr) G441-A-N: S447-A-OG            | W146-A-N: T142-A-O                                           |
| Ex-15 c.1340G > C p.(Ser447Thr) Q451-A-N: S447-A-OG            | E150-A-N: W146-A-O                                           |
| P10 Ex-7 c.556C > T p.(Arg186Trp) R186-A-NH2: L353-A-O         | A149-A-N: G146-A-O                                           |
| Ex-7 c.595C > G p.(Leu199Val) L199-A-N: T119-A-OG1             | V199-A-N: T119-A-OG1                                         |
| Ex-7 c.595C > G p.(Leu199Val) R213-A-NE: L199-A-O              | R213-A-NE: V199-A-O                                         |
| Ex-7 c.595C > G p.(Leu199Val) R213-A-NH2: L199-A-O             | R213-A-NH2: V199-A-O                                         |
unrelated patients (Patients 5 and 6). Both patients showed severe clinical and biochemical phenotype, which were presented with acute neonatal symptom. Therefore, our finding is similar with the previous reported case, in which changes to this arginine at this position are predicted to be deleterious and may affect the stability of the tetramer ASL structure [34,38]. This pathogenic variant was commonly reported in the Saudi Arabia [38] and Turkish populations [34].

Patient 7 had late-onset disease. This patient had compound heterozygous for new (Leu260Phe) and reported (Arg193Gln) pathogenic variants. This new pathogenic variant (Leu260Phe) was observed to not affecting the surrounding hydrogen bond interactions. However, this new pathogenic variant was discovered as disease-causing by in silico servers. In addition, this pathogenic variant occurs in highly conserved region, strengthening the pathogenicity of this variant. Meanwhile, the Arg193Gln was previously reported in patient associated with neonatal onset form in US and Austrian populations, and patient with the late onset form in Italian population [34,39]. However, there is no functional study for this variant [34]. Our structural analysis discovered this reported pathogenic variant affected the hydrogen bond interactions with the adjacent subunit, which indirectly disturbing the formation of tetrameric structure of ASL. Therefore, this would suggest that this pathogenic variant might contribute to the severity of this deficiency in

![Fig. 2. Hydrogen bond interactions of native and mutants with neighboring residues. Native residue is presented in orange stick and sphere, while mutant residue is presented in purple stick and sphere. Hydrogen bonds are indicated by black dotted lines. Subunits A, B, C and D are presented in green, cyan, magenta and yellow cartoons, respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)(Click to enlarge)](image-url)
Patient 8 was associated with severe neonatal onset form of disease. This patient had a compound heterozygous for new (Ser447Thr) and reported (Arg146Trp) pathogenic variants. The reported pathogenic variant, Arg146Trp was previously reported in US and Turkish populations in the literature [34]. However, there is no information stated about the functional study for this variant. In our structural analysis, both pathogenic variants (Ser447Thr and Arg146Trp) affected the ASL structure by altering the bonding patterns with the neighboring residues. In addition, in silico servers predicted both pathogenic variants as disease-causing. Consistent with this, both pathogenic variants were suggested to cause loss of protein function and resulted in a severe form in our patient.

A new variant in homozygous state (Arg146Gly) was observed in Patient 9, which had hyperammonemia episodes at day 7. This new pathogenic variant was predicted as disease-causing by in silico servers. Furthermore, the substitution of long side chain of arginine to small glycine has disturbed the bonding pattern of ASL structure. This indicates the impact of this new pathogenic variant to the stability of the ASL structure, which possibly causing the ASLD. The same position was also affected with pathogenic variant (Arg146Trp), which was reported in our study, US and Turkish populations [34]. Another pathogenic variant (Arg146Gln) located at the same position was previously reported in homozygous and heterozygous states in UK population [35].

A new variant, Leu199Val was observed not to affect the surrounding hydrogen bond interactions of ASL structure in Patient 10. However, this pathogenic variant was found as compound heterozygous which formed along with the reported pathogenic variant, Arg186Trp. As mentioned above, the reported pathogenic variant, Arg186Trp was discovered as recurrent pathogenic variant disturbing the stability of the tetramer ASL structure [34,38]. This indicates that this reported pathogenic variant might participate to the severity of disease in this patient. Furthermore, in silico analysis exhibited the new pathogenic variant, Leu199Val as disease-causing. The impacts of both pathogenic variants were in line with the clinical phenotype of this patient, in which Patient 10 had a late-onset disease.

5. Conclusion

In conclusion, this study has characterized the clinical and molecular aspects of argininosuccinate aciduria in Malaysia patients. Eight different missense variants from 10 patients were identified. All the new and reported pathogenic variants discovered in this study may potentially disturb the ASL structure. This study has expanded the mutation spectrum of ASL mutations. Establishing a precise molecular diagnosis is beneficial not only for accurate genetic counselling but also for accurate carrier testing as early intervention is important to minimize disease progression.

Declaration of Competing Interest

All the authors declared that they have no conflicts of interest to this work.

Acknowledgment

We thank the Director General of Health, Malaysia for the permission to publish this paper. Our special thanks to Director of IMR, Head Center of Nutrition, Metabolism and Cardiovascular Research and Head Unit of Inborn Error of Metabolism and Genetics for careful reading of this manuscript. We would like to extend our gratitude to Mohd Khairul Nizam Mohd Khalid and Nural Farahana Rosli for excellent technical assistance. We are also grateful to all clinicians for their contributions in this study. This study was funded by Diagnostics Services, Institute for Medical Research, Ministry of Health, Malaysia.
D.S. Goodsell, J.D. Westbrook, J. Woo, J. Young, C. Zardecki, H.M. Berman, P.E. Bourne, S.K. Burley, The RCSB protein data Bank: views of structural biology for basic and applied research and education, Nucleic Acids Res. 43 (2015) D345–D356. https://doi.org/10.1093/nar/gku1214.

[25] L.M. Sampaleanu, P.W. Codding, V.D. Lobsanov, M. Tsai, G.D. Smith, C. Horvatin, P.L. Howell, Structural studies of duck delta2 crystallin mutants provide insight into the role of Thr161 and the 280s loop in catalysis, Biochem. J. 384 (2004) 437–447. https://doi.org/10.1042/BJ20040656.

[26] M.A. Turner, A. Simpson, R.R. McInnes, P.L. Howell, Human argininosuccinate lyase: a structural basis for intragenic complementation, Proc. Natl. Acad. Sci. U. S. A. 94 (1997) 9063–9068.

[27] A. Waterhouse, M. Bertoni, S. Bienert, G. Studer, G. Tauriello, R. Gumhold, F.T. Heer, T.A. De Beer, C. Rempfer, L. Bordoli, R. Lepore, T. Schwede, SWISS-MODEL: homology modelling of protein structures and complexes, Nucleic Acids Res. 46 (2018) W296–W303. https://doi.org/10.1093/nar/gky427.

[28] W.L. DeLano, The PyMOL Molecular Graphics System, Delano Sci, San Carlos, 2002.

[29] P. Rostami, J. Häberle, A. Setoudeh, J. Zachoecke, F. Sayarifard, Two novel mutations in the argininosuccinate lyase gene in Iranian patients and literature review, Iran. J. Pediatr. 27 (2017) 26–28. https://doi.org/10.5812/ijp.7666.

[30] E.Trevisson, A. Burlina, M. Doimo, V. Pertegato, A. Casarin, L. Cesaro, P. Navas, G. Basso, G. Sartori, L. Salvianti, Functional complementation in yeast allows molecular characterization of missense argininosuccinate lyase mutations, J. Biol. Chem. 284 (2009) 28926–28934. https://doi.org/10.1074/jbc.M109.050495.

[31] M. Wiederstein, M.J. Sippl, ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins, Nucleic Acids Res. 35 (2007) W407–W410. https://doi.org/10.1093/nar/gkm290.

[32] C. Colovos, T. O Yeates, Verification of protein structures: patterns of nonbonded atomic interactions, Protein Sci. 2 (1993) 1511–1519.