Received: 30 May 2023 | Revised: 9 October 2023 | Accepted: 10 October 2023

DOI: 10.1002/jimd.12687

ORIGINAL ARTICLE

JIMD SSIEM WILEY

Datamining approaches for examining the low prevalence of N-acetylglutamate synthase deficiency and understanding transcriptional regulation of urea cycle genes

Ljubica Caldovic[1,2] | **Julie J. Ahn**[3] | **Jacklyn Andricovic**[3] | **Veronica M. Balick**[4] | **Mallory Brayer**[5] | **Pamela A. Chansky**[6] | **Tyson Dawson**[6,7] | **Alex C. Edwards**[6,8] | **Sara E. Felsen**[6,8] | **Karim Ismat**[1,2] | **Sveta V. Jagannathan**[6] | **Brendan T. Mann**[9] | **Jacob A. Medina**[6] | **Toshio Morizono**[10] | **Michio Morizono**[10] | **Shatha Salameh**[11,12] | **Neerja Vashist**[1,2] | **Emily C. Williams**[3,13] | **Zhe Zhou**[14] | **Hiroki Morizono**[1,2]

[1]Center for Genetic Medicine Research, Children's National Research Institute, Children's National Hospital, Washington, DC, USA

[2]Department of Genomics and Precision Medicine, School of Medicine and Health Sciences, The George Washington University, Washington, DC, USA

[3]Department of Anatomy and Cell Biology, The George Washington University School of Medicine and Health Sciences, Washington, DC, USA

[4]Department of Biochemistry and Molecular Medicine, The George Washington University School of Medicine and Health Sciences, Washington, DC, USA

[5]Department of Biological Sciences, The George Washington University, Washington, DC, USA

[6]The Institute for Biomedical Science, School of Medicine and Health Sciences, George Washington University, Washington, DC, USA

[7]AMPEL BioSolutions LLC, Charlottesville, Virginia, USA

[8]Center for Neuroscience Research, Children's National Research Institute, Children's National Hospital, Washington, DC, USA

[9]Department of Microbiology, Immunology, and Tropical Medicine, School of Medicine and Health Sciences, George Washington University, Washington, DC, USA

[10]College of Science and Engineering, University of Minnesota, Minneapolis, Minnesota, USA

[11]Department of Pharmacology & Physiology, School of Medicine and Health Sciences, George Washington University, Washington, DC, USA

[12]Sheikh Zayed Institute for Pediatric Surgical Innovation, Children's National Hospital, Washington, DC, USA

[13]The George Washington University Cancer Center, School of Medicine and Health Sciences, George Washington University, Washington, DC, USA

[14]Department of Civil and Environmental Engineering, The George Washington University, Washington, DC, USA

Correspondence
Hiroki Morizono, Center for Genetic Medicine Research, Children's National Research Institute, 1 Inventa Pl, Silver Spring, MD 20910, USA.
Email: hmorizono@childrensnational.org

Abstract

Ammonia, which is toxic to the brain, is converted into non-toxic urea, through a pathway of six enzymatically catalyzed steps known as the urea cycle. In this pathway, N-acetylglutamate synthase (NAGS, EC 2.3.1.1) catalyzes the formation of N-acetylglutamate (NAG) from glutamate and acetyl

Funding information
Public Health Service, Grant/Award Number: K01DK076846; National Institute of Diabetes Digestive and Kidney Diseases; National Institutes of Health; Recordati Rare Disease

Communicating Editor: Manuel Schiff

coenzyme A. *NAGS* deficiency (NAGSD) is the rarest of the urea cycle disorders, yet is unique in that ureagenesis can be restored with the drug N-carbamylglutamate (NCG). We investigated whether the rarity of NAGSD could be due to low sequence variation in the NAGS genomic region, high NAGS tolerance for amino acid replacements, and alternative sources of NAG and NCG in the body. We also evaluated whether the small genomic footprint of the NAGS catalytic domain might play a role. The small number of patients diagnosed with NAGSD could result from the absence of specific disease biomarkers and/or short NAGS catalytic domain. We screened for sequence variants in *NAGS* regulatory regions in patients suspected of having NAGSD and found a novel *NAGS* regulatory element in the first intron of the *NAGS* gene. We applied the same datamining approach to identify regulatory elements in the remaining urea cycle genes. In addition to the known promoters and enhancers of each gene, we identified several novel regulatory elements in their upstream regions and first introns. The identification of *cis*-regulatory elements of urea cycle genes and their associated transcription factors holds promise for uncovering shared mechanisms governing urea cycle gene expression and potentially leading to new treatments for urea cycle disorders.

KEYWORDS
AMPK, NAGS, NAGS deficiency, nitrogen load, transcriptional regulation, urea cycle

1 | INTRODUCTION

Ammonia, a neurotoxic product of protein and nucleic acid catabolism is converted in the liver into non-toxic urea by the urea cycle using six enzymes and two mitochondrial solute carriers. The conversion starts in the mitochondria where carbamylphosphate synthetase 1 (CPS1) and ornithine transcarbamylase (OTC) catalyze the formation of citrulline from ammonia, bicarbonate, and ornithine.[1] Citrulline is then transported to the cytoplasm by ornithine transporter (ORNT) which is encoded by the *SLC25A15* gene.[1] Argininosuccinate synthase 1 (ASS1) catalyzes formation of argininosuccinate from citrulline and aspartate, which is transported from mitochondria to cytoplasm by *SLC25A13*, also known as citrin or ARALAR2.[1] Argininosuccinate lyase (ASL) and arginase 1 (ARG1) convert argininosuccinate into urea and ornithine, which is transported to mitochondria by ORNT to be a substrate of OTC.[1] N-acetylglutamate (NAG), produced by NAG synthase (NAGS) is an essential allosteric activator of CPS1.[2,3] Genetic defects in any of the urea cycle enzymes and transporters can result in hyperammonemia, which if untreated can cause irreversible brain injury and death.[1] The X-linked *OTC* deficiency affects approx. 1:60 000–70 000 people[4–7] and is the most prevalent urea cycle defect (UCD). The prevalence of *ASS1* and *ASL* deficiencies, which have autosomal recessive inheritance, is estimated to be approx. 1:200 000.[7] The true prevalence of NAGS and ARG1 deficiencies are unknown and have been estimated to be 1:950 000 and less than 1:2000 000, respectively, based on the number of reported cases.[7] Early interventions to reduce blood ammonia concentration minimize hyperammonemic brain damage. Therefore, swiftly diagnosing every patient with a UCD is critical for good patient outcomes.[1,8]

Diagnosing every patient with *NAGS* deficiency is especially important since it is the only UCD where a single drug, N-carbamylglutamate (NCG), can restore ureagenesis.[9,10] Normalization of the blood ammonia concentration upon administration of NCG can be used to distinguish *NAGS* deficiency from other UCDs.[8,11] The absence of specific biochemical markers of NAGS deficiency makes the diagnosis of the disease challenging and heavily dependent on DNA sequencing.[12] For most cases of *NAGS* deficiency, the disease is caused by pathogenic sequence variants in the exons and splice sites.[12] The small genomic footprint of the *NAGS* gene permitted examination of non-coding regions for pathogenic sequence variants well before whole genome sequencing became widely available. This led to discovery of eight deleterious sequence variants in the *NAGS* splicing regions and *cis*-acting regulatory elements (cCRE).[13–16] Although ureagenesis in patients with NAGSD can be restored with a single drug, because it is the rarest UCD, we wanted to determine if genetic or biochemical factors contribute to the low prevalence of the disease or

whether an absence of specific biomarkers make diagnosis of the disease difficult.

Patients with clinical and biochemical symptoms of *OTC*, *CPS1*, *ASL*, and *SLC25A13* (citrin) deficiency for whom pathogenic sequence variants in the coding regions and canonical splice sites cannot be found[17–25] are likely to have the disease due to genetic defects in the non-coding regions of those genes. In some of these patients, deep intronic sequence variants that affect mRNA splicing have been found through sequencing of cDNA isolated from their fibroblasts.[26] Others have pathogenic sequence variants in the cCREs that were identified using either whole genome sequencing[13] or targeted sequencing of known regulatory elements.[13,16,27,28] Identification and functional testing of pathogenic sequence variants in the regulatory regions of urea cycle genes requires an understanding of their transcriptional regulation. Transcription factors that bind promoters and enhancers of most urea cycle genes have been identified using reporter gene and DNA binding assays, transgenic and knockout animals (Table 1 and references therein). More recently, a novel regulatory element in the first intron of the *NAGS* gene has been identified through data mining of the ENCODE project results.[13] Therefore, we queried the ENCODE database for transcription factors that bind known and predicted regulatory elements of urea cycle genes in the human liver. Knowledge of the regulation of expression of urea cycle genes will aid in diagnosing patients with non-coding pathogenic sequence variants.

The transcription and translation of urea cycle genes are known to be coordinately regulated upon changes in protein catabolism and ammonia production (nitrogen load) such as amount of dietary protein intake and/or altered cellular protein degradation due to illness.[29–35] Tabulation of transcription factors known to regulate expression of urea cycle genes (Table 1) does not easily yield a molecular mechanism for their coordinated expression. From the ENCODE database, we identified several transcription factors that bind to regulatory elements of all eight urea cycle genes that may be factors in the molecular mechanism responsible for coordinated expression of urea cycle genes.

2 | RESULTS AND DISCUSSION

2.1 | Pathogenic and common sequence variants in the *NAGS* Gene

NAGS deficiency is an extremely rare inborn error of metabolism. To date, 105 known cases have been reported. Sequencing of the *NAGS* gene has been used to diagnose *NAGS* deficiency in 78 patients from 58 families (Table S1). Additional three patients with NAGS deficiency were reported in 2012 but their genotypes and the method used to diagnose the disease were not described.[36] NAGS deficiency has been reported in an additional 20 patients from nine families before a conclusive molecular diagnostic test became available.[37–47] Of the 57 sequence variants found in patients with NAGS deficiency, 12 were small insertions and deletions that disrupted the *NAGS* reading frame and caused premature termination of NAGS translation, three were mutations that affected splicing, two affected base pairs in the vicinity of canonical splice sites, 6 were in the *NAGS* regulatory elements, four were nonsense, and 29 were missense variants (Table S1). The majority (69%) of the patients with NAGS deficiency are homozygous and 31 had neonatal onset disease (Table S1). Clinical and biochemical symptoms of 98 patients with NAGS deficiency have been

TABLE 1 Transcription factors that regulate expression of urea cycle genes in adult liver cells.

		Transcription factors		
Gene	**Transcription initiation**	**Promoter and proximal enhancer**	**Distal enhancer**	**References**
NAGS	SP1	CREB, FXR	HNF-1, NF-Y	[14,109,146]
CPS1	TATA	GR, C/EBP	C/EBP, GR, HNF-3, P3	[111–121]
OTC	TATA	HNF-4, COUP-TF, SP1, C/EBP	C/EBP, HNF-4, COUP-TF	[124–128,147]
ASS	SP1	AP2	CREB	[110,148,149]
ASL	SP1	NF-Y		[150]
Arg1	SP1	C/EBP, NF-Y, NF-1	C/EBP, P1, P2, NF-Y	[98,128,151,152]
ORNT	–[a]	–	–	–
Citrin	SP1	USF1, HNF3b	–	[122,123]

Note: Involvement of transcription factors in regulation of urea cycle genes was examined using reporter gene assays, DNA binding assays, and transgenic and gene knockout animals.

[a]Querying PubMed and Google Scholar with ([ornithine transporter OR SLC25A15] AND [enhancer OR promoter]) yielded no results.

reviewed recently.[12] Query of the gnomAD database[48] revealed four common sequence variants (MAF≥1%) in the coding region, introns, and 3′-untranslated region of the human *NAGS* gene (Table S2).

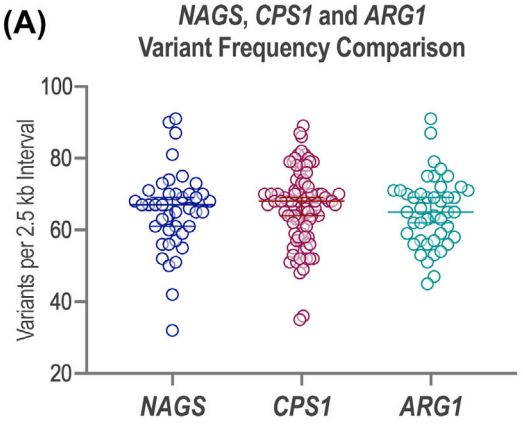
Since *NAGS* deficiency is the only UCD where ureagenesis can be restored with a single drug, we considered factors contributing to the small number of diagnosed patients compared to other UCDs. We hypothesized that these factors might include genetic and biochemical properties of the human *NAGS* gene and enzyme such as natural variation of the *NAGS* locus, tolerance of NAGS to amino acid substitutions, the amount of NAG sufficient for effective ureagenesis, and/or alternative sources of either NAG or NCG. At the same time, we could not rule out the small number of diagnosed patients with *NAGS* deficiency could be due to the absence of specific biomarkers of the disease and clinical symptoms such as nausea, vomiting, lethargy, irritability, and ataxia present in hyperammonemia, which is rare, but also occur in more frequently seen conditions.[1,12]

2.2 | Natural variation in the *NAGS* genomic region

Human NAGS gene and/or the surrounding genomic region may be more resistant to DNA sequence changes than other genomic regions resulting in fewer NAGS sequence variants. We tested this hypothesis in two ways: by comparing frequencies of sequence variants in genomic regions harboring *NAGS*, *CPS1*, and *ARG1* genes and by comparing the frequency of synonymous, missense, and nonsense variants in *NAGS*, *CPS1*, and *ARG1* coding regions in individuals without known genetic diseases. The analysis focused on these three urea cycle genes because their only known function is in the urea cycle and because of their autosomal locations.[49,50] To compare the frequency of sequence variants in genomic regions harboring *NAGS*, *CPS1*, and *ARG1*, we obtained data from the 1000 Genomes Project[51] using the chromosomal locations of sequence variants within human *NAGS*, *CPS1*, *ARG1*, and 50 kb flanking each gene (Table S3). Each genomic region was then divided into 2500 bp intervals followed by counting the number of sequence variants in each 2500 bp interval (Figure S1). The frequencies of sequence variants in genomic regions harboring *NAGS*, *CPS1*, and *ARG1*, reported as the number of sequence variants in each 2500 bp interval along the three genomic regions, were similar (Figures 1A and S1).

To compare natural variation within *NAGS*, *CPS1*, and *ARG1* coding regions, we obtained lists of synonymous, missense, and nonsense variants in each protein from gnomAD[48] (Tables S4–S6) and mapped their positions on the protein sequences of NAGS, CPS1, and ARG1 (Figures S2–S4). The frequency of synonymous, missense, and nonsense variants, calculated as the number of variants per 100 amino acids were similar for NAGS, CPS1, and ARG1 proteins (Figures 1B and S2–S4). The distribution of synonymous and missense variants was uniform across the entire length of all three proteins (Figures S2–S4). Additionally, the proportion of NAGS, CPS1, and ARG1 amino acid positions affected by synonymous, missense and nonsense variants were similar (Figures 1B and S2–S4).

Given similar rates of natural variation in the *NAGS*, *CPS1*, and *ARG1*, an explanation for the small number of patients with NAGS deficiency could be the small size of the NAGS domain responsible for substrate binding and catalysis.[52] In addition to the mitochondrial targeting signal and variable segment, NAGS has two structural domains, the N-acetyltransferase (NAT) domain that binds substrates and catalyzes NAG formation and the amino acid kinase (AAK) domain that binds NAGS allosteric activator arginine.[15,52–54] The mitochondrial targeting signal, variable segment, and AAK domain map to



(B)

	NAGS	CPS1	ARG1
Number of exons	7	38	8
Protein length	534	1500	332
Missense variants	262	745	186
Synonymous variants	152	341	72
Nonsense variants	14	16	5
Variants per 100 amino acids:			
• missense	49	50	56
• synonymous	28	23	22
Fraction of amino acids affected by variants (%)	74.5	71.4	75.1

FIGURE 1 Natural variation of *NAGS*, *CPS1*, and *ARG1*. (A) Number of sequence variants in 2500 bp intervals within genomic regions harboring *NAGS*, *CPS1*, and *ARG1* genes. (B) Comparison of the numbers of sequence variants in the coding regions of *NAGS*, *CPS1*, and *ARG1*.

residues 1–372 or 69.5% of the NAGS sequence while the NAT domain maps to residues 376–534 or 29.5% of the NAGS sequence.[52,55] In gnomAD, missense variants are uniformly distributed across the length of NAGS; 177 (68%) missense variants map to the mitochondrial targeting sequence, variable segment, and AAK domain, while 84 (32%) map to the NAT domain.

The NAT domain is disproportionately affected by pathogenic missense variants[12,52]; it harbors 14 of the 28 missense variants found in patients with *NAGS* deficiency (Table S1). Variants p.L442V, p.W484R, p.S410P, and p.R414P affect residues that bind NAGS substrates and catalyze NAG formation.[56] Variants p.G457D and p.L430P reduced solubility of the recombinant bacterial NAGS homolog[52] while p.T431I and p.R509Q variants reduced enzymatic activity and increased the K_m for glutamate in recombinant human NAGS.[57] Four of the remaining six variants (p.L319R, p.S398C, p.E433S, and p.E433G) affect amino acids conserved in eukaryotic and vertebrate-like bacterial NAGS, while the remaining two (p.Y512C, and p.A518T) affect residues that are not highly conserved.[58] Variant p.E360D in the AAK domain affects allosteric regulation of NAGS due to reduced bunding of arginine,[15,54,59] the p.350I variant increased stability while variants p.V173E, p.P260L, and p.I291L reduced solubility and/or stability of the bacterial NAGS homolog.[52] Compared to NAGS, the catalytic and regulatory domains of CPS1 and ARG1 are 1084 and 322 amino acids long, respectively.[60,61] The larger sizes of CPS1 and ARG1 catalytic domains than the NAGS NAT domain and similar rates of natural variation of *NAGS*, *CPS1*, and *ARG1* could be sufficient to explain the low prevalence of *NAGS* deficiency compared to *CPS1* and *ARG1* deficiencies.

2.3 | Only select *NAGS* variants result in *NAGS* deficiency

Since NAG is an essential allosteric activator of CPS1, a small amount could suffice to activate CPS1 and initiate efficient ureagenesis.[3] In vitro biochemical characterization of mutant NAGS proteins found in patients with *NAGS* deficiency revealed that residual NAGS activity of 5% can result in milder, late-onset disease in some individuals.[62] Given that such low enzyme activity can result in late-onset disease suggests that NAGS variants retaining significantly greater specific activity may not result in a noticeable disease phenotype. We addressed this by examining the NAGS tolerance to amino acid substitutions and potential for alternative sources of either NAG or NCG in the body.

Most amino acid replacements in human NAGS may not result in a marked reduction of its enzymatic activity and NAGS deficiency because the NAGS protein fold can tolerate substitutions of amino acids. A corollary to this would be an expected lower conservation of NAGS compared to CPS, OTC, and arginase across phyla. We compared the degree of protein sequence conservation of NAGS, CPS, OTC, and arginase because their only known functions in humans are in the urea cycle and arginine biosynthesis.[50] Protein sequences of NAGS, CPS, OTC, and arginase were collected from 58 species of bacteria, fungi, plants, invertebrates, and vertebrates (Table S7) and the percent sequence identity was calculated for all pairs of protein sequences and visualized as heatmaps (Figure 2). Bacterial species that were included in the analysis were chosen because their NAGS proteins were either characterized biochemically[59,63–65] or have known three-dimensional structure[66–68] although the genomes of *Neisseria gonorrhoeae*, *Escherichia coli*, *Pseudomonas aeruginosa*, and *Ralstonia eutropha* do not encode an arginase gene. The overall sequence identity among NAGS proteins across a wide range of phyla was lower than sequence identities between either carbamylphosphate synthetase, OTC, or arginase proteins from the same organisms (Figure 2). Conservation of carbamylphosphate synthetase, OTC, and arginase oligomerization states across phyla could contribute to the higher degree of their sequence conservation. Carbamylphosphate synthetase from *E. coli* and human CPS1 can form dimers through their allosteric domains[60,69]; mammalian and bacterial anabolic OTC are trimers[69–73]; eukaryotic and some bacterial arginases are trimers,[61,74,75] arginase from *Helicobacter pylori* is a monomer[76] while arginases from hyperthermophiles form hexamers.[77,78] However, neither *H. pylori* nor hyperthermophiles with hexameric arginases have genes with sequence similarity to *E. coli* N-acetylglutamate synthase. The NAGS protein fold appears to be more tolerant to changes in amino acid sequence and oligomerization states than carbamylphosphate synthetase, OTC, and arginase since NAGS monomers from different organisms have similar three-dimensional structures with only about 20% sequence identity and can form either hexamers or tetramers.[66–68,79] This tolerance of NAGS structure to amino acid substitutions suggests that many NAGS variants may have sufficient residual activity to avoid hyperammonemia.

The incidence of NAGSD is similar to incidences of citrin and ORNT deficiencies,[7] two conditions where other amino acid transporters may compensate for the defects in the two transporters. The earliest studies of the urea cycle postulated that NCG was essential for urea production.[80,81] Subsequent studies identified N-acetylaspartate (NAA) as an activator of CPS1[82] and NAG as the essential allosteric activator of CPS1.[3] If another enzyme could catalyze the formation of either NAG or NCG in hepatocytes, perhaps as a moonlighting function, NAGS deficiency would manifest in humans only if both enzymatic

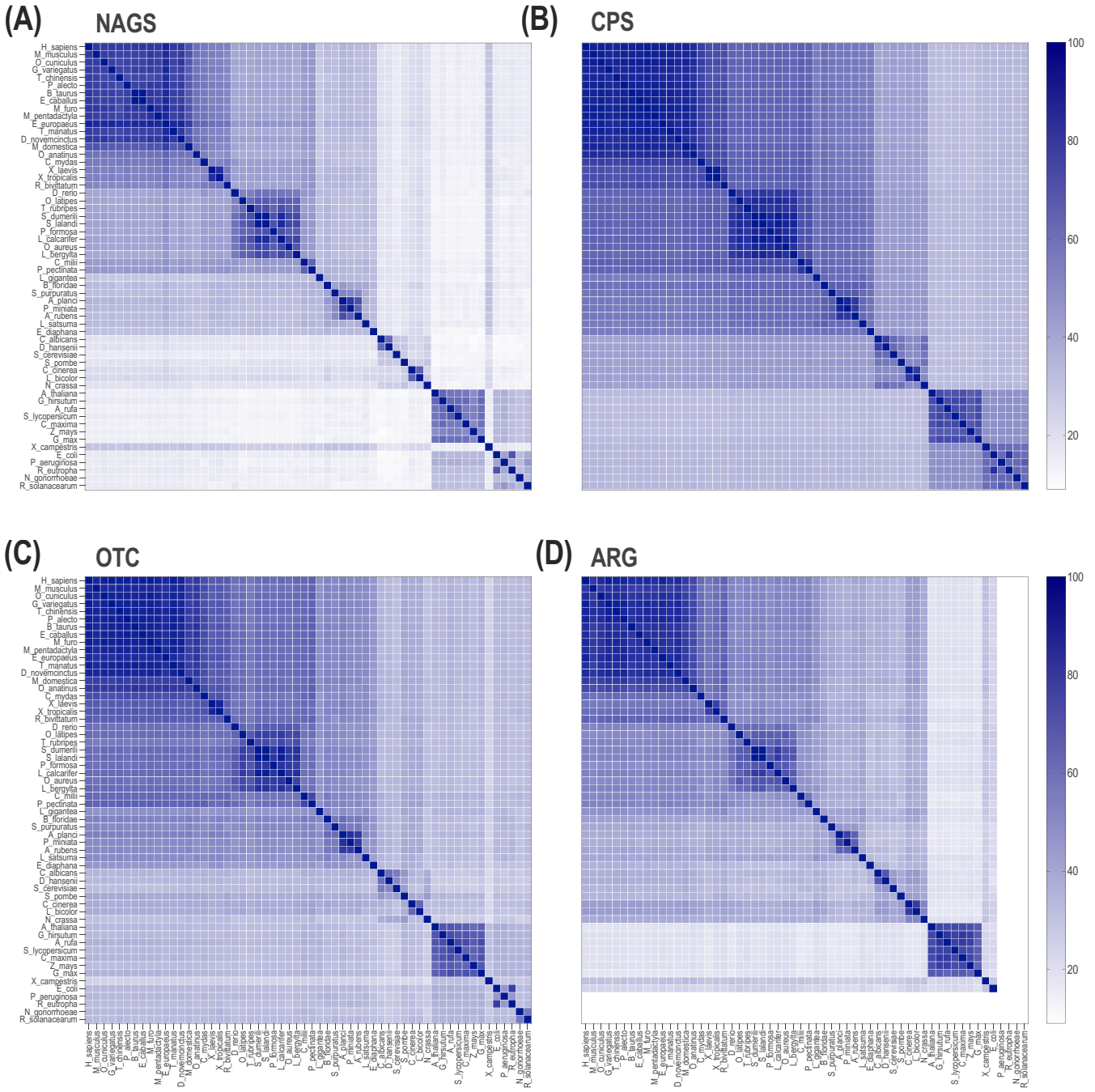


FIGURE 2 Conservation of NAGS (A), carbamylphosphate synthetase (CPS; B), OTC (C) and arginase (ARG; D) protein sequences in 58 species of mammals, reptiles, amphibians, fish, invertebrates, fungi, plants, and bacteria. The species names are listed along the *x*- and *y*-axes as well as in Table S7. Genomes of *E. coli*, *P. aeruginosa*, *R. eutropha*, and *N. gonorrhoeae* do not have arginase genes.

activities are reduced or absent. While it is possible that humans possess a second enzyme with the ability to catalyze the formation of allosteric activators of CPS1 that is not present in mice, this is unlikely given the similar phenotype of other urea cycle disorders in humans and mice and the Mendelian segregation of NAGS knockout allele in mice.[83–85] NAA and NAG are brain metabolites[86,87] that have been detected in the blood.[88,89] Therefore, it is possible that circulating NAA and NAG can enter hepatic mitochondria and activate CPS1, although cytoplasmic deacetylases[90] may decrease the efficiency of this process.

In some bacteria, NCG can be formed by L-hydantoinase as an intermediate of histidine catabolism[91] (Figure 3). NCG has been detected in the metabolomes of the human gut and oral microbiota.[92–95] Therefore, it is possible that NCG could

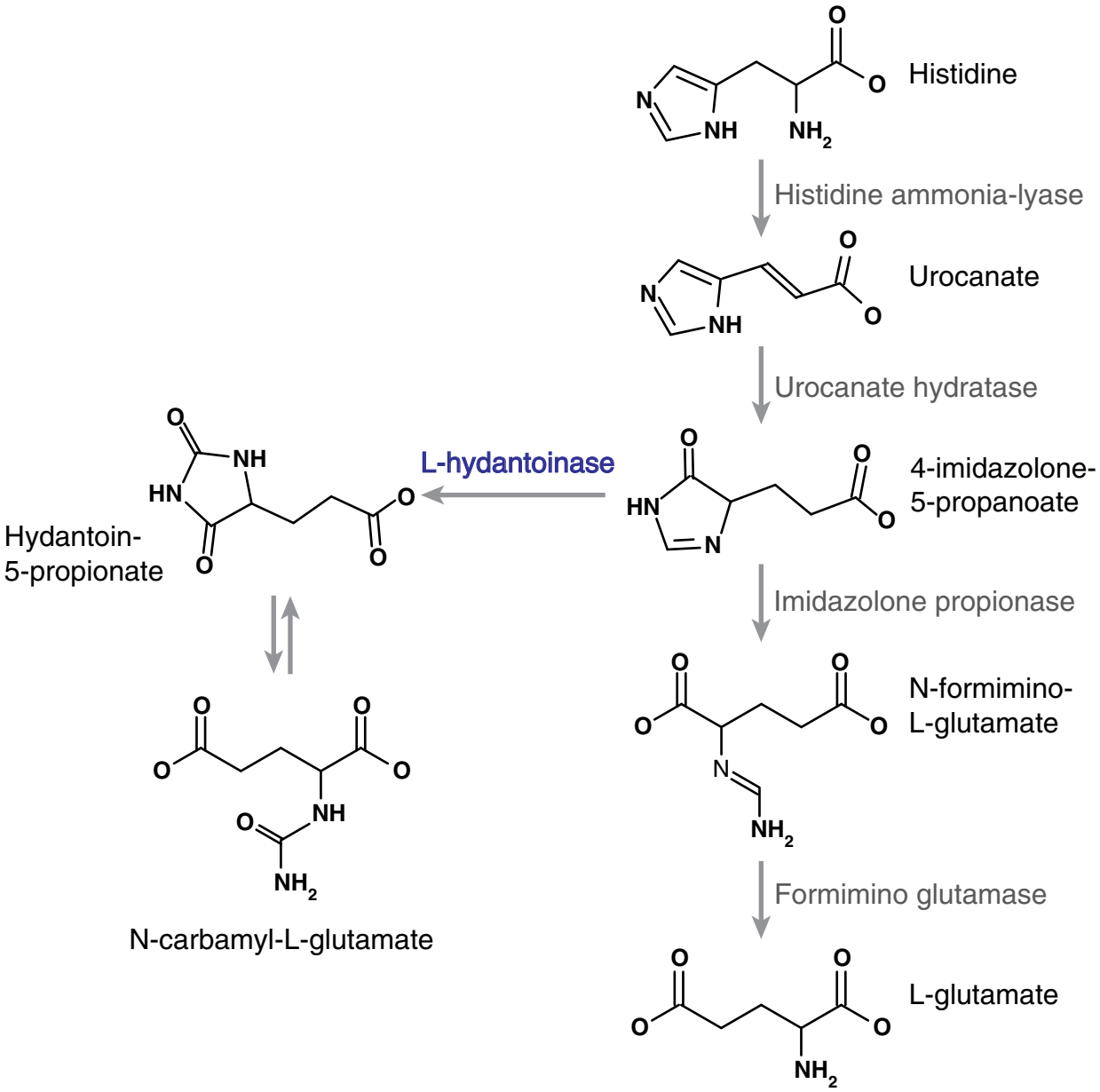


FIGURE 3 N-carbamyl-L-glutamate as a product of histidine catabolism in bacteria. Enzyme names are shown in gray and blue.

be produced by the human microbiota, absorbed, and transported to the liver to activate CPS1.

2.4 | Under-diagnosis of *NAGS* deficiency

While structural and biochemical properties of NAGS protein could contribute to the small number of patients diagnosed with NAGS deficiency, we hypothesize that the difficulty of diagnosis due to the absence of specific biochemical markers may be the main reason for the observed low incidence of the disease. Therefore, increasing awareness that nausea, vomiting, lethargy, irritability, and ataxia are clinical symptoms of hyperammonemia together with increased genetic testing could lead to an increase in diagnosis of all urea cycle disorders including *NAGS* deficiency. Although most of the pathogenic *NAGS* sequence variants reside in the coding region and splice sites, 10% have been found in the NAGS regulatory regions through a combination of targeted sequencing and data mining.[13,14,16] Two NAGS regulatory elements, located 3 kb upstream of the *NAGS* transcription start site and in the first intron of *NAGS*, have been identified based on the presence of pathogenic sequence variants in patients with *NAGS* deficiency.[13,14]

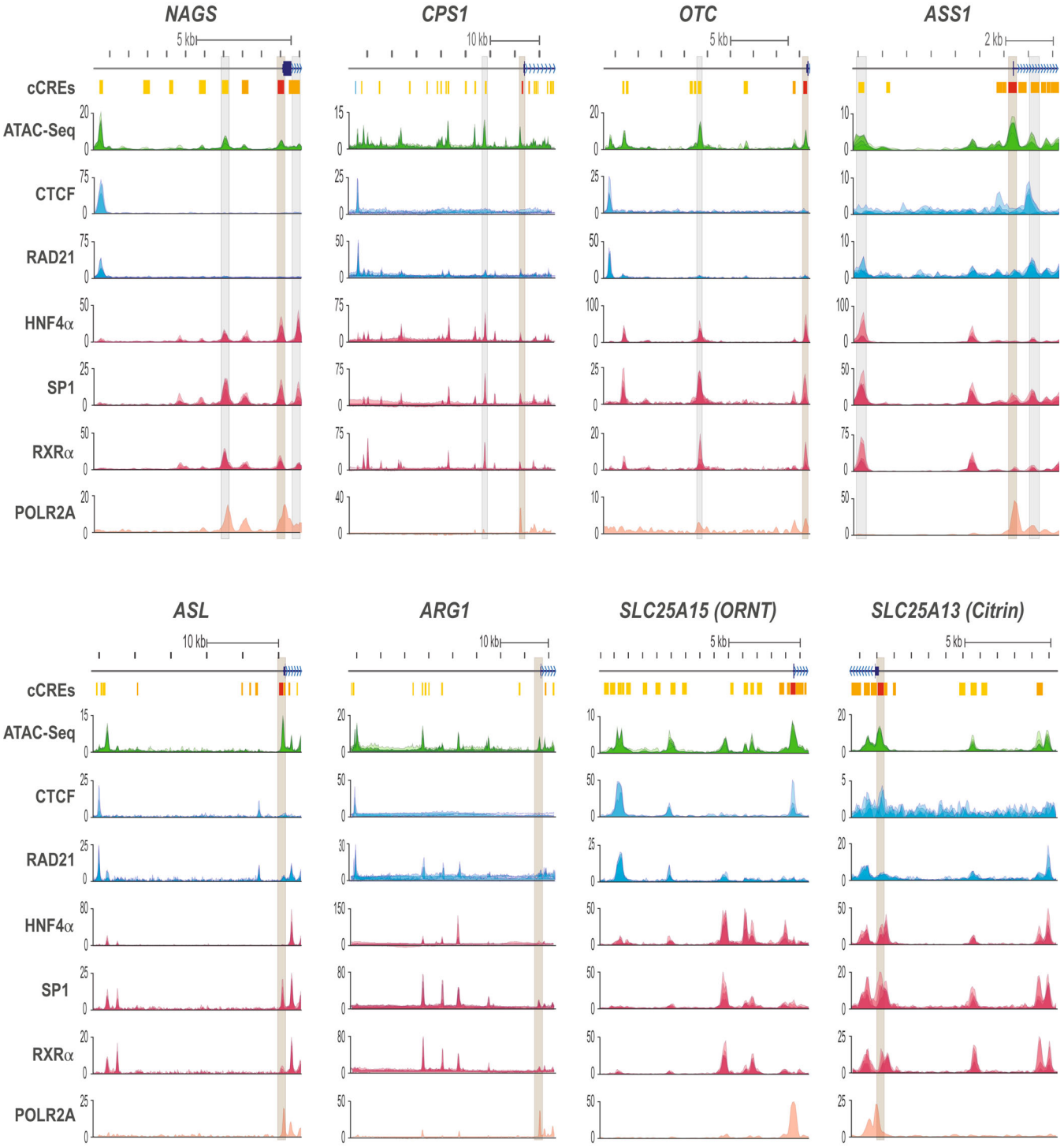


FIGURE 4 Transcriptional regulation of urea cycle genes in human liver tissue. Chromatin accessibility (ATAC-Seq), binding of cohesion complex subunits CTCF and RAD21, binding of transcription factors HNF4α, SP1, and RXRα, and binding of the 2A subunit of RNA polymerase II (POLR2A) are shown for the following genomic regions: *NAGS* (chr17:43994838-44 005 554); *CPS1* (chr2:210520320-210 563 758); *OTC* (chrX:38335687-38 353 060); *ASS1* (chr9:130438404-130 446 866); *ASL* (chr7:66049813-66 078 206); *ARG1* (chr6:131533519-131 576 359); *SLC25A15* (chr13:40776255-40 790 513); *SLC25A13* (chr7:96320513-96 332 077). Characterized regulatory elements: promoters—tan; enhancers—gray. Predicted cCREs: promoters—red, proximal enhancers—orange, distal enhancers—yellow, CTCF binding sites—blue.

However, patients with *NAGS* deficiency do exist with only a single mutant *NAGS* allele (Table S1). While it is possible that sequence variants in these patients have a dominant negative effect on NAGS function,[52] the presence of pathogenic sequence variants in yet-to-be-discovered regulatory elements cannot be ruled out. Specific epigenetic

histone modifications and binding of transcription factors are hallmarks of the *cis*-acting gene regulatory elements identified in the ENCODE project.[96] Since the ENCODE project has been instrumental in rationalizing a functional basis for several pathogenic non-coding sequence variants found in patients with *NAGS* and *OTC* deficiencies, we sought to identify novel regulatory elements in all eight urea cycle genes by data mining of the ENCODE database.

2.5 | Datamining ENCODE to identify novel regulatory elements of urea cycle genes

Since a combination of DNA sequencing and datamining approaches led to identification of two NAGS regulatory elements, we queried the UCSC Genome Browser for candidate cCREs predicted to regulate urea cycle gene expression and the ENCODE database for transcription factors that bind to regulatory elements of urea cycle genes. Our goals were to identify novel regulatory elements of urea cycle genes and transcription factors that govern coordinated expression of urea cycle genes in response to changes in protein catabolism.

Identification of cCREs in the ENCODE project was based on epigenetic data from cell lines and human tissues such as the locations of DNase hypersensitive sites, DNA methylation and histone modifications.[97] Some of the cCREs mapped to known, experimentally identified promoters and enhancers of *NAGS*, *CPS1*, *OTC*, *ASS1*, *ASL*, and *SLC25A13* (citrin) genes (Figure 4, gray and tan highlights and references in Table 1). Interestingly, none of the cCREs mapped to the *ARG1* promoter (Figure 4). This could be because expression of *ARG1* is liver specific[98,99] and the location of the *ARG1* promoter could not be predicted using epigenetic data from cell lines and tissues that do not express the gene. Next, we queried the ENCODE database for transcription factor binding and chromatin accessibility in the regulatory regions of urea cycle genes in human liver tissue. DNA binding data were available for 16 transcription factors, CTCF and RAD21 chromatin modifiers, and RNA polymerase subunit 2A (POLR2A) (Table 2 and Figures 4 and S5–S12). We also collected DNase-Seq and ATAC-Seq data about chromatin accessibility in the upstream regulatory regions of urea cycle genes in the liver. DNase-Seq and ATAC-Seq peaks coincided with some, but not all, cCREs (Figures 4 and S5–S12). The cCREs that do not coincide with DNase-Seq and ATAC-Seq peaks may regulate expression of urea cycle genes in other tissues or indicate aberrant expression of urea cycle genes in cancer cells and cell lines.[100–106]

Chromatin modifiers CTCF and RAD21 are components of the cohesion complex that binds chromatin insulators to regulate chromatin organization into topologically associating domains[107,108] (TADs). Chromatin insulators and TADs prevent interactions between promoters and distant regulatory elements from a different gene.[107,108] CTCF and RAD21 bind to predicted chromatin insulator upstream of the *CPS1* gene and to predicted enhancers in the upstream regions of other urea cycle genes (Figures 4 and S6). It is possible that cCREs that bind CTCF and RAD21 act as chromatin insulators of urea cycle genes in the liver but not in other cell types.

Of the 16 transcription factors whose DNA binding was analyzed in human liver, SP1, CREB, HNF3, COUP-TF, and HNF-4 were known to bind to, and regulate expression of urea cycle genes (Table 1). As expected, transcription factor SP1 was bound to promoters of *NAGS*, *ASS1*, *ASL*, *ARG1*, and SLC25A13 (citrin) genes that lack the TATA-box motif (Figures 4, S5, S8–S10, and S12). However, SP1 was also bound to *CPS1* and *OTC* promoters which do have the TATA-box (Figures 4 and S6, S7) as well as to enhancers of NAGS, CPS1, OTC, and ASS1 that were not previously known to bind this transcription factor (Figures 4 and S5–S8). ChIP-Seq peaks that indicate CREB binding to *NAGS* promoter[109] and *ASS1* enhancer[110] were absent from the ENCODE data. This discrepancy could be due to the different experimental systems used in characterization of *NAGS* and *ASS1* transcriptional regulation and in the ENCODE project. Transcriptional regulation of *NAGS* and *ASS1* genes was studied in HepG2 and HuH7 hepatoma cell lines which recapitulate many, but not all, characteristics of transcriptional regulation of urea cycle genes in hepatocytes. Transcription factor HNF-3β was bound to *CPS1* enhancer (Figure S6) and to *SLC25A13* (citrin) promoter (Figure S12) as expected from previous studies.[111–123] HNF3β also bound to cCREs located approx. 5.5 and 10 kb upstream of the *SLC25A13* (citrin) promoter (Figure S12). ENCODE data also indicate that transcription factors HNF3α and HNF3β regulate expression of *ARG1* (Figure S10 and Table 2). Both transcription factors were bound to *ARG1* promoter and a cCRE located approx. 25 kb upstream of the *ARG1* promoter. HNF3α and HNF3β binding sites were also present approx. 15 kb and 20 kb upstream of the *ARG1* promoter in the same location as DNase-Seq and ATAC-Seq peaks (Figure S10). COUP-TF is a negative regulator of *OTC*[124] and its binding was not detected in the human liver tissue (Table 2). The COUP-TF transcription factor appears to regulate expression of *ASS1* and *SLC25A15 (ORNT)* genes. For *ASS1*, COUP-TF bound in two sites, one located at a cCRE approximately 13 kb and another site approximately 4 kb upstream of the promoter. The latter COUP-TF binding site was in the same location as DNAse-Seq and ATAC-Seq peaks (Figure S8). For *SLC25A15 (ORNT)*, the COUP-TF

TABLE 2 Transcription factors that bind to experimentally verified and predicted regulatory elements of urea cycle genes.

	NAGS	*CPS1*	*OTC*	*ASS1*	*ASL*	*ARG1*	*ORNT*	*Citrin*
ATF3/CREB	–	++[a]	+/–[b]	-[c]	–	++	–	+/–
COUP-TF2	–	+/–	–	++	–	+/–	++	+/–
EGR1[d]	–	+/–	–	–	–	+/–	+/–	+/–
GABPA	–	–	–	–	–	–	++	+/–
HNF3α	–	+/–	–	–	–	++	–	+/–
HNF3β	–	++	–	–	+/–	++	+/–	++
HNF4α	++	++	++	++	++	++	++	++
HNF4γ	+/–	–	+/–	+/–	+/–	+/–	–	–
JUND	–	++	++	+/–	–	++	++	++
MAX	+/–	++	–	+/–	–	++	+/–	+/–
REST	+/–	++	++	++	+/–	++	–	++
RXRα	++	++	++	++	++	++	++	++
SP1	++	++	++	++	++	++	++	++
TAF1	+/–	+/–	–	++	++	++	+/–	+/–
YY1	+/–	++	+/–	+/–	+/–	++	++	+/–
ZBTB33	–	+/–	–	–	–	++	–	+/–
Data visualization	Figure S5	Figure S6	Figure S7	Figure S8	Figure S9	Figure S10	Figure S11	Figure S12

[a]Fold Change over Control >20.
[b]Fold Change over Control 10–20.
[c]Fold Change over Control <10.
[d]ChIP-seq peaks not visualized in Figures S5–S12 because of the low signal in most urea cycle genes.

binding site was approximately 5 kb upstream of the predicted ORNT promoter and in the same location as DNAse-Seq and ATAC-Seq peaks (Figure S11). HNF-4α was bound to *OTC* promoter and enhancer (Figures 4, S7, and Table 2), which is consistent with the known role of this transcription factor in regulation of *OTC* expression.[124–128] The ENCODE data also revealed strong binding of JunD, MAX, REST, TAF, YY1, and ZBTB33 transcription factors to cCREs in urea cycle genes (Table 2 and Figures S5–S12). Each of the six transcription factors binds cCREs in subsets of urea cycle genes and their role in regulation of ureagenesis remains to be determined. Such studies will require an improved model system that recapitulates the hepatocyte nuclear environment more closely than the shortcomings seen with HepG2 and HuH7 cells.

Since our analysis revealed novel cCREs and involvement of additional transcription factors in regulation of urea cycle gene expression we next compared expression patterns of urea cycle genes in human tissues to known expression patterns in model organisms. RNA-seq and quantitative proteomics data from 19 human tissues were retrieved from the GTEx database. Tissue-specific expression patterns between human, mouse, and rat tissues for citrin and six urea cycle genes that encode urea cycle enzymes were similar (Figure S13). As for these rodents,[99,129–133] urea cycle enzymes, and *SLC25A13* (citrin) are highly expressed in human liver (Figure S13). Human and rodent *NAGS/Nags*, *CPS1/Cps1*, and *OTC/Otc* mRNA and proteins are expressed in the small intestine where they function in citrulline biosynthesis[129,134] (Figure S13). Expression of human and rodent *ASS1/Ass1*, *ASL/Asl*, and *SLC25A13/Slc25A13* (citrin) mRNA and proteins in the kidney and small intestine suggests similar functions of the three genes in arginine biosynthesis[99,132,133] (Figure S13). The tissue specificity of *ORNT* gene expression has only been studied in humans.[135] Despite some differences in the biochemical symptoms of patients with *ARG1* and citrin deficiencies and phenotype of *Arg1* and citrin knockout mice,[136,137] similar expression patterns of urea cycle genes suggest conservation of regulatory mechanisms of gene expression between humans and rodents. Therefore, cultured cells and animal models could be used to examine the involvement of JunD, MAX, REST, TAF, YY1, and ZBTB33 in the regulation of ureagenesis.

2.6 | Molecular mechanisms coordinating regulation of urea cycle gene expression

The urea cycle is involved in detoxification of the ammonia produced in the catabolism of proteins and amino acids which allows them to be used as an energy source.

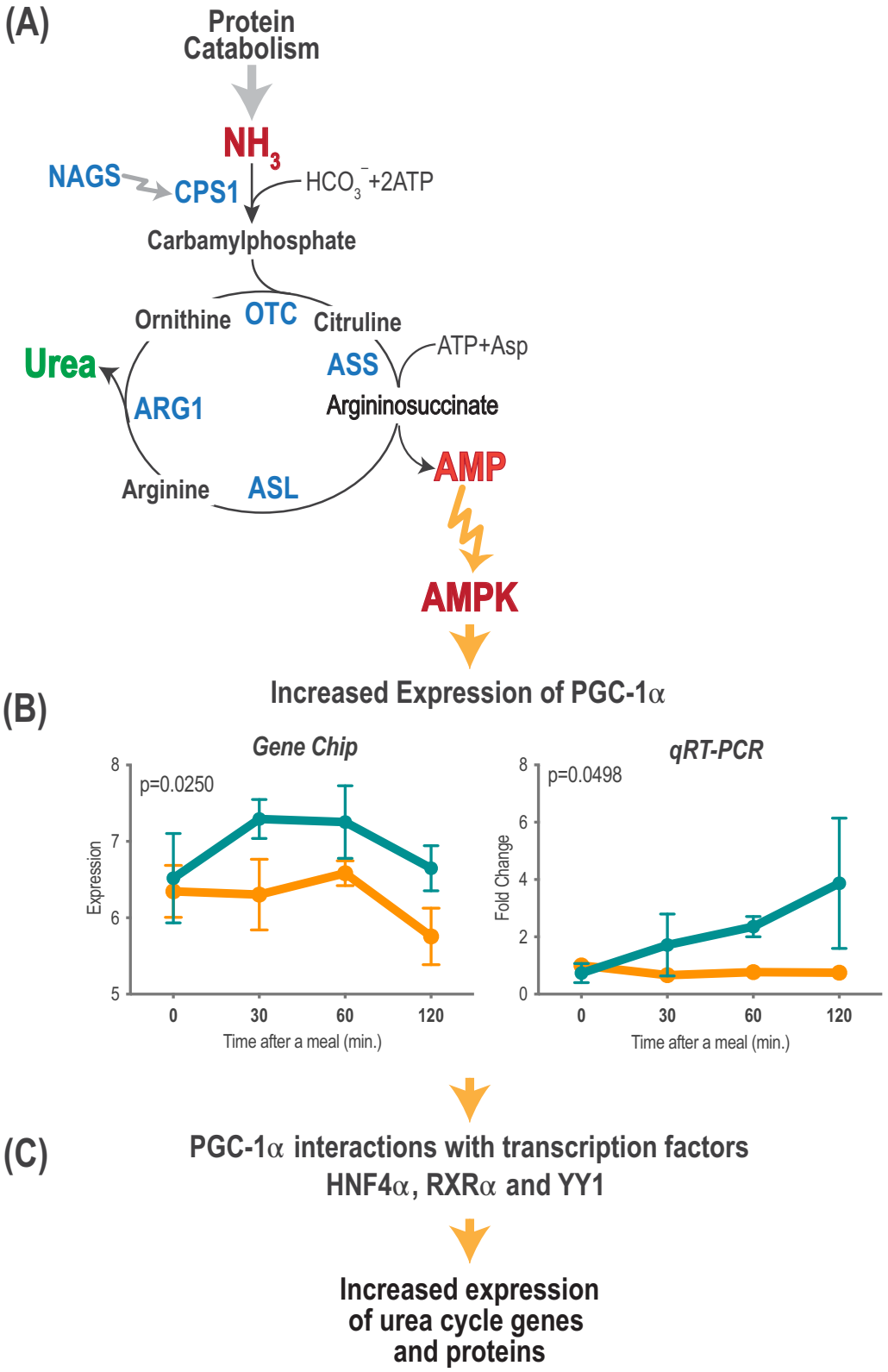


FIGURE 5 Molecular mechanism of urea cycle adaptation to changing nitrogen load. (A) Increased flux through the urea cycle results in increased production of AMP by ASS1, which activates AMPK. ARG1, arginase 1; ASL, argininosuccinate lyase; ASS, argininosuccinate synthase; CPS1, carbamylphosphate synthetase; NAGS, N-acetylglutamate synthase; OTC, ornithine transcarbamylase. (B) Expression of *Pgc-1a* mRNA in mice fed either a high (teal) or low (orange) protein diet. The *Pgc-1α* mRNA abundance was measured using Affymetrix microarrays (left) and validated using qRT-PCT (right). (C) Interactions between *Pgc-1α* and transcription factors HNF4α, RXRα, and YY1 could result in increased expression of urea cycle genes and their products.

In rodents and non-human primates, the changes in dietary protein intake and in the catabolism of cellular proteins, commonly referred to as nitrogen load, cause the expression of urea cycle genes and the abundance of urea cycle enzymes to adapt in response.[29–32] As the flux through the urea cycle increases, the increased production of AMP from the ASS1-catalyzed formation of argininosuccinate has been shown to activate AMP-activated kinase (AMPK)[138] (Figure 5A). Since the AMPK signaling cascade affects multiple transcription factors that regulate expression of genes in carbohydrate and metabolism, we reasoned the adaptation of the urea cycle to changes in nitrogen load could involve one or more transcription factors common to all urea cycle genes. Transcription factors SP1, RXRα, HNF4α, and YY1 bind to regulatory elements of all eight urea cycle genes (Table 2 and Figure 4) and we sought published confirmation of which of these four transcription factors are targeted by the AMPK signaling cascade. SP1 is a target of AMPK signaling in hepatocytes[139] and glioblastoma cells.[140] Therefore, we hypothesize that AMPK signaling could upregulate urea cycle genes through SP1 in response to changing nitrogen load. The activity of HNF4α in the liver is regulated by AMPK and mediated by the phosphorylation of peroxisome proliferator-activated receptor gamma coactivator 1α (PGC-1α).[141] Phosphorylated PGC-1a interacts with many transcription factors, including HNF4α, RXRα, and YY1, to regulate genes involved in the adaptation to changes in cellular state and/or environment.[141–143] PGC-1α is also one of the genes in our experiments that were markedly upregulated in the livers of mice fed a high protein diet.[29] Upregulation of *Ppargc1a* mRNA, which encodes PGC1α, was detected using microarray transcriptional profiling and validated by qRT-PCR (Figure 5B). Therefore, we propose the following model of urea cycle adaptation to changing nitrogen load: AMPK regulates expression of urea cycle genes directly via SP1 and indirectly via PGC-1α dependent interactions with transcription factors HNF4α, RXRα, and/or YY1 (Figure 5C). This model requires experimental verification using animal models and cultured cells adapted to different nitrogen loads. Understanding how the urea cycle genes are coordinately regulated in response to changing nitrogen load can lead to new therapeutic approaches that will stimulate ureagenesis in patients with hyperammonemia due to liver failure and inborn errors of protein catabolism.

3 | CONCLUSIONS

A combination of small sequence length of the NAGS catalytic domain, non-specific clinical symptoms of hyperammonemia, biochemical symptoms, elevated ammonia and glutamine, and low citrulline in the blood, that are shared with other urea cycle disorders appear to be the primary basis for the low incidence of *NAGS* deficiency compared to other urea cycle disorders. Since NAGS deficiency is the only urea cycle disorder that can be treated with a single drug, understanding these factors is important in order to identify every patient with the disease. Rapid genomic sequencing in combination with reliable functional testing and annotation of *NAGS* sequence variants will aid in making diagnoses. Our search for cCREs began with two patients with clinical and biochemical symptoms of urea cycle disorders but no pathogenic sequence variants in the coding regions and splice sites of urea cycle genes, which led to the identification of two *NAGS* regulatory regions. This in turn prompted a wider investigation of urea cycle gene regulation by data mining of the ENCODE database. We identified four transcription factors that regulate all eight urea cycle genes, and uncovered a plausible role for AMPK signaling, upregulation of PGC-1α and its interactions with one or more of the four transcription factors as the molecular mechanism for adaptation of ureagenesis to changing nitrogen load.

4 | MATERIALS AND METHODS

4.1 | Datamining

GnomAD database was queried for *NAGS*, *CPS1*, and *ARG1* missense, synonymous and nonsense variants and the lists of variants were downloaded as CSV files. Single nucleotide variants in genomic regions surrounding NAGS, CPS1, and ARG1 genes were retrieved using the UCSC Genome Browser Table Browser tool. Lists of single nucleotide variants were retrieved from 1000G Ph3 Vars track and tgpPhase3 table for genomic regions chr17:43954622-44 059 068, chr2:210506599-210 729 107, and chr6:131523226-131 634 329 of the GRCh38/hg38 human genome assembly. Clustal Omega and default parameters were used for multiple sequence alignments of NAGS, CPS, OTC, and arginase proteins from 58 organisms. Percent Identity Matrices, generated by Clustal Omega, were used to visualize conservation of the four proteins.

The ENCODE Project Functional Genomics Portal was queried for the availability of data for DNase sensitivity and hypersensitivity sites, CTCF binding sites, RNA polymerase II, and transcription factor binding sites in the human liver. The following filters were applied to ENCODE Experimental Matrix: DNA binding and DNA accessibility for assay type; TF-ChIP-seq, DNase-seq, and ATAC-seq for assay title; *Homo sapiens* for Organism; tissue for biosample classification; liver for biosample; liver

for organ. BigWig fold-change files for biological and technical replicates for the following ChIP-Seq experiments were down-loaded: CTCF, RAD21, RNAP2A, ATF3, COUP-TF2, EGR1, GABPA, HNF3α, HNF3β, HNF4α, HNF4γ, JUND, MAX, REST, RXRα, SP1, TAF1, YY1, ZBTB33 as well as DNase-seq, ATAC-seq. Visualization of the ChIP-Seq data was carried out using custom Python scripts, available at https://github.com/MIMOR02/bigwig-file-vizualizations, as described before.[144]

4.2 | *Pgc-1α* gene expression analysis

The Institutional Animal Care and Use Committee of the Children's National Hospital approved all experimental procedures involving mice. All institutional and national guidelines for the care and use of laboratory animals were followed.

Transcriptional profiling used to determine differentially expressed genes, including *Pgc-1α*, in the livers of mice fed high and low protein diets on a restricted feeding schedule (18 h feeding and 6 h fasting) and validation of the expression changes have been described previously.[29] Four mice per diet were analyzed at each time point (fasting, 30, 60, and 120 min after a meal) for a total of 32 animals. RNA was isolated from frozen livers using TRIzol reagent (Invitrogen), converted into cDNA, labeled with biotin and hybridized to GeneChip Mouse Genome 430 2.0 (Affymetrix). Fluorescent images were scanned and analyzed using Probe Microarray Suite (MAS) version 5.0. Partek software package (Partek Incorporated) was used to identify differentially expressed genes in the livers of mice fed high and low-protein diets.[29] Normalized fluorescence intensities for the probe set 1456394_at were used to visualize the *Pgc-1a* gene expression in Figure 5C.

The *Pgc-1a* mRNA expression differences in the livers of mice fed high and low protein diets were validated using RNA extracted with Trizol Reagent (Invitrogen) from a separate set of 32 liver samples.[29] *Pgc-1a* transcripts were quantified using Applied Biosystems 384 custom gene card array on a 7900HT Fast Real-Time PCR System (Applied Biosystems, Inc.). The $\Delta\Delta C_t$ method[145] was to calculate the difference in *Pgc1a* gene expression levels. GraphPad Prism and two-way ANOVA were used for statistical analysis.

AUTHOR CONTRIBUTIONS

Ljubica Caldovic conceived and directed the study and wrote the manuscript. Julie J. Ahn performed data mining to identify *NAGS* expression patterns and regulatory elements. Jacklyn Andricovic performed data mining to identify *SLC25A15* expression patterns and regulatory elements. Veronica M. Balick, Sveta V. Jagannathan, and Emily C. Williams collected synonymous, missense, and nonsense sequence variants from gnomAD. Tyson Dawson performed data mining to identify *CPS1* expression patterns and regulatory elements. Alex C. Edwards, Pamela A. Chansky, and Mallory Brayer carried out analysis of NAGS, CPS1, OTC, and arginase protein conservation. Sara E. Felsen, Karim Ismat, and Shatha Salameh carried out data mining for alternative sources of NAGS and NCG and performed a literature review of *NAGS* deficiency cases. Brendan T. Mann performed datamining to identify *ASL* expression patterns and regulatory elements. Jacob A. Medina performed datamining to identify *ASS1* expression patterns and regulatory elements. Toshio Morizono performed datamining to identify *SLC25A13* expression patterns and regulatory elements. Michio Morizono wrote Python scripts that were used for visualization of ChIP-Seq data. Neerja Vashist performed data mining to identify *ARG1* expression patterns and regulatory elements. Zhe Zhou performed data mining to identify *OTC* expression patterns and regulatory elements. Hiroki Morizono conceived and directed the study and edited the manuscript.

ACKNOWLEDGEMENTS

We are grateful to Dr. Clare Woodward for helpful discussions and advice related to this project.

FUNDING INFORMATION

This work was supported by the Public Health Service Grant K01DK076846 from the National Institute of Diabetes Digestive and Kidney Diseases, National Institutes of Health, and the Recordati Rare Disease, Inc. Special Purpose Fund.

CONFLICT OF INTEREST STATEMENT

Dr. Caldovic's work was supported in part by the Recordati Rare Diseases, Inc. which manufactures and sells NCG as Carbaglu® (carglumic acid), which is used for treatment of *NAGS* deficiency. Julie J. Ahn, Jacklyn Andricovic, Veronica M. Balick, Mallory Bryer, Pamela A. Chansky, Tyson Dawson, Alex C. Edwards, Sara E. Felsen, Karim Ismat, Sveta V Jagannathan, Brendan T Mann, Jacob A Medina, Toshio Morizono, Michio Morizono, Shatha Salameh, Neerja Vashist, Emily C. Williams, Zhe Zhou, and Hiroki Morizono declare that they have no conflict of interest.

DATA AVAILABILITY STATEMENT

All data used in this study have been included in supplementary files. Custom Python scripts used for analysis of the data from the ENCODE Project are available at https://github.com/MIMOR02/bigwig-file-vizualizations.

ETHICS STATEMENT

This article does not contain any studies with human subjects performed by the any of the authors.

PATIENT CONSENT STATEMENT

This article does not contain any studies with human subjects performed by the any of the authors.

ANIMAL RIGHTS

The Institutional Animal Care and Use Committee of the Children's National Hospital approved all experimental procedures involving mice. All institutional and national guidelines for the care and use of laboratory animals were followed.

ORCID

Ljubica Caldovic https://orcid.org/0000-0002-9140-5585
Julie J. Ahn https://orcid.org/0000-0001-5168-6717
Jacklyn Andricovic https://orcid.org/0000-0003-2945-5095
Veronica M. Balick https://orcid.org/0000-0002-7640-3784
Mallory Brayer https://orcid.org/0009-0002-0133-0784
Pamela A. Chansky https://orcid.org/0000-0002-8169-992X
Tyson Dawson https://orcid.org/0000-0003-4888-9673
Alex C. Edwards https://orcid.org/0000-0001-8608-0133
Sara E. Felsen https://orcid.org/0000-0003-3208-7860
Karim Ismat https://orcid.org/0000-0003-4398-0101
Sveta V. Jagannathan https://orcid.org/0000-0001-5393-0162
Brendan T. Mann https://orcid.org/0000-0003-0694-7353
Jacob A. Medina https://orcid.org/0000-0001-9563-4994
Toshio Morizono https://orcid.org/0009-0005-2375-0685
Michio Morizono https://orcid.org/0009-0005-9452-4060
Shatha Salameh https://orcid.org/0000-0001-9624-4277
Neerja Vashist https://orcid.org/0000-0002-1887-270X
Emily C. Williams https://orcid.org/0000-0002-5561-6553
Zhe Zhou https://orcid.org/0000-0003-3253-6355
Hiroki Morizono https://orcid.org/0000-0002-9678-5564

REFERENCES

1. Ah Mew N, Simpson KL, Gropman AL, Lanpher BC, Chapman KA, Summar ML. Urea cycle disorders overview. In: Adam MP, Ardinger HH, Pagon RA, et al., eds. GeneReviews®; 1993.
2. Caldovic L, Tuchman M. N-acetylglutamate and its changing role through evolution. *Biochem J*. 2003;372(Pt 2):279-290.
3. Hall LM, Metzenberg RL, Cohen PP. Isolation and characterization of a naturally occurring cofactor of carbamyl phosphate biosynthesis. *J Biol Chem*. 1958;230:1013-1021.
4. Balasubramaniam S, Rudduck C, Bennetts B, Peters G, Wilcken B, Ellaway C. Contiguous gene deletion syndrome in a female with ornithine transcarbamylase deficiency. *Mol Genet Metab*. 2010;99(1):34-41. doi:10.1016/j.ymgme.2009.08.007
5. Dionisi-Vici C, Rizzo C, Burlina AB, et al. Inborn errors of metabolism in the Italian pediatric population: a national retrospective survey. *J Pediatr*. 2002;140(3):321-327.
6. Keskinen P, Siitonen A, Salo M. Hereditary urea cycle diseases in Finland. *Acta Paediatr*. 2008;97(10):1412-1419. doi:10.1111/j.1651-2227.2008.00923.x
7. Summar ML, Koelker S, Freedenberg D, et al. The incidence of urea cycle disorders. *Mol Genet Metab*. 2013;110(1–2):179-180. doi:10.1016/j.ymgme.2013.07.008
8. Haberle J, Burlina A, Chakrapani A, et al. Suggested guidelines for the diagnosis and management of urea cycle disorders: first revision. *J Inherit Metab Dis*. 2019;42(6):1192-1230. doi:10.1002/jimd.12100
9. Caldovic L, Morizono H, Daikhin Y, et al. Restoration of ureagenesis in N-acetylglutamate synthase deficiency by N-carbamylglutamate. *J Pediatr*. 2004;145(4):552-554.
10. Tuchman M, Caldovic L, Daikhin Y, et al. N-carbamylglutamate markedly enhances ureagenesis in N-acetylglutamate deficiency and propionic acidemia as measured by isotopic incorporation and blood biomarkers. *Pediatr Res*. 2008;64(2):213-217. doi:10.1203/PDR.0b013e318179454b
11. Haberle J, Boddaert N, Burlina A, et al. Suggested guidelines for the diagnosis and management of urea cycle disorders. *Orphanet J Rare Dis*. 2012;7:32. doi:10.1186/1750-1172-7-32
12. Kenneson A, Singh RH. Presentation and management of N-acetylglutamate synthase deficiency: a review of the literature. *Orphanet J Rare Dis*. 2020;15(1):279. doi:10.1186/s13023-020-01560-z
13. Haberle J, Moore MB, Haskins N, et al. Noncoding sequence variants define a novel regulatory element in the first intron of the N-acetylglutamate synthase gene. *Hum Mutat*. 2021;42(12):1624-1636. doi:10.1002/humu.24281
14. Heibel SK, Ah Mew N, Caldovic L, Daikhin Y, Yudkoff M, Tuchman M. N-carbamylglutamate enhancement of ureagenesis leads to discovery of a novel deleterious mutation in a newly defined enhancer of the NAGS gene and to effective therapy. *Hum Mutat*. 2011;32(10):1153-1160. doi:10.1002/humu.21553
15. Sonaimuthu P, Senkevitch E, Haskins N, et al. Gene delivery corrects N-acetylglutamate synthase deficiency and enables insights in the physiological impact of L-arginine activation of N-acetylglutamate synthase. *Sci Rep*. 2021;11(1):3580. doi:10.1038/s41598-021-82994-8
16. Williams M, Burlina A, Rubert L, et al. N-Acetylglutamate synthase deficiency due to a recurrent sequence variant in the N-acetylglutamate synthase enhancer region. *Sci Rep*. 2018;8(1):15436. doi:10.1038/s41598-018-33457-0
17. Haberle J, Shchelochkov OA, Wang J, et al. Molecular defects in human carbamoy phosphate synthetase I: mutational spectrum, diagnostic and protein structure considerations. *Hum Mutat*. 2011;32(6):579-589. doi:10.1002/humu.21406
18. Yamaguchi S, Brailey LL, Morizono H, Bale AE, Tuchman M. Mutations and polymorphisms in the human ornithine transcarbamylase (OTC) gene. *Hum Mutat*. 2006;27(7):626-632.

19. Shchelochkov OA, Li FY, Geraghty MT, et al. High-frequency detection of deletions and variable rearrangements at the ornithine transcarbamylase (OTC) locus by oligonucleotide array CGH. *Mol Genet Metab*. 2009;96(3):97-105. doi:10.1016/j.ymgme.2008.11.167
20. Funghini S, Thusberg J, Spada M, et al. Carbamoyl phosphate synthetase 1 deficiency in Italy: clinical and genetic findings in a heterogeneous cohort. *Gene*. 2012;493(2):228-234. doi:10.1016/j.gene.2011.11.052
21. Kurokawa K, Yorifuji T, Kawai M, et al. Molecular and clinical analyses of Japanese patients with carbamoylphosphate synthetase 1 (CPS1) deficiency. *J Hum Genet*. 2007;52(4):349-354. doi:10.1007/s10038-007-0122-9
22. Kido J, Häberle J, Sugawara K, et al. Clinical manifestation and long-term outcome of citrin deficiency: report from a nationwide study in Japan. *J Inherit Metab Dis*. 2022;45(3):431-444. doi:10.1002/jimd.12483
23. Lin WX, Zeng H-S, Zhang Z-H, et al. Molecular diagnosis of pediatric patients with citrin deficiency in China: SLC25A13 mutation spectrum and the geographic distribution. *Sci Rep*. 2016;6:29732. doi:10.1038/srep29732
24. Balmer C, Pandey AV, Rüfenacht V, et al. Mutations and polymorphisms in the human argininosuccinate lyase (ASL) gene. *Hum Mutat*. 2014;35(1):27-35. doi:10.1002/humu.22469
25. Bernal AC, Tubio MC, Crespo C, Eiroa HD. Clinical and genetic characterization and biochemical correlation at presentation in 48 patients diagnosed with urea cycle disorders at the hospital Juan P Garrahan, Argentina. *J Inborn Errors Metab Screen*. 2021;9:e20200026. doi:10.1590/2326-4594-JIEMS-2020-0026
26. Isler J, Rüfenacht V, Gemperle C, Allegri G, Häberle J. Improvement of diagnostic yield in carbamoylphosphate synthetase 1 (CPS1) molecular genetic investigation by RNA sequencing. *JIMD Rep*. 2020;52(1):28-34. doi:10.1002/jmd2.12091
27. Jang YJ, LaBella AL, Feeney TP, et al. Disease-causing mutations in the promoter and enhancer of the ornithine transcarbamylase gene. *Hum Mutat*. 2018;39(4):527-536. doi:10.1002/humu.23394
28. Han ST, Anderson KJ, Bjornsson HT, Longo N, Valle D. A promoter variant in the OTC gene associated with late and variable age of onset hyperammonemia. *J Inherit Metab Dis*. 2022;45(4):710-718. doi:10.1002/jimd.12524
29. Heibel SK, McGuire PJ, Haskins N, et al. AMP-activated protein kinase signaling regulated expression of urea cycle enzymes in response to changes in dietary protein intake. *J Inherit Metab Dis*. 2019;42(6):1088-1096. doi:10.1002/jimd.12133
30. Schimke RT. Differential effects of fasting and protein-free diets on levels of urea cycle enzymes in rat liver. *J Biol Chem*. 1962;237:1921-1924.
31. Schimke RT. Adaptive characteristics of urea cycle enzymes in the rat. *J Biol Chem*. 1962;237:459-468.
32. Nuzum CT, Snodgrass PJ. Urea cycle enzyme adaptation to dietary protein in primates. *Science*. 1971;172(3987):1042-1043.
33. Saheki T, Hosoya M, Fujinami S, Katsunuma T. Regulation of urea synthesis: changes in the concentration of ornithine in the liver corresponding to changes in urea synthesis. *Adv Exp Med Biol*. 1982;153:255-263.
34. McGuire PJ, Lee H-S, Summar ML. Infectious precipitants of acute hyperammonemia are associated with indicators of increased morbidity in patients with urea cycle disorders. *J Pediatr*. 2013;163(6):1705-1710.e1. doi:10.1016/j.jpeds.2013.08.029
35. McGuire PJ, Tarasenko TN, Wang T, et al. Acute metabolic decompensation due to influenza in a mouse model of ornithine transcarbamylase deficiency. *Dis Models Mech*. 2014;7(2):205-213. doi:10.1242/dmm.013003
36. Adam S, Champion H, Daly A, et al. Dietary management of urea cycle disorders: UK practice. *J Hum Nutr Diet*. 2012;25(4):398-404. doi:10.1111/j.1365-277X.2012.01259.x
37. Bachmann C, Brandis M, Weissenbarth-Riedel E, Burghard R, Colombo JP. N-acetylglutamate synthetase deficiency, a second patient. *J Inherit Metab Dis*. 1988;11(2):191-193.
38. Bachmann C, Colombo JP, Jaggi K. N-acetylglutamate synthetase (NAGS) deficiency: diagnosis, clinical observations and treatment. *Adv Exp Med Biol*. 1982;153:39-45.
39. Bachmann C, Krähenbühl S, Colombo JP, Schubiger G, Jaggi KH, Tönz O. N-acetylglutamate synthetase deficiency: a disorder of ammonia detoxication. *N Engl J Med*. 1981;304(9):543.
40. Burlina AB, Bachmann C, Wermuth B, et al. Partial N-acetylglutamate synthetase deficiency: a new case with uncontrollable movement disorders. *J Inherit Metab Dis*. 1992;15(3):395-398.
41. Elpeleg ON, Colombo JP, Amir N, Bachmann C, Hurvitz H. Late-onset form of partial N-acetylglutamate synthetase deficiency. *Eur J Pediatr*. 1990;149(9):634-636.
42. Forget PP, Oosterhout M, Bakker JA, Wermuth B, Vies JSH, Spaapen LJM. Partial N-acetyl-glutamate synthetase deficiency masquerading as a valproic acid-induced Reye-like syndrome. *Acta Paediatr*. 1999;88(12):1409-1411.
43. Hinnie J, Colombo JP, Wermuth B, Dryburgh FJ. N-Acetylglutamate synthetase deficiency responding to carbamylglutamate. *J Inherit Metab Dis*. 1997;20(6):839-840.
44. Morris AA, Richmond SW, Oddie SJ, Pourfarzam M, Worthington V, Leonard JV. N-acetylglutamate synthetase deficiency: favourable experience with carbamylglutamate. *J Inherit Metab Dis*. 1998;21(8):867-868.
45. Pandya AL, Koch R, Hommes FA, Williams JC. N-acetylglutamate synthetase deficiency: clinical and laboratory observations. *J Inherit Metab Dis*. 1991;14(5):685-690.
46. Schubiger G, Bachmann C, Barben P, Colombo JP, Tönz O, Schüpbach D. N-acetylglutamate synthetase deficiency: diagnosis, management and follow-up of a rare disorder of ammonia detoxication. *Eur J Pediatr*. 1991;150(5):353-356.
47. Vockley J, Vockley CMW, Lin SP, et al. Normal N-acetylglutamate concentration measured in liver from a new patient with N-acetylglutamate synthetase deficiency: physiologic and biochemical implications. *Biochem Med Metab Biol*. 1992;47(1):38-46.
48. Karczewski KJ, Francioli LC, Tiao G, et al. The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature*. 2020;581(7809):434-443. doi:10.1038/s41586-020-2308-7
49. Brusilow SW, Horwich AL. Urea cycle enzymes. In: Beaudet AL, Sly WS, Valle D, eds. *CR scriver*. McGraw-Hill; 2001:1909-1963.

50. Meijer AJ, Lamers WH, Chamuleau RA. Nitrogen metabolism and ornithine cycle function. *Physiol Rev*. 1990;70(3):701-748.
51. Abecasis GR et al. An integrated map of genetic variation from 1,092 human genomes. *Nature*. 2012;491(7422):56-65. doi:10.1038/nature11632
52. Sancho-Vaello E, Marco-Marín C, Gougeard N, et al. Understanding N-acetyl-L-glutamate synthase deficiency: mutational Spectrum, impact of clinical mutations on enzyme functionality, and structural considerations. *Hum Mutat*. 2016;37(7):679-694. doi:10.1002/humu.22995
53. Haskins N, Mumo A, Brown PH, Tuchman M, Morizono H, Caldovic L. Effect of arginine on oligomerization and stability of N-acetylglutamate synthase. *Sci Rep*. 2016;6:38711. doi:10.1038/srep38711
54. Haskins N, Panglao M, Qu Q, et al. Inversion of allosteric effect of arginine on N-acetylglutamate synthase, a molecular marker for evolution of tetrapods. *BMC Biochem*. 2008;9:24.
55. Caldovic L, Morizono H, Gracia Panglao M, et al. Cloning and expression of the human N-acetylglutamate synthase gene. *Biochem Biophys Res Commun*. 2002;299(4):581-586.
56. Zhao G, Jin Z, Allewell NM, Tuchman M, Shi D. Crystal structure of the N-acetyltransferase domain of human N-acetyl-L-glutamate synthase in complex with N-acetyl-L-glutamate provides insights into its catalytic and regulatory mechanisms. *PLoS One*. 2013;8(7):e70369. doi:10.1371/journal.pone.0070369
57. Caldovic L, Morizono H, Panglao MG, et al. Late onset N-acetylglutamate synthase deficiency caused by hypomorphic alleles. *Hum Mutat*. 2005;25(3):293-298.
58. Caldovic L, Morizono H, Tuchman M. Mutations and polymorphisms in the human N-acetylglutamate synthase (NAGS) gene. *Hum Mutat*. 2007;28(8):754-759.
59. Sancho-Vaello E, Fernández-Murga ML, Rubio V. Site-directed mutagenesis studies of acetylglutamate synthase delineate the site for the arginine inhibitor. *FEBS Lett*. 2008;582(7):1081-1086.
60. de Cima S, Polo LM, Díez-Fernández C, et al. Structure of human carbamoyl phosphate synthetase: deciphering the on/off switch of human ureagenesis. *Sci Rep*. 2015;5:16950. doi:10.1038/srep16950
61. di Costanzo L, Sabio G, Mora A, et al. Crystal structure of human arginase I at 1.29-A resolution and exploration of inhibition in the immune response. *Proc Natl Acad Sci USA*. 2005;102(37):13058-13063. doi:10.1073/pnas.0504027102
62. Schmidt E, Nuoffer JM, Häberle J, et al. Identification of novel mutations of the human N-acetylglutamate synthase gene and their functional investigation by expression studies. *Biochim Biophys Acta*. 2005;1740(1):54-59.
63. Marvil DK, Leisinger T. N-acetylglutamate synthase of *Escherichia coli*: purification, characterization, and molecular properties. *J Biol Chem*. 1977;252(10):3295-3303.
64. Qu Q, Morizono H, Shi D, Tuchman M, Caldovic L. A novel bifunctional N-acetylglutamate synthase-kinase from Xanthomonas campestris that is closely related to mammalian N-acetylglutamate synthase. *BMC Biochem*. 2007;8:4.
65. Hata A, Tsuzuki T, Shimada K, Takiguchi M, Mori M, Matsuda I. Structure of the human ornithine transcarbamylase gene. *J Biochem*. 1988;103(2):302-308.
66. Min L, Jin Z, Caldovic L, et al. Mechanism of allosteric inhibition of N-acetyl-L-glutamate synthase by L-arginine. *J Biol Chem*. 2009;284(8):4873-4880. doi:10.1074/jbc.M805348200
67. Shi D, Sagar V, Jin Z, et al. The crystal structure of N-acetyl-L-glutamate synthase from *Neisseria gonorrhoeae* provides insights into mechanisms of catalysis and regulation. *J Biol Chem*. 2008;283(11):7176-7184.
68. Zhao G, Haskinsm N, Jin Z, Allewell NM, Tuchman M, Shi D. Structure of N-acetyl-L-glutamate synthase/kinase from *Maricaulis maris* with the allosteric inhibitor L-arginine bound. *Biochem Biophys Res Commun*. 2013;437(4):585-590. doi:10.1016/j.bbrc.2013.07.003
69. Kim J, Raushel FM. Allosteric control of the oligomerization of carbamoyl phosphate synthetase from *Escherichia coli*. *Biochemistry*. 2001;40(37):11030-11036. doi:10.1021/bi011121u
70. de Gregorio A, Battistutta R, Arena N, et al. Functional and structural characterization of ovine ornithine transcarbamoylase. *Org Biomol Chem*. 2003;1(18):3178-3185. doi:10.1039/b304901a
71. Shi D, Morizono H, Ha Y, Aoyagi M, Tuchman M, Allewell NM. 1.85-A resolution crystal structure of human ornithine transcarbamoylase complexed with N-phosphonacetyl-L-ornithine. Catalytic mechanism and correlation with inherited deficiency. *J Biol Chem*. 1998;273(51):34247-34254.
72. Ha Y, McCann MT, Tuchman M, Allewell NM. Substrate-induced conformational change in a trimeric ornithine transcarbamoylase. *Proc Natl Acad Sci USA*. 1997;94(18):9550-9555. doi:10.1073/pnas.94.18.9550
73. Itoh Y, Soldati L, Stalon V, et al. Anabolic ornithine carbamoyltransferase of *Pseudomonas aeruginosa*: nucleotide sequence and transcriptional control of the argF structural gene. *J Bacteriol*. 1988;170(6):2725-2734. doi:10.1128/jb.170.6.2725-2734.1988
74. Chang C, Evdokimova E, Mcchesney M, Joachimiak A, Savchenko A, Center for Structural Genomics of Infectious Diseases (CSGID). Crystal structure of arginase from *Bacillus cereus* 2018 [updated 15 may 2019]. Available from. doi:10.2210/pdb6nbk/pdb
75. Kanyo ZF, Scolnick LR, Ash DE, Christianson DW. Structure of a unique binuclear manganese cluster in arginase. *Nature*. 1996;383(6600):554-557. doi:10.1038/383554a0
76. Zhang J, Zhang X, Wu C, et al. Expression, purification and characterization of arginase from *Helicobacter pylori* in its apo form. *PLoS One*. 2011;6(10):e26205. doi:10.1371/journal.pone.0026205
77. Bewley MC, Jeffrey PD, Patchett ML, Kanyo ZF, Baker EN. Crystal structures of *Bacillus caldovelox* arginase in complex with substrate and inhibitors reveal new insights into activation, inhibition and catalysis in the arginase superfamily. *Structure*. 1999;7(4):435-448. doi:10.1016/s0969-2126(99)80056-2
78. Kumarevel TS, Karthe P, Kuramitsu S, Yokoyama S, RIKEN Structural Genomics/Proteomics Initiative (RSGI). Crystal structure of the arginase from *Thermus thermophilus*: RCSB. 2007 [Updated 13 July 2011]. Available from. doi:10.2210/pdb2ef4/pdb
79. Shi D, Li Y, Cabrera-Luque J, et al. A novel N-acetylglutamate synthase architecture revealed by the crystal structure of the bifunctional enzyme from *Maricaulis maris*. *PLoS One*. 2011;6(12):e28825. doi:10.1371/journal.pone.0028825

80. Grisolia S, Cohen PP. The catalytic role of carbamyl glutamate in citrulline biosynthesis. *J Biol Chem*. 1952;198(2):561-571.
81. Grisolia S, Cohen PP. Catalytic role of of glutamate derivatives in citrulline biosynthesis. *J Biol Chem*. 1953;204(2):753-757.
82. Forman HJ, Waddell R, Hamilton PB, Grisolia S. Activation of carbamoyl phosphate synthase by N-acetyl-L-aspartate. *Biochem J*. 1974;143(1):63-66. doi:10.1042/bj1430063
83. Deignan JL, Cederbaum SD, Grody WW. Contrasting features of urea cycle disorders in human patients and knockout mouse models. *Mol Genet Metab*. 2008;93(1):7-14. doi:10.1016/j.ymgme.2007.08.123
84. Senkevitch E, Cederbaum SD, Grody WW. A novel biochemically salvageable animal model of hyperammonemia devoid of N-acetylglutamate synthase. *Mol Genet Metab*. 2012;106(2):160-168. doi:10.1016/j.ymgme.2012.03.004
85. Pinner JR, Freckmann M-L, Kirk EP, Yoshino M. Female heterozygotes for the hypomorphic R40H mutation can have ornithine transcarbamylase deficiency and present in early adolescence: a case report and review of the literature. *J Med Case Rep*. 2010;4:361. doi:10.1186/1752-1947-4-361
86. Alonso E, García-Pérez MA, Bueso J, Rubio V. N-acetyl-L-glutamate in brain: assay, levels, and regional and subcellular distribution. *Neurochem Res*. 1991;16(7):787-794.
87. Moreno A, Ross BD, Bluml S. Direct determination of the N-acetyl-L-aspartate synthesis rate in the human brain by (13) C MRS and [1-(13)C]glucose infusion. *J Neurochem*. 2001;77(1):347-350. doi:10.1046/j.1471-4159.2001.t01-1-00282.x
88. de Tommaso M, Ceci E, Pica C, et al. Serum levels of N-acetyl-aspartate in migraine and tension-type headache. *J Headache Pain*. 2012;13(5):389-394. doi:10.1007/s10194-012-0448-3
89. Enthoven LF, Shi Y, Fay EE, et al. The effects of pregnancy on amino acid levels and nitrogen disposition. *Metabolites*. 2023;13(2):242. doi:10.3390/metabo13020242
90. Reglero A, Rivas J, Mendelson J, Wallace R, Grisolia S. Deacylation and transacetylation of acetyl glutamate and acetyl ornithine in rat liver. *FEBS Lett*. 1977;81(1):13-17.
91. Hassall H, Greenberg DM. The bacterial metabolism of L-Hydantoin-5-propionic acid to carbamylglutamic acid and glutamic acid. *J Biol Chem*. 1963;238:3325-3329.
92. Pei J, Li F, Xie Y, Liu J, Yu T, Feng X. Microbial and metabolomic analysis of gingival crevicular fluid in general chronic periodontitis patients: lessons for a predictive, preventive, and personalized medical approach. *EPMA J*. 2020;11(2):197-215. doi:10.1007/s13167-020-00202-5
93. Menni C, Zierer J, Pallister T, et al. Omega-3 fatty acids correlate with gut microbiome diversity and production of N-carbamylglutamate in middle aged and elderly women. *Sci Rep*. 2017;7(1):11079. doi:10.1038/s41598-017-10382-2
94. Bowerman KL, Rehman SF, Vaughan A, et al. Disease-associated gut microbiome and metabolome changes in patients with chronic obstructive pulmonary disease. *Nat Commun*. 2020;11(1):5886. doi:10.1038/s41467-020-19701-0
95. Visconti A, Le Roy CI, Rosa F, et al. Interplay between the human gut microbiome and host metabolism. *Nat Commun*. 2019;10(1):4505. doi:10.1038/s41467-019-12476-z
96. Luo Y, Hitz BC, Gabdank I, et al. New developments on the encyclopedia of DNA elements (ENCODE) data portal. *Nucleic Acids Res*. 2020;48(D1):D882-D889. doi:10.1093/nar/gkz1062
97. EP Consortium, Moore JE, Purcaro MJ, Pratt HE, et al. Expanded encyclopaedias of DNA elements in the human and mouse genomes. *Nature*. 2020;583(7818):699-710. doi:10.1038/s41586-020-2493-4
98. Ohtake A, Takiguchi M, Shigeto Y, Amaya Y, Kawamoto S, Mori M. Structural organization of the gene for rat liver-type arginase. *J Biol Chem*. 1988;263(5):2245-2249.
99. Yu Y, Terada K, Nagasaki A, Takiguchi M, Mori M. Preparation of recombinant argininosuccinate synthetase and argininosuccinate lyase: expression of the enzymes in rat tissues. *J Biochem*. 1995;117(5):952-957. doi:10.1093/oxfordjournals.jbchem.a124826
100. Lee JS, Adler L, Karathi H, et al. Urea cycle dysregulation generates clinically relevant genomic and biochemical signatures. *Cell*. 2018;174(6):1559-1570. doi:10.1016/j.cell.2018.07.019
101. Nagamani SC, Erez A. A metabolic link between the urea cycle and cancer cell proliferation. *Mol Cell Oncol*. 2016;3(2):e1127314. doi:10.1080/23723556.2015.1127314
102. Rabinovich S, Adler L, Yizhak K, et al. Diversion of aspartate in ASS1-deficient tumours fosters de novo pyrimidine synthesis. *Nature*. 2015;527(7578):379-383. doi:10.1038/nature15529
103. Rabinovich S, Silberman A, Adler L, et al. The mitochondrial carrier Citrin plays a role in regulating cellular energy during carcinogenesis. *Oncogene*. 2020;39(1):164-175. doi:10.1038/s41388-019-0976-2
104. Celiktas M, Tanaka I, Tripathi SC, et al. Role of CPS1 in cell growth, metabolism and prognosis in LKB1-inactivated lung adenocarcinoma. *J Natl Cancer Inst*. 2017;109(3):1-9. doi:10.1093/jnci/djw231
105. Kim J, Hu Z, Cai L, et al. CPS1 maintains pyrimidine pools and DNA synthesis in KRAS/LKB1-mutant lung cancer cells. *Nature*. 2017;546(7656):168-172. doi:10.1038/nature22359
106. Makris G, Kayhan S, Kreuzer M, et al. Impact of small molecule-mediated inhibition of ammonia detoxification on lung malignancies and liver metabolism. *Cancer Commun*. 2023. Epub 2023/01/29;43:508-512. doi:10.1002/cac2.12402
107. Popay TM, Dixon JR. Coming full circle: on the origin and evolution of the looping model for enhancer-promoter communication. *J Biol Chem*. 2022;298(8):102117. doi:10.1016/j.jbc.2022.102117
108. Chang LH, Ghosh S, Noordermeer D. TADs and their borders: free movement or building a wall? *J Mol Biol*. 2020;432(3):643-652. doi:10.1016/j.jmb.2019.11.025
109. Heibel SK, Lopez GY, Panglao M, et al. Transcriptional regulation of N-acetylglutamate synthase. *PLoS One*. 2012;7(2):e29527. doi:10.1371/journal.pone.0029527
110. Guei TR, Liu M-C, Yang C-P, Su T-S. Identification of a liver-specific cAMP response element in the human argininosuccinate synthetase gene. *Biochem Biophys Res Commun*. 2008;377(1):257-261. doi:10.1016/j.bbrc.2008.09.118
111. Goping IS, Lamontagne S, Shore GC, Nguyen M. A gene-type-specific enhancer regulates the carbamyl phosphate synthetase I promoter by cooperating with the proximal GAG activating element. *Nucleic Acids Res*. 1995;23(10):1717-1721.
112. Goping IS, Shore GC. Interactions between repressor and anti-repressor elements in the carbamyl phosphate synthetase I promoter. *J Biol Chem*. 1994;269(5):3891-3896.
113. Lagace M, Goping IS, Mueller CR, Lazzaro M, Shore GC. The carbamyl phosphate synthetase promoter contains multiple

binding sites for C/EBP-related proteins. *Gene*. 1992;118(2): 231-238.
114. Christoffels VM, Grange T, Kaestner KH, et al. Glucocorticoid receptor, C/EBP, HNF3, and protein kinase A coordinately activate the glucocorticoid response unit of the carbamoylphosphate synthetase I gene. *Mol Cell Biol*. 1998;18(11):6305-6315.
115. Christoffels VM, Habets PEMH, das AT, et al. A single regulatory module of the carbamoylphosphate synthetase I gene executes its hepatic program of expression. *J Biol Chem*. 2000; 275(51):40020-40027. doi:10.1074/jbc.M007001200
116. Christoffels VM, van den Hoff MJB, Lamers MC, et al. The upstream regulatory region of the carbamoyl-phosphate synthetase I gene controls its tissue-specific, developmental, and hormonal regulation in vivo. *J Biol Chem*. 1996;271(49):31243-31250.
117. Christoffels VM, van den Hoff MJB, Moorman AFM, Lamers WH. The far-upstream enhancer of the carbamoylphosphate synthetase I gene is responsible for the tissue specificity and hormone inducibility of its expression. *J Biol Chem*. 1995;270(42):24932-24940.
118. Hoogenkamp M, Stallen JMP, Lamers WH, Gaemers IC. In vivo footprinting of the carbamoylphosphate synthetase I cAMP-response unit indicates important roles for FoxA and PKA in formation of the enhanceosome. *Biochimie*. 2006; 88(10):1357-1366. doi:10.1016/j.biochi.2006.06.009
119. Schoneveld OJ, Gaemers IC, Das AT, et al. Structural requirements of the glucocorticoid-response unit of the carbamoylphosphate synthase gene. *Biochem J*. 2004;382(Pt 2):463-470. doi:10.1042/BJ20040471
120. Schoneveld OJ, Gaemers IC, Hoogenkamp M, Lamers WH. The role of proximal-enhancer elements in the glucocorticoid regulation of carbamoylphosphate synthetase gene transcription from the upstream response unit. *Biochimie*. 2005;87(11): 1033-1040. doi:10.1016/j.biochi.2005.02.015
121. Chen Z, Tang N, Wang X, Chen Y. The activity of the carbamoyl phosphate synthase 1 promoter in human liver-derived cells is dependent on hepatocyte nuclear factor 3-beta. *J Cell Mol Med*. 2017;21(9):2036-2045. doi:10.1111/jcmm.13123
122. Chen JL, Zhang Z-H, Li B-X, Cai Z, Zhou Q-H. Bioinformatic and functional analysis of promoter region of human SLC25A13 gene. *Gene*. 2019;693:69-75. doi:10.1016/j.gene.2019.01.023
123. Convertini P, Todisco S, de Santis F, et al. Transcriptional regulation factors of the human mitochondrial aspartate/glutamate carrier gene, isoform 2 (SLC25A13): USF1 as basal factor and FOXA2 as activator in liver cells. *Int J Mol Sci*. 2019;20(8):1888. doi:10.3390/ijms20081888
124. Kimura A, Nishiyori A, Murakami T, et al. Chicken ovalbumin upstream promoter-transcription factor (COUP-TF) represses transcription from the promoter of the gene for ornithine transcarbamylase in a manner antagonistic to hepatocyte nuclear factor-4 (HNF-4). *J Biol Chem*. 1993;268(15): 11125-11133.
125. Kimura T, Christoffels VM, Chowdhury S, et al. Hypoglycemia-associated hyperammonemia caused by impaired expression of ornithine cycle enzyme genes in C/EBPalpha knockout mice. *J Biol Chem*. 1998;273(42):27505-27510.
126. Nishiyori A, Tashiro H, Kimura A, et al. Both HNF-4 and C/EBP beta are required for liver-specific activity of the ornithine transcarbamylase enhancer. *J Biol Chem*. 1994; 269(2):1323-1331.
127. Inoue Y, Hayhurst GP, Inoue J, Mori M, Gonzalez FJ. Defective ureagenesis in mice carrying a liver-specific disruption of hepatocyte nuclear factor 4alpha (HNF4alpha). HNF4alpha regulates ornithine transcarbamylase in vivo. *J Biol Chem*. 2002;277(28):25257-25265.
128. Su Y, Chen Z, Yan L, et al. Optimizing combination of liver-enriched transcription factors and nuclear receptors simultaneously favors ammonia and drug metabolism in liver cells. *Exp Cell Res*. 2018;362(2):504-514. doi:10.1016/j.yexcr.2017.12.015
129. Caldovic L, Morizono H, Yu X, et al. Identification, cloning and expression of the mouse N-acetylglutamate synthase gene. *Biochem J*. 2002;364(Pt 3):825-831.
130. Ryall J, Rachubinski RA, Nguyen M, Rozen R, Broglie KE, Shore GC. Regulation and expression of carbamyl phosphate synthetase I mRNA in developing rat liver and Morris hepatoma 5123D. *J Biol Chem*. 1984;259(14):9172-9176.
131. Surh LC, Beaudet AL, O'Brien WE. Molecular characterization of the murine argininosuccinate synthetase locus. *Gene*. 1991;99(2):181-189. doi:10.1016/0378-1119(91)90125-u
132. Goutal I, Fairand A, Husson A. Expression of the genes of arginine-synthesizing enzymes in the rat kidney during development. *Biol Neonate*. 1999;76(4):253-260. doi:10.1159/000014166
133. del Arco A, Morcillo J, Martínez-Morales JR, et al. Expression of the aspartate/glutamate mitochondrial carriers aralar1 and citrin during development and in adult rat tissues. *Eur J Biochem*. 2002;269(13):3313-3320. doi:10.1046/j.1432-1033.2002.03018.x
134. Ryall JC, Quantz MA, Shore GC. Rat liver and intestinal mucosa differ in the developmental pattern and hormonal regulation of carbamoyl-phosphate synthetase I and ornithine carbamoyl transferase gene expression. *Eur J Biochem*. 1986; 156(3):453-458.
135. Camacho JA, Rioseco-Camacho N, Andrade D, Porter J, Kong J. Cloning and characterization of human ORNT2: a second mitochondrial ornithine transporter that can rescue a defective ORNT1 in patients with the hyperornithinemia-hyperammonemia-homocitrullinuria syndrome, a urea cycle disorder. *Mol Genet Metab*. 2003;79(4):257-271. doi:10.1016/s1096-7192(03)00105-7
136. Iyer RK, Yoo PK, Kern RM, et al. Mouse model for human arginase deficiency. *Mol Cell Biol*. 2002;22(13):4491-4498. doi:10.1128/mcb.22.13.4491-4498.2002
137. Sinasac DS, Moriyama M, Jalil MA, et al. Slc25a13-knockout mice harbor metabolic deficits but fail to display hallmarks of adult-onset type II citrullinemia. *Mol Cell Biol*. 2004;24(2):527-536. doi:10.1128/MCB.24.2.527-536.2004
138. Madiraju AK, Alves T, Zhao X, et al. Argininosuccinate synthetase regulates hepatic AMPK linking protein catabolism and ureagenesis to hepatic lipid metabolism. *Proc Natl Acad Sci USA*. 2016;113(24):E3423-E3430. doi:10.1073/pnas.1606022113
139. Gonzalez-Gonzalez L, Gallego-Gutiérrez H, Martin-Tapia D, et al. ZO-2 favors hippo signaling, and its re-expression in the steatotic liver by AMPK restores junctional sealing. *Tissue Barriers*. 2022;10(2):1994351. doi:10.1080/21688370.2021.1994351
140. Hashimoto T, Urushihara Y, Murata Y, Fujishima Y, Hosoi Y. AMPK increases expression of ATM through transcriptional

factor Sp1 and induces radioresistance under severe hypoxia in glioblastoma cell lines. *Biochem Biophys Res Commun*. 2022;590:82-88. doi:10.1016/j.bbrc.2021.12.076

141. Yoon JC, Puigserver P, Chen G, et al. Control of hepatic gluconeogenesis through the transcriptional coactivator PGC-1. *Nature*. 2001;413(6852):131-138. doi:10.1038/35093050
142. Cunningham JT, Rodgers JT, Arlow DH, Vazquez F, Mootha VK, Puigserver P. mTOR controls mitochondrial oxidative function through a YY1-PGC-1alpha transcriptional complex. *Nature*. 2007;450(7170):736-740. doi:10.1038/nature06322
143. Puigserver P, Wu Z, Park CW, Graves R, Wright M, Spiegelman BM. A cold-inducible coactivator of nuclear receptors linked to adaptive thermogenesis. *Cell*. 1998;92(6):829-839. doi:10.1016/s0092-8674(00)81410-5
144. Owusu-Ansah M, Guptan N, Alindogan D, Morizono M, Caldovic L. NAGS, CPS1, and SLC25A13 (citrin) at the crossroads of arginine and pyrimidines metabolism in tumor cells. *Int J Mol Sci*. 2023;24(7):6754. doi:10.3390/ijms24076754
145. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-delta delta C(T)) method. *Methods*. 2001;25(4):402-408. doi:10.1006/meth.2001.1262
146. Renga B, Mencarelli A, Cipriani S, et al. The nuclear receptor FXR regulates hepatic transport and metabolism of glutamine and glutamate. *Biochim Biophys Acta*. 2011;1812(11):1522-1531. doi:10.1016/j.bbadis.2011.06.009
147. Murakami T, Nishiyori A, Takiguchi M, Mori M. Promoter and 11-kilobase upstream enhancer elements responsible for hepatoma cell-specific expression of the rat ornithine transcarbamylase gene. *Mol Cell Biol*. 1990;10(3):1180-1191.
148. Anderson GM, Freytag SO. Synergistic activation of a human promoter in vivo by transcription factor Sp1. *Mol Cell Biol*. 1991;11(4):1935-1943. doi:10.1128/mcb.11.4.1935-1943.1991
149. Jinno Y, Matuo S, Nomiyama H, Shimada K, Matsuda I. Novel structure of the 5′ end region of the human argininosuccinate synthetase gene. *J Biochem*. 1985;98(5):1395-1403. doi:10.1093/oxfordjournals.jbchem.a135407
150. Matsubasa T, Takiguchi M, Matsuda I, Mori M. Rat argininosuccinate lyase promoter: the dyad-symmetric CCAAT box sequence CCAATTGG in the promoter is recognized by NF-Y. *J Biochem*. 1994;116(5):1044-1055. doi:10.1093/oxfordjournals.jbchem.a124626
151. Chowdhury S, Gotoh T, Mori M, Takiguchi M. CCAAT/enhancer-binding protein beta (C/EBP beta) binds and activates while hepatocyte nuclear factor-4 (HNF-4) does not bind but represses the liver-type arginase promoter. *Eur J Biochem*. 1996;236(2):500-509. doi:10.1111/j.1432-1033.1996.00500.x
152. Takiguchi M, Mori M. In vitro analysis of the rat liver-type arginase promoter. *J Biol Chem*. 1991;266(14):9186-9193.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Caldovic L, Ahn JJ, Andricovic J, et al. Datamining approaches for examining the low prevalence of N-acetylglutamate synthase deficiency and understanding transcriptional regulation of urea cycle genes. *J Inherit Metab Dis*. 2024;47(6):1175-1193. doi:10.1002/jimd.12687