VIEWS & REVIEWS OPEN ACCESS

Genotype-structure-phenotype relationships diverge in paralogs *ATP1A1*, *ATP1A2*, and *ATP1A3*

Kathleen J. Sweadner, PhD, Elena Arystarkhova, PhD, John T. Penniston, PhD, Kathryn J. Swoboda, MD, Allison Brashear, MD, and Laurie J. Ozelius, PhD

Correspondence
Dr. Sweadner
sweadner@helix.mgh.harvard.edu

Neurol Genet 2019;5:e303. doi:10.1212/NXG.0000000000000303

Abstract

Objective

We tested the assumption that closely related genes should have similar pathogenic variants by analyzing >200 pathogenic variants in a gene family with high neurologic impact and high sequence identity, the Na,K-ATPases *ATP1A1*, *ATP1A2*, and *ATP1A3*.

Methods

Data sets of disease-associated variants were compared. Their equivalent positions in protein crystal structures were used for insights into pathogenicity and correlated with the phenotype and conservation of homology.

Results

Relatively few mutations affected the corresponding amino acids in 2 genes. In the membrane domain of *ATP1A3* (primarily expressed in neurons), variants producing milder neurologic phenotypes had different structural positions than variants producing severe phenotypes. In *ATP1A2* (primarily expressed in astrocytes), membrane domain variants characteristic of severe phenotypes in *ATP1A3* were absent from patient data. The known variants in *ATP1A1* fell into 2 distinct groups. Sequence conservation was an imperfect indicator: it varied among structural domains, and some variants with demonstrated pathogenicity were in low conservation sites.

Conclusions

Pathogenic variants varied between genes despite high sequence identity, and there is a genotype-structure-phenotype relationship in *ATP1A3* that correlates with neurologic outcomes. The absence of "severe" pathogenic variants in *ATP1A2* patients predicts that they will manifest either in a different tissue or by death in utero and that new *ATP1A1* variants will produce additional phenotypes. It is important that some variants in poorly conserved amino acids are nonetheless pathogenic and could be incorrectly predicted to be benign.

From the Department of Neurosurgery (K.J. Sweadner, E.A., J.T.P.), Center for Human Genetics Research (K.J. Swoboda), and Department of Neurology, (K.J. Swoboda, L.J.O.) Massachusetts General Hospital, Boston; and the Department of Neurology (A.B.), Wake Forest School of Medicine, Winston-Salem, NC.

Funding information and disclosures are provided at the end of the article. Full disclosure form information provided by the authors is available with the full text of this article at Neurology.org/NG.

The Article Processing Charge was funded by the authors.

Glossary

AHC = alternating hemiplegia of childhood; **APA** = aldosterone-producing adenoma; **CAPOS** = cerebellar atrophy, areflexia, pes cavus, optic nerve atrophy, and sensorineural deafness; **CMT** = Charcot-Marie-Tooth; **RDP** = rapid-onset dystonia–parkinsonism.

We expect that genes that diverged only slightly after past duplications should have similar pathogenic variants. The Na,K-ATPase gene family provides a good test data set because there are >200 different human mutations (in >800 patients or families), and the proteins and many mutations are well studied.[1–4] All known pathogenic variants are de novo or have dominant inheritance.

Na,K-ATPase is the membrane protein responsible for producing sodium and potassium ion gradients. The 3 paralogs (isoforms) of the catalytic subunit (α) in the human brain are 87%–88% identical among α1, α2, and α3. Except at the N-terminus, their proteins align without insertions or deletions of more than 1 amino acid (figure e-1, links.lww.com/NXG/A134). In mice, knockouts are lethal in utero[5] or at birth.[6,7]

ATP1A1 is expressed in most tissues, including neurons and glia in the brain. *ATP1A2* is expressed in astrocytes; skeletal, cardiac, vascular, and smooth muscle; and adipocytes. *ATP1A3* is not only in neurons but also in cardiac and several other tissues. Few pathogenic variants have been discovered in *ATP1A1* so far in Charcot-Marie-Tooth neuropathy and aldosterone-producing adenomas. In contrast, there are multiple germline neurologic diseases of *ATP1A2* and *ATP1A3*.

The Na,K-ATPase paralogs have the same fundamental ion transport mechanism.[8] Sodium-bound and potassium-bound conformations are shown in different crystal structures (figure e-2, links.lww.com/NXG/A134). Variant locations in crystal structures can sometimes be related to specific enzyme functions. Mutations can produce effects on Na,K-ATPase activity, ion affinity, ion leakage, or biosynthesis. We investigated whether there are correlations between pathogenic variants (missense and 1-codon deletions), their positions in protein structure, evolutionary conservation, and the spectrum of disease phenotypes.

Methods

Sources of disease variants

The literature was searched for mutations by gene symbol. Disease variants in *ATP1A1* were from aldosterone-producing adenomas (APAs)[3] and cases of Marie-Charcot-Tooth type 2.[4] Variants in *ATP1A2* were from familial (FHM2), sporadic hemiplegic migraine, and epilepsy syndromes.[2,9] Variants in *ATP1A3* included alternating hemiplegia of childhood (AHC), rapid-onset dystonia–parkinsonism (RDP), and other milder and more severe neurologic manifestations found by genome sequencing of undiagnosed patients.[1,10,11] Some variants in *ATP1A3* were from our own patients, and a few alternative codons for known mutations were from ClinVar ncbi.nlm.nih.gov/clinvar/). The reference sequences were NM_000701.7 (*ATP1A1*); NM_000702.3 (*ATP1A2*); and NM_152296.4 (*ATP1A3*). Many suspected mutations have been tested in the laboratory for pathogenicity, but not all. The criteria for including variants in the analysis are detailed in supplementary information. All variants that met the criteria were included regardless of diagnosis.

Analysis of structure and conservation

Na,K-ATPase protein crystal structures in Na^+-bound and K^+-bound conformations at the highest available resolutions, 3WGU, 3KDP, and 2ZXE,[12–14] were from the PDB database (rcsb.org). The positions of variants were studied in Swiss PDB Viewer 4.1 based on the homology of the aligned sequences (figure e-1, links.lww.com/NXG/A134). Sequence conservation was computed with the ConSurf server (consurf.tau.ac.il/2016/) with default parameters, using the structure of the pig α1 Na,K-ATPase (3KDP) and its sequence as input. ConSurf searches for homologous sequences, and highly similar sequences are removed so that a broad spectrum of evolution is covered[15]; more detail is in supplementary data and figure e-3. The resulting conservation scores were graphed with GraphPad Prism with aligned human α1, α2, and α3.

Standard protocol approvals, registrations, and patient consents

No human experimentation was performed. Written informed consent was obtained from all patients or their guardians for cases not taken from the literature.

Data availability

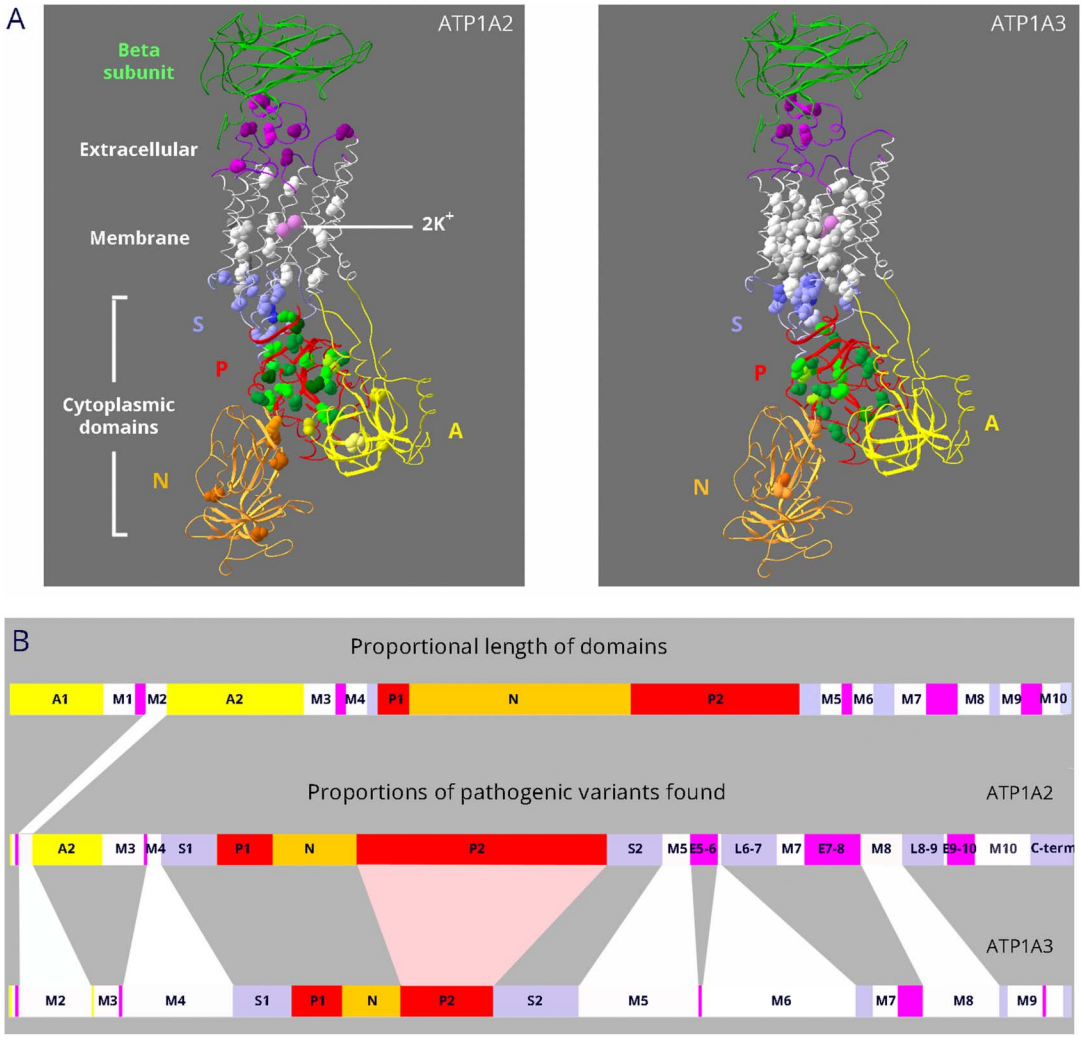
All variants are listed in supplementary information tables links.lww.com/NXG/A134.

Results

Variant distributions in *ATP1A2* and *ATP1A3*

The premise that high sequence conservation predicts variant pathogenicity implies that paralogs with high sequence identity should have similar variant distributions in the protein structure. *ATP1A2* and *ATP1A3* have the most disease-associated variants for comparison (figure 1A): 77 in *ATP1A2* and 130 in *ATP1A3* at this writing, including alternative substitutions (e.g., p.D801 to p.D801N and p.D801Y). There are striking differences, however, in the proportion of disease

Figure 1 Dissimilar mutation distributions in 88% identical ATP1A2 and ATP1A3 proteins



(A) Ribbon diagrams of Na,K-ATPase in the K^+-bound conformation. Bound K^+ is shown as pink spheres. Extracellular portions include the β-subunit (green; transmembrane portion removed for clarity) and extracellular loops of the α-subunit (magenta). The 10 transmembrane α-helices are white. The cytoplasmic domains are the stalk (S, lavender), phosphorylation domain (P, red), nucleotide binding domain (N, gold), and actuator domain (A, yellow). Spacefill residues are the backbones of each amino acid mutated in ATP1A2 or ATP1A3. Color shading is varied to help distinguish different amino acids. For ATP1A2, 76 different DNA variants occur in 66 amino acids, i.e., 10 amino acids have 2 alternative codon changes. Three of the mutations are single amino acid deletions. For ATP1A3, 130 different DNA variants occur in 86 amino acids; 5 are single amino acid deletions, and there are 25 amino acids with 2–6 alternative codon changes. All 18 overlapping residue pairs (including 3 where one was found in ClinVar) were identical before the mutation. In 6 pairs, the amino acid change was different; in 8, it was the same; and in 4, both identical and different substitutions were found (i.e., alternative codon changes). (B) Linear diagram of the domain substructure. Both the A domain (yellow) and the P domain (red) are composed of 2 parts that are separated in the linear structure. As in (A), the extracellular loops are magenta. The stalk domain comprises 5 separate parts: the cytoplasmic extensions of the M4 and M5 transmembrane spans S1 and S2; the short intracellular loops L6-7 and L8-9; and the C-terminus (lavender). For ATP1A2 and ATP1A3, domain lengths are proportional to the number of distinct variants found (including alternative codon changes). Although variants are found in most domains, ATP1A2 has an excess in the P domain, and ATP1A3 has an excess in the membrane domain.

variants found in different protein domains of ATP1A2 and ATP1A3 (figure 1B). Such differences would not be predicted by shared enzyme mechanisms and high sequence identity. Disease-causing variants in ATP1A2 and ATP1A3 should frequently appear at the corresponding amino acids. Instead, of 154 mutated amino acid positions in the 2 genes (68 in ATP1A2 and 86 in ATP1A3), only 18 mutated amino acids overlapped.

Although more such overlap should emerge with time, the low level of overlap is not simply that there are not enough cases yet: it contrasts sharply with the number of recurring variants. First, there were many cases in which alternative codons produced different substitutions of an amino acid: 25 amino acids in ATP1A3 had alternative substitutions (up to 6), and 13 amino acids in ATP1A2 had 2 alternative substitutions. Second, for ~600 independent occurrences of variants in *ATP1A3* patients (a family is 1 occurrence), 62% have recurred twice or more, and of those, more than half have recurred 5 or more times. *ATP1A2* has many fewer reported occurrences (~160), but 20% of their mutations recurred. Of the 18 pairs of variants overlapping in *ATP1A2* and *ATP1A3*,

none involved highly recurrent mutations. These data suggest that there is little bias toward crucial residues shared by ATP1A2 and ATP1A3 but strong bias for other residues.

Singletons and families in the literature were ascertained by symptoms suspicious of a genetic origin. Separate clinical diagnoses such as RDP and AHC were independently linked to *ATP1A3*.[10,16] The divergence in structure distribution suggests the existence of genotype-structure-phenotype relationships. Additional phenotypes for each gene may yet be discovered.

Structure-phenotype relationship for *ATP1A3*

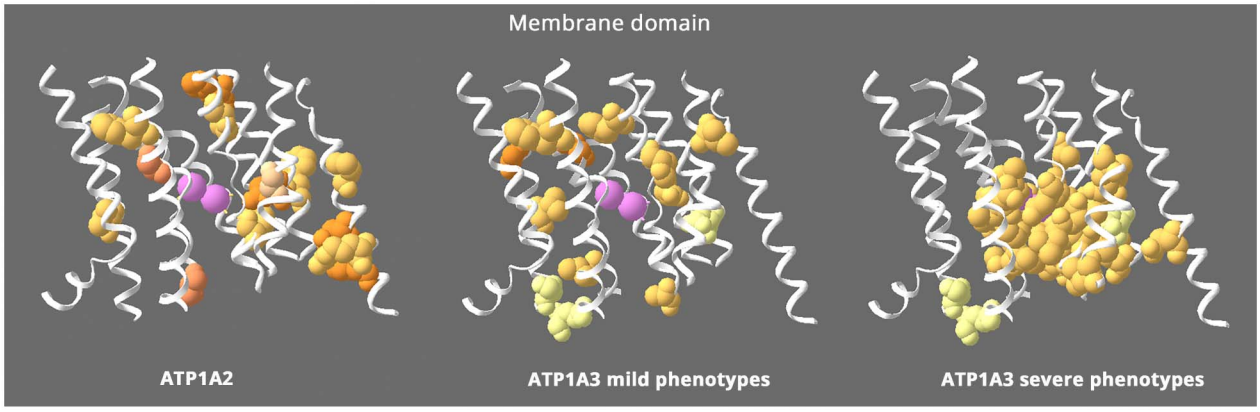
Human disease variants in *ATP1A2* and *ATP1A3* produce neurologic syndromes with onset at birth, in childhood, or in adulthood, including familial hemiplegic migraine (FHM2 [MIM:602481]),[2] RDP (MIM:128235),[1,29] AHC (MIM: 614820),[1,2] and cerebellar ataxia, areflexia, pes cavus, optic nerve atrophy, and sensorineural deafness (CAPOS [MIM: 601338]).[11] Disease variants in *ATP1A2*, expressed in astrocytes, produce FHM2 phenotypes associated with spreading depression or spreading excitation: migraine, hemiplegia, and seizures.[2] Symptom severity has not been systematically studied for *ATP1A2* mutations,[2,17] but *ATP1A3* disease variants can be divided by severity.[10,11,18–20] The most severe are defined by onset in infancy and include AHC,[1] severe infantile epilepsy,[18,21] and severe hypotonia.[22] Milder phenotypes have onset in children and adults and include RDP, CAPOS, relapsing encephalopathy and cerebellar ataxia[19] also called fever-induced weakness and encephalopathy,[23] and rapid-onset ataxia.[20,24] Ages at symptom onset were taken from original reports.

Figure 2 highlights a structure-phenotype relationship in the membrane domain. ATP1A2 and the milder phenotypes in ATP1A3 have a scattered distribution, but the ion binding site is almost devoid of variants. In contrast, the variants with severe phenotypes in ATP1A3 are clustered around the ion binding sites, as noted previously for AHC.[25–27] Three residues in ATP1A3 with either mild or severe manifestations are pale yellow and not close to the ions. The absence of known ATP1A2 mutations near the ions suggests that their effects are too severe to manifest as hemiplegic migraine.

The P domain, the site of ATP hydrolysis, is highly evolutionarily conserved. Figure 3A dissects this domain because its structure is otherwise difficult to appreciate. Two distant stretches of amino acids (P1 and P2) interdigitate in a twisted sheet of β-sheet strands encircled by α-helices. In figure 3A, the top row shows an end view of the β-sheet, the α-helices that surround it, and the 2 parts combined. The active site has an aqua magnesium ion. The second row shows a side view of the same structures. Figure 3B shows the distributions of mutated amino acids. In the β-sheet, all but 3 are clustered at the face of the active site. In the ring of α-helices, all mutations except 1, in magenta, are away from the active site. The helix mutations are confined to the central 5 of the 9 helices. The more distal helices and strands so far have no mutations.

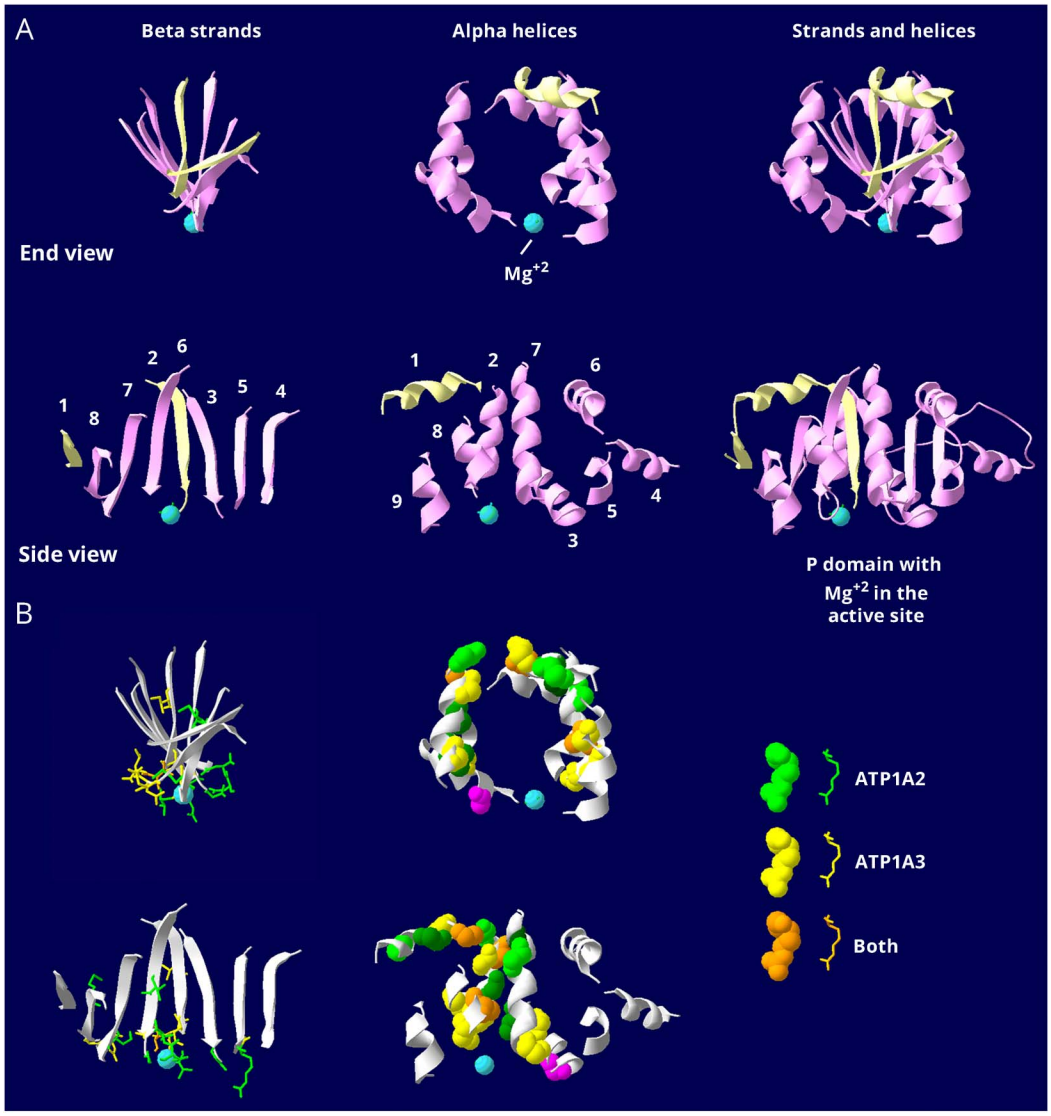
In Figure 3, color distinguishes the variants found in ATP1A2, ATP1A3, or both. All have similar patterns of distribution in the P domain. Figure e-4 (links.lww.com/NXG/A134) compares the mild and severe phenotype variants of ATP1A3. Any suggestion of differences in location would need to be

Figure 2 Structure-phenotype relationship in the membrane domain



This figure shows the transmembrane α-helices, omitting the portions that extend into the extracellular and intracellular spaces. All the ATP1A2 membrane mutations are shown (left), whereas ATP1A3's are divided into 2 groups: milder ATP1A3 phenotypes (middle) with onset from childhood to adult and severe ATP1A3 mutations (right) with onset in infancy. ATP1A2 and the milder ATP1A3 phenotypes have similar distributions that appear to exclude the residues right around the ions (pink = potassium ions). In contrast, the severe ATP1A3 mutations are usually close to the ions. Almost 70% of the severe mutations are in contiguous stretches in M4, M5, and M6 near the ions. In contrast, 100% of the mild mutations are not adjacent to any other mutations, and the few in M4, M5, and M6 are on either side of the contiguous stretches. The distributions highlight 2 unique features: that mild and severe ATP1A3 mutations have different distributions and that ATP1A2 mutations all look like "mild". In fact, hemiplegic migraine seldom has onset in infancy.[17] Equivalent "severe" mutations of ATP1A2 have not been found. Variants of 3 ATP1A3 residues in pale yellow can produce either mild or severe phenotypes. Fourteen ATP1A3 residues altogether have produced both RDP and mild AHC or an intermediate phenotype. Eleven of those were recurrent (with the same or different substitution) or appeared also in *ATP1A1* or *ATP1A2* patients. Here, 2 had different substitutions in patients with differing phenotypes (S137Y or S137F severe,[10] S137del mild[53]; Q140L severe,[10] Q140H mild[54]), and the third, D923N, reduces affinity for Na^{+}[55] and presents as a continuum between AHC and RDP,[56,57] severe or mild in the same family.[58]

Figure 3 Similar but restricted distributions of pathogenic variants in the P domain



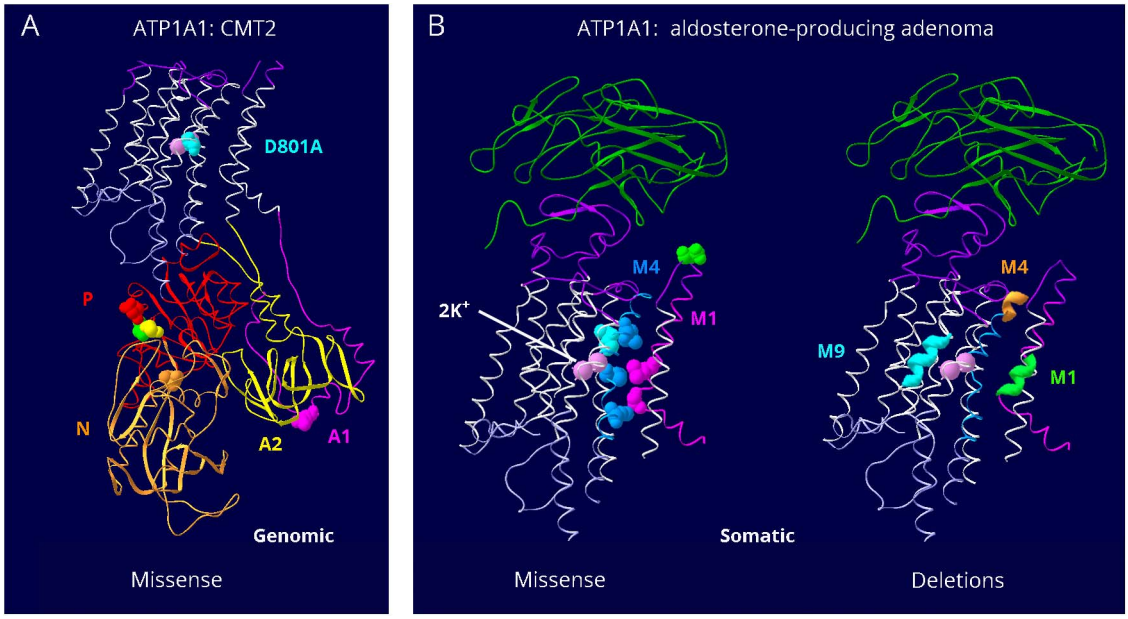
The P domain is the most conserved in the gene family. It is covalently phosphorylated during transport, and the Mg^{2+} ion (aqua) is at the active site. (A) The P domain consists of 2 halves that are separated in the linear structure but that interdigitate in a complex way when the protein folds. The first half, P1 in figure 1B, is shown in pale yellow, and the second half, P2 in figure 1B, in pink. There is a twisted β-sheet (left views) surrounded by α-helices (middle views). Strands and helices are numbered in the order found in the sequence. (B) The known mutations are displayed with color coding as indicated. Stick view is used for mutations in the β-sheet to avoid overcrowding. Note how most of the side chains of the mutated residues extend down into the active site surface. Spacefill view of the mutations, without side chains, is used for the α -helices. Note how the mutations are confined to just 5 helices: helix 1, which is associated with the stalk domain, and 4 helices that form a belt around the middle of the P domain: 2, 3, 7, and 8. There are no mutations in the end strands (1 and 4) or the end helices (4, 5, 6, and 9), which may be less important to function. The magenta mutation in an α-helix is the only helix mutation near (but not in) the active site. It is ATP1A3 T613M, a recurrent RDP mutation. The data indicate that the ATP1A2 and ATP1A3 distributions of mutations are very similar in the P domain, but there are 2 types of mutations. Most of the mutations in the β-sheet are at the active site. Most of the ones in the α-helices are distant from the active site. There was no correlation with phenotype severity.

supported by further evidence. Figure e-5 compares the ATP1A2 and ATP1A3 variants in the S domain. On the whole, P domain and S domain mutation distributions do not differ systematically between ATP1A2 and ATP1A3.

Two types of variants in *ATP1A1*

Presumably because *ATP1A1* is essential in early embryonic life, not many disease variants have been discovered. Those fall into 2 groups with different characteristics. Seven ATP1A1 putatively pathogenic variants found in Charcot-Marie-Tooth (CMT2) axonal sensorimotor neuropathy are in the cytoplasmic A, P, and N domains, and 1 is at an ion binding residue frequently mutated in ATP1A3 but with a less disruptive substitution (figure 4A). Adrenal adenomas, like many cancers, have a high somatic mutation rate, and APAs have distinctly different ATP1A1 variants (short deletions and missense), all in the membrane domain (figure 4B).[28–31] Those tested for function showed reduced Na,K-ATPase activity, cell depolarization, or a modest inward leak current.[28,29,32–34] The amino acids of the most recurrent

Figure 4 Distinct groups of ATP1A1 pathogenic variants



There are only a handful of known mutations for ATP1A1, but they fall in 2 clearly different groups. (A) Seven germline mutations were found in CMT2 patients.[4] One is at the interface between the 2 halves of the A domain (A1 dark pink, A2 yellow), where it may interfere with A domain stability. One is at the site equivalent to D801N, the most common AHC pathogenic variant, but the change is from aspartate to alanine, presumably a milder substitution. The other 5 variants are close to the boundary of P and N domains. They do not appear fundamentally different from mutations in ATP1A2 or ATP1A3. I592T in ATP1A1 (orange) is the residue equivalent to I589T in ATP1A2.[59] (B) The other group is somatic mutations found in aldosterone-producing adenomas.[3] (*Left*) Six of the missense mutations all lie along the ion path through the protein, and some alter amino acids that are believed to be gates. The seventh is at the top of the extracellular mouth (green). Two of them, V332G and L337M, are less disruptive versions of mild pathogenic variants in ATP1A3: V322D and L327del. (*Right*) Multiple different short deletions (of 2–6 amino acids) appear only in the 3 stretches of amino acids highlighted in membrane spans 1, 4, and 9. Such deletions have not been reported for ATP1A2 or ATP1A3. These somatic mutations are in a class by themselves. They are believed to generate slight leaks, large enough to lead to aldosterone production but not large enough to kill the cells.

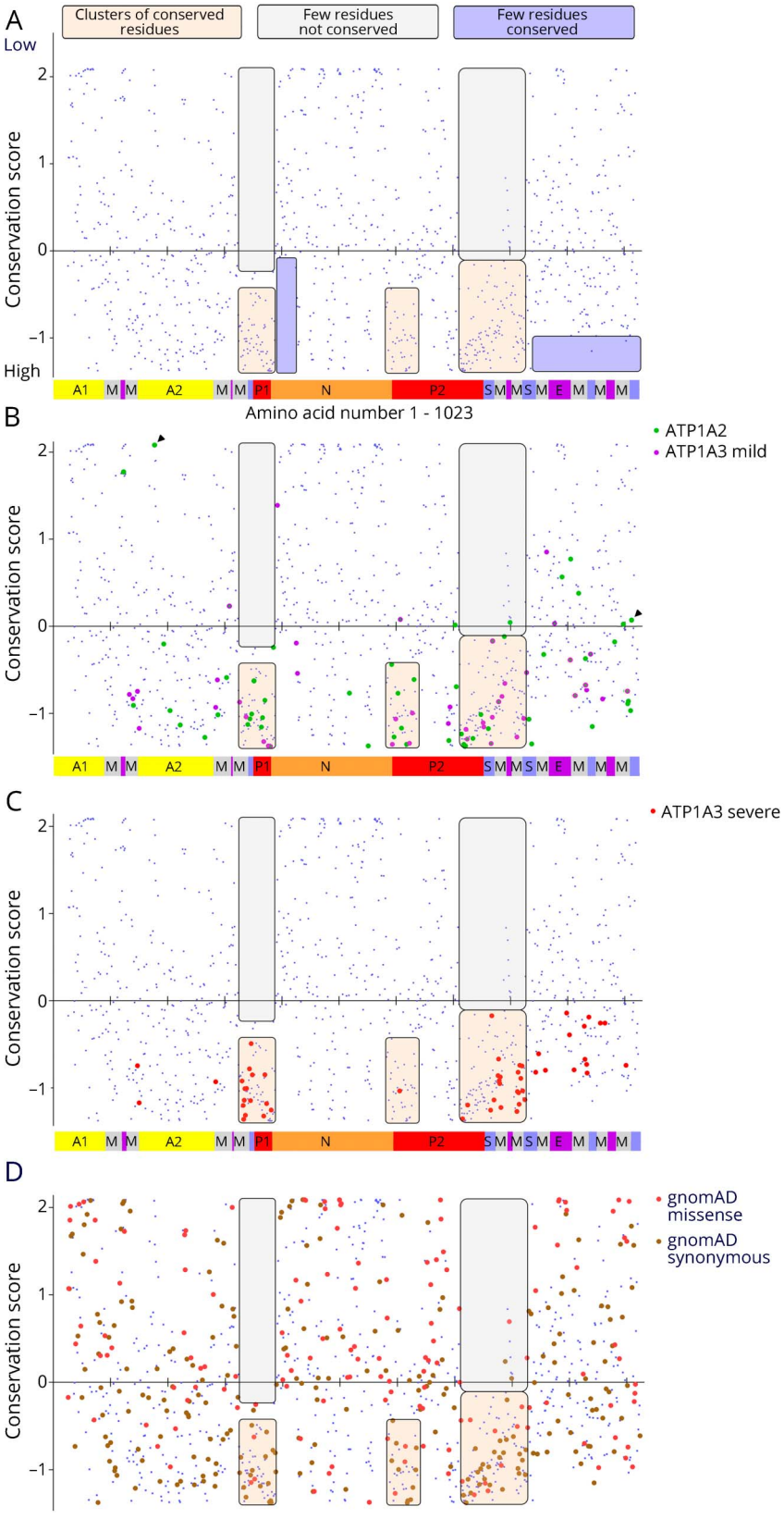
adenoma variants, p.Gly99Arg, p.Leu104Arg, and p.Val332Gly, had previously been shown to line the ion pathway or function as gates: their chemical modification blocked palytoxin-induced currents caused by the toxin's ability to bind to Na,K-ATPase and wedge it open.[35] In sum, the CMT2 ATP1A1 disease variants are likely to cause a mild loss of function, and the adenoma variants may cause a gain of function with limited currents leading to aberrant aldosterone production.

Pathogenic variants without evolutionary conservation

To determine whether mutations are in amino acids with high evolutionary conservation, we used ConSurf to estimate conservation based on 94 nonredundant homologous sequences from the P-type ATPase superfamily, from vertebrates to yeast and bacteria.[36] Figure 5A shows the conservation scores of all residues as small blue symbols, and the bar shows the locations of domains as in figure 1B. The 2D presentation highlights that for much of the protein, high conservation and low conservation residues are closely interspersed in the linear sequence. Shaded boxes highlight 2 nonrandom features. First, there are 3 large stretches of conserved residues (tan) with practically no low conservation residues (off-white). In the right off-white box, at around amino acid 800, there is a vertical streak of less-conserved residues that correspond to a loop in the extracellular domain, which is magenta on the domain bar. Second, there are 2 stretches of amino acids that are depleted of high conservation residues (lavender boxes). Because conservation was calculated using sequences that are related but have divergent functions, we infer that these regions, rather than being of lesser importance, evolved for functions not shared by more distant homologs: regions of special interest for *ATP1A1*, *ATP1A2*, and *ATP1A3*.

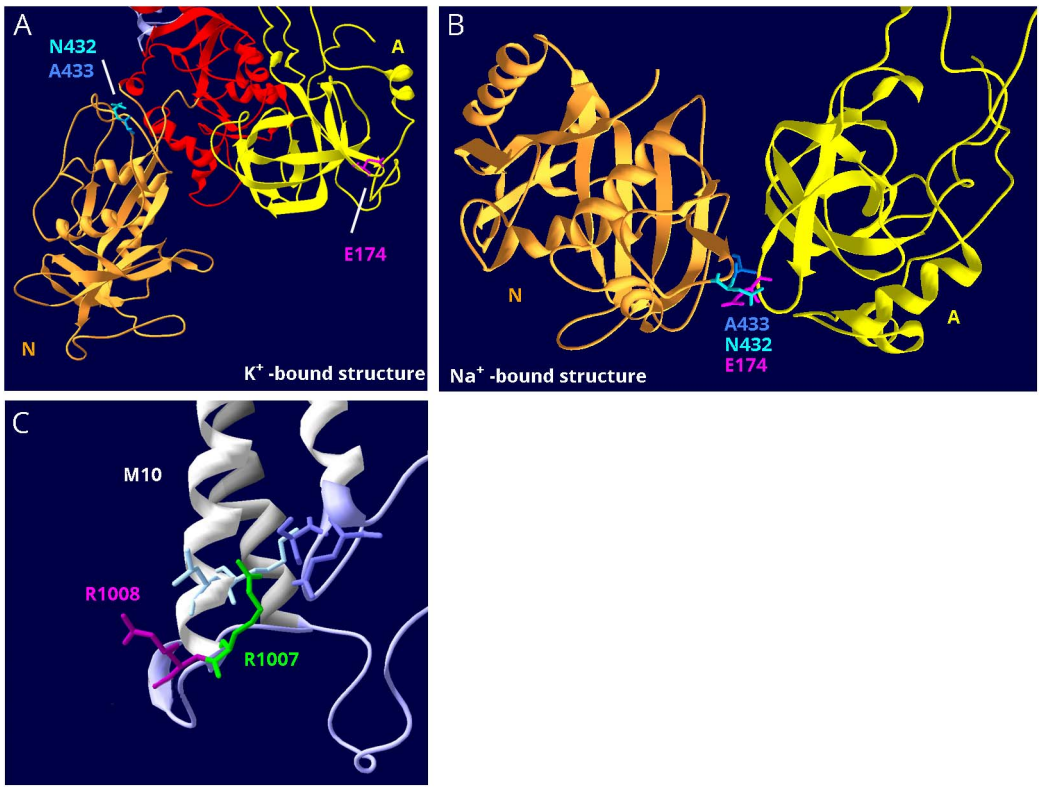
The average conservation scores for missense variants found in patients were similar for all 3 genes at the 30th percentile (data not shown). Two-thirds of the mutated residues were in the highest 20%, and only 3% were in the lowest 20%. The conservation scores were graphed to look for patterns unique to *ATP1A2* and mild-phenotype *ATP1A3* (figure 5B) and severe *ATP1A3* (figure 5C). A difference in the distribution of variants is clear both in the location and conservation score, with the severe-phenotype *ATP1A3* variants being more restricted and all falling below zero on the graph (5C). Another difference that cannot be appreciated on the graphs is that 65% of the severe *ATP1A3* variants are recurrent, whereas the proportion is 38% for mild *ATP1A3* variants and 20% for *ATP1A2*. The region at the C-terminus with no very high conservation residues has many mutations (figure 5, B and C).

Figure 5 Evolutionary sequence conservation



Human ATP1A1, ATP1A2, and ATP1A3 amino acid sequences were aligned with *Sus scrofa* ATP1A1 (the species used in protein crystal structures 3KDP and 3WGU), and each residue affected by a disease variant was assigned the conservation score calculated for ATP1A1. (A) The individual conservation scores of all aligning amino acids are shown with score on the Y axis and amino acid sequence (1–1,023) on the X axis. As seen in figure e-1 (links.lww.com/NXG/A134), there is no alignment at the N-terminus, so N-terminal residues do not appear. Each blue spot is one amino acid, and the Na,K-ATPase domains are illustrated in the bar at the bottom. Some patterns of conservation can be discerned, highlighted by transparent boxes: two regions with few or no highly conserved amino acids (lavender) and 3 regions of conserved residues (tan) with practically no low conservation residues (off-white). (B) The ATP1A2 variants (green) and the mild ATP1A3 variants (magenta) have similar distributions. There are 14 cases (~15%) where homology scores were in the less-conserved half. The C-terminal region has many pathogenic variants in both genes despite lacking the highest conservation scores. Residues E174 and R1008 from figure 6 are marked with arrowheads. (C) The severe ATP1A3 variants had a more restricted distribution. It can be seen that many align with the membrane spans. Although their conservation scores varied widely, none were in the less-conserved half. Of the points that do not lie in one of the 3 clusters of high conservation, 8 were shared by mild and severe *ATP1A3* phenotypes (visible as 2 colors per symbol at high magnification). This suggests that intermediate phenotypes tend to have less restricted structure distributions than the most severe phenotypes. (D) All the synonymous (brown; n = 281) and missense (coral pink; n = 145) *ATP1A3* variants from gnomAD (138,632 sequences) are plotted to compare to the clustering of mutations.

Figure 6 Structure-based interpretation of variants of uncertain significance



(A) E174K; N, P, and A domains are visible. In the K^+-bound conformation, the N (gold) and A (yellow) domains are far apart. (B) In the Na^+-bound conformation, the N and A domains are in contact. The P domain in the background was hidden for clarity. E174K (magenta) is a variant in ATP1A2 that has the poorest conservation score in figure 5B, and it is predicted to be tolerated by SIFT.[38] In the K^+-bound structure, it is fully exposed to the cytoplasm (A), but in the Na^+-bound structure, it is the point of closest contact between the A domain (yellow) and N domain (gold). There it interacts with 2 adjacent residues, an asparagine N432 and an alanine A433 (blue and aqua). They are also shown in (A). (C) R1008W; close-up of M10, M7, and M8 where these helices terminate in parts of the S domain (lavender). R1008W (magenta) is a variant in ATP1A2 in a patient with compound heterozygosity with R548C. Because the adjacent variant R1007W (green) is known to be pathogenic, it was suggested that R1008W might contribute to the severity.[39] R1007 has a better conservation score, but R1008's score is still in the range of other pathogenic variants (close to zero). In the structure, R1007 points into the protein where it makes close association with 2 domains: two residues in M10 (light blue) and 2 residues in L7-8 (lavender), including the protein kinase A phosphorylation site, S933. In contrast, R1008 points out into the cytoplasm, where a substitution may be tolerated.

Normally, one would expect that variants with low conservation scores are benign. In the ATP1A2 and mild ATP1A3 variants in figure 5B, 14 amino acids have positive (poor) conservation scores. Of these, 10 were supported as pathogenic by laboratory testing of function, recurrence as an alternative amino acid, or occurrence in a second gene, and 3 of those were de novo. There was no pattern in their scores. Three others had insufficient data, and a fourth was published as possibly disease causing and is analyzed further.

The missense and synonymous variants in gnomAD gave no obvious pattern (figure 5D). More than half (58%) of the synonymous variants were in the higher half of the conservation score, whereas only 39.3% of missense variants were ($p = 10^{-5}$ t test, 2 tailed, equal variance) consistent with genetic constraint.

Structure data assist tricky predictions

ATP1A2 variant E174K has a conservation score of 2.077, putting it at the top of figure 5B. The variant appeared in a family that experiences migraine with side-changing paresthesias, which is on the FHM2 spectrum.[37] Pedigree data were supportive but incomplete. When expressed in *Xenopus* oocytes, no reduction of pump activity was seen,[37] but the ion concentrations used were high enough to saturate the enzyme. In later experiments using protein expressed in insect cells, it was learned that the enzyme had altered kinetics: reduced affinity for Na^+ and increased affinity for K^+.[38] Because Na^+ concentration is rate limiting for enzyme activity in vivo, this mimics reduction of activity. Figure 6, A and B illustrate how the position of E174 in the crystal structures explains the pathogenicity. E174 is at the surface of the A domain, which rotates. In the K^+-bound structure, E174 faces the cytoplasm, and the side chain is in contact with nothing (figure 6A). In the Na^+-bound form, however, it binds 2 amino acids in the N-domain (figure 6B), i.e., it is at a domain-domain interface where the change from negative to positive charge will be damaging. This explains the abnormal kinetics because the sodium conformation would be destabilized.[38] In this example, an amino acid with the worst conservation score nonetheless is legitimately pathogenic.

The second example is ATP1A2 variant R1008W, which was found together with a known pathogenic variant, R548C, in a compound heterozygous child with unusually early onset of severe epilepsy, hemiplegia, and migraine symptoms.[39] The severity of the phenotype, not seen in a family with R548C alone[40] or in the child's mother who carried the R548C mutation, suggested that both variants contributed to disease severity. An adjacent arginine with the same substitution, R1007W, previously described in a family with hemiplegic migraine and epilepsy, was shown to be pathogenic by functional studies in *Xenopus* oocytes.[41] Although R1007 has a conservation score in figure 5B of −0.967 while R1008 is +0.069, their identical substitutions suggested that the adjacent variants might both be pathogenic.[39] However, their positions in the crystal structure tell a different story (figure 6C). R1007's side chain faces into the protein and makes contacts with 2 amino acids in M10 and 2 in the S domain. One of those contacts is the protein kinase A site, and R1007 had been shown previously by mutagenesis to play a complex role in controlling Na,K-ATPase.[42] In contrast, the adjacent R1008 side chain points away from the protein, where it may associate with phospholipid headgroups.[43] Adjacent amino acids can have very different propensities to be pathogenic. In this example, a variant with some apparently credible features, nonetheless, may be benign.

Discussion

It was not expected that pathogenic variants seen in Na,K-ATPase paralogs would differ. The concepts of conservation of sequence homology and genetic constraint (intolerance of substitution) are tightly linked. The premise is that evolutionary conservation occurs because variants causing functional damage are eliminated by selection. This study documents that sets of paralogs can violate that logic at the disease level: sequence similarity between paralogs is not always a strong predictor for the discovery and interpretation of new human pathogenic variants. The Na,K-ATPase catalytic subunit gene family instead shows partial divergence in pathogenic variants. Ascertainment bias plays a role: patients must survive with the mutation, and symptoms must be uncommon enough to suggest monogenic disease. Mutations outside of these constraints would occur but be undetected.

It is fundamental that mutation phenotypes can appear in one tissue or cell type and not another because of differences in how physiology is affected. Partial loss of Na,K-ATPase activity may have more impact in a cell with high transport activity than one with low activity; in inhibitory rather than excitatory neurons; or presynaptically rather than postsynaptically. An elevation of intracellular Na^+ due to a reduction of sodium affinity may have more effect on dendrites than on action potentials or have different secondary effects on transmitter reuptake. Similar mutations in astrocytes and neurons can have different impacts on intracellular Ca^{2+}, pH, or Cl^- because of the expression of different sodium-dependent carriers. When comparing different mutations, however, variants that simply reduce activity should have a ceiling of physiologic impact; e.g., a mutation inactivating the ion binding site should have the same physiologic effect as a mutation inactivating the ATP hydrolysis site, provided both changes produce inert enzyme. Loss of function alone, however, does not explain why many *ATP1A2* and *ATP1A3* neurologic variants are biased toward specific mutations as suggested by their independent recurrence.

To understand the genotype-structure-phenotype differences displayed by *ATP1A2* and *ATP1A3* and by mild and severe *ATP1A3* mutations, we postulate a gain-of-function defect in the severe mutations. Certain mutations of Na,K-ATPase may cause a greater perturbation of physiology than pump inactivation alone by alteration of ion selectivity or stoichiometry; by depolarizing inward currents carried by protons or Na^+; by leaks through the ion pathway because of defective gating; or by protein instability or misfolding in the membrane.[28,29,33,34,44] Small inward proton currents occur in wild-type enzyme and are sometimes blocked in mutants, but their physiologic impact is still uncertain.[26] Larger leaks or an outward proton current would be disruptive for both neurons and glia and potentially toxic. The extreme toxicity of palytoxin is due to the opening of the ion pathway through Na,K-ATPase like a channel.[45] Any leaks in surviving patients should be very slow relative to palytoxin and might manifest only at physiologic temperatures or higher. The ATP1A3 variants that are tightly clustered around the ion sites would be candidates. So would sensitive spots such as G358 on the P1 helix embedded in the S domain, where a G358V mutation produced fatal infantile epilepsy[18] and G358C, S, and D produce AHC. So would the highly recurrent AHC mutations E815K and G947R that have dominant negative effects[46] and are in a network of residues affecting the cytoplasmic ion pathway.[26,47] The importance of pathogenic leaks is supported by temperature-sensitive mutations studied in Ca^{+2} ATPases in *Drosophila*[48] in addition to the aldosterone-secreting adenoma *ATP1A1* mutations.[28,34] Many other mutations, such as in the P domain, N domain, or extracellular surface, may affect only biosynthesis or activity.

One significant observation is that the membrane domain pathogenic variants causing severe phenotypes in *ATP1A3* are entirely missing from the *ATP1A2* patient population. It is not that hemiplegic migraine *ATP1A2* mutations are never fully inactivating: a number had no detectable activity in laboratory studies.[2] Even among the variants shared by *ATP1A2* and *ATP1A3*, however, none were close to the ion binding sites. The logical prediction is that severe variants of *ATP1A2* occur but have unrecognized disease consequences. For example, they may produce membrane depolarization that interferes with astrocyte roles in synapse development and remodeling,[49] and this clinically supersedes the cortical spreading depression that contributes to migraine.[2] Other possibilities are a phenotype in other *ATP1A2*-expressing tissues (skeletal,

vascular, smooth, or cardiac muscle; fat) or lethality, perinatally or in utero.

The variants of *ATP1A3* that we have called "milder" nonetheless produce disabling neurologic disease with stable, progressive, or episodic motor symptoms, often with cognitive deficits and psychiatric manifestations.[1,50,51] No pathogenic variants have been found yet, however, in the N-terminus, the A domain, the linkers to the A domain, or some of the peripheral membrane spans, suggesting that their functional impact is too small (or too great) to produce recognized diseases.

Genetically defined loss-of-function mutations (LoF: frameshifts, premature stops, and deletions) are almost absent from gnomAD for *ATP1A1* and *ATP1A3* and reduced in *ATP1A2*. This is evidence of purifying selection, and that the genes are highly constrained, however, LoF mutations are also largely missing in the patient populations. Recessive variants are missing so far from the whole gene family. This is puzzling because partial reduction of activity of such important enzymes should have pathophysiologic effects if homozygous. A possibility is that recessive variants will be found for phenotypes or risk factors that are not yet recognized. *ATP1A1* recessive variants could manifest in any tissue with high ion transport. *ATP1A3* variants might affect CNS functions already affected by dominant mutations in RDP and AHC, such as depression, behavioral and cognitive deficits, or psychosis.[50–52]

This gene family's mutational landscape shows that the spectrum of pathologic variants varies among paralogs. Important pathologic variants can occur in evolutionarily recent specializations that are not conserved at all. The mutations in *ATP1A2* and *ATP1A3* and the germline mutations in *ATP1A1* were overwhelmingly missense or deletions of 1 codon, which suggests that the mutated protein must be present for the known manifestations. It will be interesting to see whether recessive mutations have a structural distribution more representative of sequence identity in the gene family.

Author contributions

Design and execution of the study: K.J. Sweadner, E.A., J.T.P., and L.J.O. Contribution to genetic and phenotypic data and interpretation of data: K.J. Swoboda and A.B. Writing of the manuscript: K.J. Swoboda, E.A., and L.J.O. Revisions of the manuscript: all authors.

Acknowledgment

This work would not be possible without the patients and their families for their eagerness to contribute to research. The authors thank the Genome Aggregation Database (gnomAD) and the groups that provided exome and genome variant data to this resource. A full list of contributing groups can be found at gnomad.broadinstitute.org/about. They also thank the laboratories that produced crystal structures of Na,K-ATPase and Drs. S. Bressman, C. Brownstein, J. Damásio, S. DeBrosse, S. Demarest, W.B. Dobyns, L.M. Frank, S. Frucht, M. Mikati, F. Mochel, R. Schiffman, N.J.C. Smith, C. Toro, and L.A. Wolfe for sharing unpublished disease variants.

Study funding

Supported by NIH NS058949 to A.B.

Disclosure

Dr. Sweadner serves on the advisory boards of DANDRITE and the Nordic EMBL Partnership for Molecular Medicine and has received salary support from NINDS R01NS058949 (AB). Dr. Arystarkhova has received salary support from NINDS R01NS058949 (AB). Dr. Penniston reports no disclosures. Dr. Swoboda reports grants and personal fees from Biogen, unrelated to the submitted work. Dr. Brashear consults about protocol development with Revance and Ipsen and has received salary support from NINDS R01NS058949 (AB). Her conflict of interest is managed by Wake Forest School of Medicine. Dr. Ozelius reports no disclosures. Full disclosure form information provided by the authors is available with the full text of this article at Neurology.org/NG.

Publication history

Received by *Neurology: Genetics* June 1, 2018. Accepted in final form November 8, 2018.

References

1. Heinzen EL, Arzimanoglou A, Brashear A, et al. Distinct neurological disorders with *ATP1A3* mutations. Lancet Neurol 2014;13,503–514.
2. Friedrich T, Tavraz NN, Junghans C. ATP1A2 mutations in migraine: seeing through the facets of an ion pump onto the neurobiology of disease. Front Physiol 2016;7:239.
3. Azizan EA, Brown MJ. Novel genetic determinants of adrenal aldosterone regulation. Curr Opin Endocrinol Diabetes Obes 2016;23:209–217.
4. Lassuthova P, Rebelo AP, Ravenscroft G, et al. Mutations in *ATP1A1* Cause Dominant Charcot-Marie-Tooth Type 2. Am J Hum Genet 2018;102:505–514.
5. Barcroft LC, Moseley AE, Lingrel JB, Watson AJ. Deletion of the Na/K-ATPase α1-subunit gene (*Atp1a1*) does not prevent cavitation of the preimplantation mouse embryo. Mech Dev 2004;121:417–426.
6. Moseley AE, Lieske SP, Wetzel RK, et al. The Na,K-ATPase α2 isoform is expressed in neurons, and its absence disrupts neuronal activity in newborn mice. J Biol Chem 2003;278:5317–5324.
7. Moseley A, Williams MT, Schaefer TL, et al. Deficiency in Na,K-ATPase α isoform genes alters spatial learning, motor activity and anxiety in mice. J Neurosci 2007;27:616–626.
8. Blanco G. Na,K- ATPase subunit heterogeneity as a mechanism for tissue-specific ion regulation. Semin Nephrol 2005;25:292–303.
9. Prontera P, Sarchielli P, Caproni S, et al. Epilepsy in hemiplegic migraine: Genetic mutations and clinical implications. Cephalalgia 2018;38,361–373.
10. Heinzen EL, Swoboda KJ, Hitomi Y, et al. *De novo* mutations in *ATP1A3* cause alternating hemiplegia of childhood. Nat Genet 2012;44,1030–1034.
11. Brashear A, Sweadner KJ, Cook JF, Swoboda KJ, Ozelius LJ. *ATP1A3*-related neurologic disorders. 2008 Feb 7 [Updated 2018 Feb 22]. In: Adam MP, Ardinger HH, Pagon RA, et al., eds. *GeneReviews®* [Internet]. Seattle, WA: University of Washington; 1993–2019. Available at: https://www.ncbi.nlm.nih.gov/books/NBK1115/.
12. Shinoda T, Ogawa H, Cornelius F, Toyoshima C. Crystal structure of the sodium-potassium pump at 2.4 Å resolution. Nature 2009;459:446–450.
13. Kanai R, Ogawa H, Vilsen B, Cornelius F, Toyoshima C. Crystal structure of a Na^+-bound Na^+,K^+-ATPase preceding the E1P state. Nature 2013;502:201–206.
14. Morth JP, Pedersen BP, Toustrup-Jensen MS, et al. Crystal structure of the sodium-potassium pump. Nature 2007;450,1043–1049.
15. Landau M, Mayrose I, Rosenberg Y, et al. ConSurf 2005: the projection of evolutionary conservation scores of residues on protein structures. Nucleic Acids Res 2005;33:W299–W302.
16. de Carvalho Aguiar P, Sweadner KJ, Penniston JT, et al. Mutations in the Na,K-ATPase α3 gene *ATP1A3* are associated with rapid-onset dystonia parkinsonism. Neuron 2004;43:169–175.
17. Pelzer N, Haan J, Stam AH, et al. Clinical spectrum of hemiplegic migraine and chances of finding a pathogenic mutation. Neurology 2018;90,e575–e582.
18. Paciorkowski AR, McDaniel SS, Jansen LA, et al. Novel mutations in *ATP1A3* associated with catastrophic early life epilepsy, episodic prolonged apnea, and postnatal microcephaly. Epilepsia 2015;56,422–430.

19. Dard R, Mignot C, Durr A, et al. Relapsing encephalopathy with cerebellar ataxia related to an *ATP1A3* mutation. Dev Med Child Neurol 2015;57,1183–1186.
20. Sweadner KJ, Toro C, Whitlow CT, et al. ATP1A3 mutation in adult rapid-onset ataxia. PLoS One 2016;11:e0151429.
21. Sasaki M, Ishii A, Saito Y, et al. Genotype-phenotype correlations in alternating hemiplegia of childhood. Neurology 2014;82:482–490.
22. Sweney MT, Newcomb TM, Swoboda KJ. The expanding spectrum of neurological phenotypes in children with *ATP1A3* mutations, alternating hemiplegia of childhood, rapid-onset dystonia-parkinsonism, CAPOS and beyond. Pediatr Neurol 2015;52:56–64.
23. Yano ST, Silver K, Young R, et al. Fever-induced paroxysmal weakness and encephalopathy, a new phenotype of ATP1A3 mutation. Pediatr Neurol 2017;73,101–105.
24. Rosewich H, Ohlenbusch A, Huppke P, et al. The expanding clinical and genetic spectrum of ATP1A3-related disorders. Neurology 2014;82:945–955.
25. Yang X, Gao H, Zhang J, et al. *ATP1A3* mutations and genotype-phenotype correlation of alternating hemiplegia of childhood in Chinese patients. PLoS One 2014;9: e97274.
26. Holm R, Toustrup-Jensen MS, Einholm AP, Schack VR, Andersen JP, Vilsen B. Neurological disease mutations of alpha3 Na^+,K^+-ATPase: structural and functional perspectives and rescue of compromised function. Biochim Biophys Acta 2016;1857:1807–1828.
27. Clausen MV, Hilbers F, Poulsen H. The structure and function of the Na,K-ATPase isoforms in health and disease. Front Physiol 2017;8:371.
28. Azizan EA, Poulsen H, Tuluc P, et al. Somatic mutations in *ATP1A1* and *CACNA1D* underlie a common subtype of adrenal hypertension. Nat Genet 2013;45;1055–1060.
29. Beuschlein F, Boulkroun S, Osswald A, et al. Somatic mutations in *ATP1A1* and *ATP2B3* lead to aldosterone-producing adenomas and secondary hypertension. Nat Genet 2013;45:440–442.
30. Williams TA, Monticone S, Schack VR, et al. Somatic *ATP1A1, ATP2B3*, and *KCNJ5* mutations in aldosterone-producing adenomas. Hypertension 2014;63:188–195.
31. Nishimoto K, Tomlins SA, Kuick R, et al. Aldosterone-stimulating somatic gene mutations are common in normal adrenal glands. Proc Natl Acad Sci USA 2015;112: E4591–E4599.
32. Kopec W, Loubet B, Poulsen H, Khandelia H. Molecular mechanism of Na^+,K^+-ATPase malfunction in mutations characteristic of adrenal hypertension. Biochemistry 2014;53:746–754.
33. Stindl J, Tauber P, Sterner C, Tegtmeier I, Warth R, Bandulik S. Pathogenesis of adrenal aldosterone-producing adenomas carrying mutations of the Na^+/K^+-ATPase. Endocrinology 2015;156:4582–4591.
34. Meyer DJ, Gatto C, Artigas P. On the effect of hyperaldosteronism-inducing mutations in Na/K pumps. J Gen Physiol 2017;149:1009–1028.
35. Takeuchi A, Reyes N, Artigas P, Gadsby DC. The ion pathway through the opened Na^+,K^+-ATPase pump. Nature 2008;456:413–416.
36. Ashkenazy H, Abadi S, Martz E, et al. ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. Nucleic Acids Res 2016;44:W344–W350.
37. Todt U, Dichgans M, Jurkat-Rott K, et al. Rare missense variants in *ATP1A2* in families with clustering of common forms of migraine. Hum Mutat 2005;26:315–321.
38. Swarts HG, Weigand KM, Venselaar H, van den Maagdenberg AM, Russel FG, Koenderink JB. Familial hemiplegic migraine mutations affect Na,K-ATPase domain interactions. Biochim Biophys Acta 2013;1832:2173–2179.
39. Wilbur C, Buerki SE, Guella I, et al. An infant with epilepsy and recurrent hemiplegia due to compound heterozygous variants in ATP1A2. Pediatr Neurol 2017;75:87–90.
40. Lebas A, Guyant-Maréchal L, Hannequin D, Riant F, Tournier-Lasserve E, Parain D. Severe attacks of familial hemiplegic migraine, childhood epilepsy and *ATP1A2* mutation. Cephalalgia 2008;28:774–777.
41. Pisano T, Spiller S, Mei D, et al. Functional characterization of a novel C-terminal *ATP1A2* mutation causing hemiplegic migraine and epilepsy. Cephalalgia 2013;33; 1302–1310.
42. Einholm AP, Nielsen HN, Holm R, Toustrup-Jensen MS, Vilsen B. Importance of a potential protein kinase A phosphorylation site of Na^+,K^+-ATPase and its interaction network for Na^+ binding. J Biol Chem 2016;291:10934–10947.
43. Norimatsu Y, Hasegawa K, Shimizu N, Toyoshima C. Protein-phospholipid interplay revealed with crystals of the calcium pump. Nature 2017;545:193–198.
44. Vedovato N, Gadsby DC. The two C-terminal tyrosines stabilize occluded Na/K pump conformations containing Na or K ions. J Gen Physiol 2010;136:63–82.
45. Artigas P, GadsbyDC. Na^+/K^+-pump ligands modulate gating of palytoxin-induced ion channnels. Proc Natl Acad Sci U S A 2003;100:501–505.
46. Li M, Jazayeri D, Corry B, et al. A functional correlate of severity in alternating hemiplegia of childhood. Neurobiol Dis 2015;77:88–93.
47. Meier S, Tavraz NN, Durr KL, Friedrich T. Hyperpolarization-activated inward leakage currents caused by deletion or mutation of carboxy-terminal tyrosines of the Na^+/K^+-ATPase α subunit. J Gen Physiol 2010;135:115–134.
48. Kaneko M, Desai BS, Cook B. Ionic leakage underlies a gain-of-function effect of dominant disease mutations affecting diverse P-type ATPases. Nat Genet 2014;46: 144–151.
49. Allen NJ, Eroglu C. Cell biology of astrocyte-synapse interactions. Neuron 2017;96: 697–708.
50. Brashear A, Cook JF, Hill DF, et al. Psychiatric disorders in rapid-onset dystonia-parkinsonism. Neurology 2012;79,1168–1173.
51. Cook JF, Hill DF, Snively BM, et al. Cognitive impairment in rapid-onset dystonia-parkinsonism. Mov Disord 2014;29:344–350.
52. Smedemark-Margulies N, Brownstein CA, Vargas S, et al. A novel de novo mutation in *ATP1A3* and childhood-onset schizophrenia. Cold Spring Harb Mol Case Stud 2016; 2:a001008.
53. Wilcox R, Brænne I, Brüggemann N, et al. Genome sequencing identifies a novel mutation in *ATP1A3* in a family with dystonia in females only. J Neurol 2015;262: 187–193.
54. Prange L, Shashi V, Herman K, et al. D-DEMO, a novel and distinct phenotype caused by ATP1A3 mutations. American Academy of Neurology 2017; Meeting abstract, 2017.
55. Einholm AP, Toustrup-Jensen MS, Holm R, Andersen JP, Vilsen B. The rapid-onset dystonia parkinsonism mutation D923N of the Na^+,K^+-ATPase α3 isoform disrupts Na^+ interaction at the third Na^+ site. J Biol Chem 2010;285:26245–26254.
56. Zanotti-Fregonara P, Vidailhet M, Kas A, et al. [123I]-FP-CIT and [99mTc]-HMPAO single photon emission computed tomography in a new sporadic case of rapid-onset dystonia-parkinsonism. J Neurol Sci 2008;273:148–151.
57. Brashear A, Mink JW, Hill DF, et al. ATP1A3 mutations in infants: a new rapid-onset dystonia-Parkinsonism phenotype characterized by motor delay and ataxia. Dev Med Child Neurol 2012;54,1065–1067.
58. Roubergue A, Roze E, Vuillaumier-Barrot S, et al. The multiple faces of the *ATP1A3*-related dystonic movement disorder. Mov Disord 2013;28,1457–1459.
59. Al-Bulushi B, Al-Hashem A, Tabarki B. A wide clinical phenotype spectrum in patients with *ATP1A2* mutations. J Child Neurol 2014;29:265–268.