Amino acid–level signal-to-noise analysis of incidentally identified variants in genes associated with long QT syndrome during pediatric whole exome sequencing reflects background genetic noise

Andrew P. Landstrom, MD, PhD,* Ernesto Fernandez, MD,* Jill A. Rosenfeld, MS,† Yaping Yang, PhD,‡ Andrew L. Dailey-Schwartz, MD, MPH,* Christina Y. Miyake, MD,* Hugh D. Allen, MD,* Daniel J. Penny, MD, PhD,* Jeffrey J. Kim, MD*

From the *Department of Pediatrics, Section of Pediatric Cardiology, Baylor College of Medicine, Houston, Texas, and †Baylor Miraca Genetic Laboratories, Houston, Texas.

BACKGROUND Due to rapid expansion of clinical genetic testing, an increasing number of genetic variants of undetermined significance and unclear diagnostic value are being identified in children. Variants found in genes associated with heritable channelopathies, such as long QT syndrome (LQTS), are particularly difficult to interpret given the risk of sudden cardiac death associated with pathologic mutations.

OBJECTIVE The purpose of this study was to determine whether variants in LQTS-associated genes from whole exome sequencing (WES) represent disease-associated biomarkers or background genetic “noise.”

METHODS WES variants from Baylor Genetics Laboratories were obtained for 17 LQTS-associated genes. Rare variants from healthy controls were obtained from the GnomAD database. LQTS case variants were extracted from the literature. Amino acid–level mapping and signal-to-noise calculations were conducted. Clinical history and diagnostic studies were analyzed for WES subjects evaluated at our institution.

RESULTS Variants in LQTS case-associated genes were present in 38.3% of 7244 WES probands. There was a similar frequency of variants in the WES and healthy cohorts for LQTS1–3 (11.2% and 12.9%, respectively) and LQTS4–17 (27.1% and 38.4%, respectively). WES variants preferentially localized to amino acids altered in control individuals compared to cases. Based on amino acid–level analysis, WES-identified variants are indistinguishable from healthy background variation, whereas LQTS1 and 2 case-identified variants localized to clear pathologic “hotspots.” No individuals who under- went clinical evaluation had clinical suspicion for LQTS.

CONCLUSION The prevalence of incidentally identified LQTS-associated variants is ~38% among WES tests. These variants most likely represent benign healthy background genetic variation rather than disease-associated mutations.

KEYWORDS Genetic testing; Genetics; Long QT syndrome; Mutation; Variant of undetermined significance; Whole exome sequencing

Introduction

Congenital long QT syndrome (LQTS) is characterized by delayed cardiac repolarization in the absence of an underlying syndrome or structural heart disease, which can manifest as a prolonged QT interval on a resting electrocardiogram. This repolarization defect does not typically have a direct pathologic impact on the heart unless the patient is exposed to triggers such as physical exertion, emotion, auditory stimuli, QT-prolonging drugs, or the postpartum period, which electrically destabilize the heart and lead to the potentially lethal dysrhythmia of torsades de pointes. This places the individual at risk for syncope, seizures, and sudden cardiac death. A relatively common arrhythmia syndrome, LQTS affects as many as 1 in 2000 persons. LQTS is a prototypical cardiac channelopathy, and heritable mutations in genes encoding cardiac channelopathies are believed to be its cause. It has been traditionally held that 65–75% of LQTS cases are due to mutations in 1 of 3 major genes: KCNQ1-encoded I\textsubscript{Ks} potassium channel (Kv7.1, LQTS1), KCNH2-encoded I\textsubscript{Kr} potassium channel (Kv11.1, LQTS2), and SCN5A-encoded I\textsubscript{Na} sodium channel (Nav1.5, LQTS3).
These ion channels play key roles in modulating the cardiac action potential, and LQTS-associated genetic defects in these channels delay cellular repolarization. In addition to these 3 major genes, more than a dozen others have been identified as rare causes of LQTS. To date, hundreds of mutations have been identified in 17 LQTS-susceptibility genes; however, genes associated with LQTS4 through LQTS17 are rare among documented LQTS probands. Because of the clear link between heritable mutations in these 17 genes and the pathogenesis of LQTS, a clinical “gene panel” genetic test has been developed. Recent advances in genetic sequencing technology via whole exome sequencing (WES) have made genetic interrogation of a larger number of genes achievable with increasing utilization in the clinical setting among patients with a heterogeneous, or otherwise unclear, phenotypic presentation. Although expansive genetic sequencing tests have provided an increased ability to detect pathologic genetic variation, they have also drastically increased the identification of incidentally identified variants of unclear diagnostic significance (VUS).

To address this issue, we compiled genetic variants found in LQTS-associated genes reported in one of the world’s largest cohorts of clinical WES testing. We then compared this WES cohort to a cohort of genetic variants identified in subjects diagnosed with LQTS and ostensibly healthy individuals. We found remarkable similarity between WES cohort variants and healthy genetic variation and no evidence of LQTS in individuals with a variant who underwent clinical evaluation. We conclude that incidentally identified variants likely represent healthy background variation.

**Methods**

**Genes and gene variants**

Genes and gene variants are detailed in the Supplemental Methods.

**Study cohorts**

This study, which was approved by the institutional review board, included 3 cohorts. It has been previously described and is fully detailed in the Supplemental Methods.

**Whole exome sequencing**

Sequencing was performed as previously described and is detailed in the Supplemental Methods.

**Nomenclature**

Variant nomenclature was done as previously described and is detailed in the Supplemental Methods.

**Primary sequence comparison and statistics**

Primary sequence comparison and statistics were performed as previously described and are detailed in the Supplemental Methods.

**Results**

**Prevalence of incidentally identified LQTS1–17 variants in a large WES cohort**

A large cohort of clinical WES genetic test referrals was interrogated to establish the background variance in LQTS1–17 variants found incidentally in next-generation sequencing referrals. The overall WES cohort included 7938 total individuals with 7244 probands. Among probands, the median age at genetic testing was 6.1 years, with 54.0% males, 45.2% females, and 0.8% fetal samples (Figure 1A and Supplementary Table 1). Individuals undergoing WES testing were referred from various institutions with approximately 90% originating within the United States and approximately 50% from the state of Texas.

LQTS1–17 variants were identified and deemed either “likely pathogenic” or a VUS by WES testing in 2775 individuals (frequency 38.8% [95% CI of 37.3–40.0]) with a total of 3743 unique variants identified. Variants localizing to the major genes commonly associated with LQTS, LQTS1–3, were identified in 813 individuals (11.2% [10.5–12.0] of WES probands). There were 1962 individuals (27.1% [26.1–28.1]) hosting variants localizing to LQTS4–17 genes without an LQTS1–3 variant (Figure 1B). Among variant-positive individuals, the majority hosted a single variant (2012; 72.5% [70.8–74.2] among variant-positive probands), whereas 603 (21.7% [20.2–23.3]) hosted 2 variants, 122 (4.4% [3.6–5.2]) hosted 3, 31 (1.1% [0.8–1.6]) hosted 4, and 7 individuals (0.3% [0.1–0.5]) hosted 5 unique variants. These results are summarized in Figure 1C. Taken together, these results suggest that LQTS-associated variants are incidentally identified in a significant proportion of WES tests and are predominantly VUSs found in LQTS4–17-associated genes.

**Incidentally identified LQTS-associated variants have a similar gene frequency to healthy individuals and not LQTS cases**

To compare the frequency of incidentally identified WES variants to known cases of LQTS, a LQTS case cohort was compiled from a number of international cohort-based studies, including those of Splawski et al (N = 262 individuals), Tester et al (N = 541), Hedley et al (N = 44), Al-Hassan et al (N = 56), Lieve et al (N = 855), and Napolitano et al (N = 430), for a total of 2188 LQTS cases. Among these cases, 992 (45.3% [43.2–47.5]) were positive for a likely disease-associated variant in an LQTS1–3-associated gene. This was approximately 4-fold higher than the WES cohort frequency of 11.2% (P < .0001). Although comprehensive genotyping data for rare causes of LQTS were not available for all studies, LQTS4–17 mutation-positive probands were rare and had a yield of approximately 1.9% [1.3–2.5] among LQTS cases.

Given the marked disparity between the frequency of variants in LQTS1–3 genes among LQTS cases and among incidentally identified individuals, we next calculated the frequencies of rare LQTS-associated gene variants among ostensibly healthy individuals from the GnomAD cohort.
There was a statistically significant, although modest, difference between the prevalence of WES variants in LQTS1–3 genes and the GnomAD control cohort (11.2% [10.5–12.0] vs 12.9% [12.8–13.1], respectively; \( P < .0001 \)). Furthermore, there was a significant difference between the prevalence of incidentally identified variants in LQTS4–17 genes and rare variants in the control cohort. These results are summarized in Figure 2. This was consistent with the findings using the ExAC database (Supplemental Results). Taken together, these results suggest that approximately 13% of ostensibly healthy individuals will have a rare variant found in LQTS1–3 genes, which is modestly different from the rate of incidentally identified variants found in WES testing. In comparison, genes associated with rare causes of LQTS are common, with a rate of approximately 38% in healthy individuals and 27% in WES testing.

**LQTS-associated WES variants localize to genes that are rare causes of LQTS and reflect healthy background prevalence**

To determine the gene-specific prevalence of variants in LQTS-associated genes in next-generation sequencing, we next stratified variant prevalence by genetic loci. Within the WES cohort, we found a small number of individuals with a “likely pathologic” variant, whereas the vast majority hosted VUSs. These VUSs predominantly localized to a small number of LQTS-associated genes, which are traditionally held to be rare causes of LQTS. These include AKAP9 (11.0% [10.3–11.8] of all WES referrals) and ANK2 (9.7% [9.0–10.4]). The 2 most common genetic causes of LQTS, KCNQ1 and KCNH2, had a relatively low frequency of WES-identified VUSs, with a frequency of 2.3% [2.0–2.7] and 2.8% [2.4–3.2], respectively. These results are summarized in Figures 3A and 3B (further detailed in the Supplemental Results). Overall, this subset analysis demonstrates that WES-identified variants tended to be VUSs and localized to genes that are rare causes of LQTS, whereas
genes commonly associated with LQTS pathogenicity had a relatively low rate of incidentally identified WES variants.

We next compared the gene-specific variant prevalence within these cohorts to determine what degree of similarity was reflected at the individual gene level. As with WES variants, control variants were predominantly missense (46.8% [46.6–47.0] variant frequency) and rarely radical (3.2% [3.1–3.2]) in any LQTS-associated gene. These rare variants were found to predominantly localize to LQTS genes with a higher prevalence of WES-identified variants and were rare causes of LQTS. Healthy rare variants were found most commonly in AKAP9 (11.2% [11.1–1.3] variant frequency) and ANK2 (9.9% [9.7–10.0]). These results are summarized in Figure 3C. Among all LQTS genes, there were widely congruent gene frequencies between the WES and control individuals, which reinforces the remarkable similarity between the prevalence of incidentally identified WES variants and rare variants found in healthy individuals.

Given the notable similarity between WES and healthy variants, we evaluated the possibility that the variation in variant frequency was directly related to coding sequence length. AKAP9, the gene locus with the highest frequency of rare variants among the WES (20.1% relative frequency) and GnomAD (19.2%) cohorts demonstrated a nearly identical relative coding nucleotide length of 23.2%. Similarly, AKAP9 composed 23.0% of the relative primary sequence length, 23.1% of the WES variants, and 21.8% of the GnomAD variants. Overall, we found that the relative protein length between LQTS1–17 gene products correlated tightly with the relative frequency of WES and GnomAD rare variants (Supplementary Figure 1).

**Figure 3**  WES and control cohort gene-specific variant prevalence. A: Bar graph of WES variants (blue fill) deemed “likely pathogenic” at the time of genetic testing for each LQTS-associated gene. B: Variants deemed variants of undetermined significance (VUS). C: Rare variants among GnomAD control cohort (white fill). For all panels missense (blue/white), intronic and untranslated regions (UTR; gray), and radical mutations (black) are noted. Error bars denote 95% confidence interval. LQTS = long QT syndrome; WES = whole exome sequencing.

**Gene-specific signal-to-noise ratios for LQTS-associated genes**

Given the similarity of LQTS gene frequency in WES variants and the background genetic variation noted in healthy individuals, we next calculated gene-specific signal-to-noise ratios by normalizing the frequency of variants in LQTS case and WES cohorts, respectively, against the background rate of variants in healthy individuals. In keeping with previous studies, the genes with the highest ratios were KCNQ1 (6.3 [5.9–6.9]) and KCNH2 (7.0 [6.3–7.6]), 2 of the 3 canonical causes of LQTS. SCN5A demonstrated a ratio of 0.76 [0.64–0.92]. Interestingly, KCNJ2 (3.4 [1.5–7.6]) also demonstrated a relatively high signal-to-noise ratio. Among genes only evaluated in cohorts of LQTS-positive, genotype-negative individuals, only CALM1 (32.8 [14.9–71.9]) demonstrated a signal-to-noise ratio >2. These results are summarized in Figure 4A. These results were similar, with an expected higher signal-to-noise ratio when a lower
GnomAD MAF threshold of <0.001 was used (Supplemental Results).

With the high signal-to-noise ratio among truly pathologic variants in canonical LQTS-associated genes, we next examined WES variants. Only CAV3 demonstrated a signal-to-noise ratio >1 (1.8% [1.4–2.2]). This was driven by a high recurrence of the WES variant CAV3-T78M, which made up 50% of all CAV3 variants within the WES cohort. This variant was found in the GnomAD cohort with a variant MAF of 0.0027. These results are summarized in Figure 4B. Overall, these results suggest that incidentally identified WES variants rarely rise above the frequency of rare background variation on an individual gene basis.

**Variant co-localization between cohorts**

To determine whether variants identified among the clinical WES cohort more frequently altered amino acids affected by control or case variants, we mapped the variants along the primary sequence of Kv7.1, Kv11.1, and Nav1.5. There was a relatively low degree of overlap in variant location between WES and LQTS cases, including 35.4% [23.9–48.2] for Kv7.1, 28.9% [19.8–39.4] for Kv11.1, and 11.9% [7.4–17.8] for Nav1.5. In comparison, there was a markedly higher frequency of variant overlap among WES and control residues. Specifically, 75.4% [63.1–85.2] for Kv7.1, 82.2% [72.7–89.5] for Kv11.1, and 90.4% [86.0–94.5] for Nav1.5. Overall, WES cohort variants in Kv7.1 were 3-fold more likely to co-localize to residues found in controls, 4-fold more likely in Kv11.1, and 13-fold more likely in Nav1.5 (P <.0001). Findings are shown in Figure 5.

**Amino acid level signal-to-noise calculations among Kv7.1, Kv11.1, and Nav1.5**

Given the amino acid level similarities between WES and control variants and previous evidence that LQTS-associated mutations localize to pathologic hotspots within the ion channel, we next conducted amino acid level signal-to-noise comparisons. Interestingly, for Kv7.1, signal-to-noise ratios for disease-associated variants peaked in the second transmembrane domain, the pore region, and the KCNE1-binding site. Within the pore region, 3 distinct peaks were seen localizing to the S5 and S6 transmembrane domains of the pore, with a large peak immediately N-terminal to the Kv7.1 ion selectivity filter. There were no such peaks when WES variants were normalized to controls. In Kv11.1, large LQTS disease-associated peaks were noted at the Per-Arnt-Sim domain and the pore domain. Within the pore domain, the region immediately C-terminal to the selectivity filter represented the highest signal within the peptide sequence. Interestingly, we observed a high signal-to-noise peak localizing to Kv11.1 residues 912–930 for both LQTS and WES variants. This region has no known functional domain. Finally, for Nav1.5, there were no significant signal-to-noise peaks for either LQTS or WES variants. These findings are shown in Figure 6, with individual
Overall, these findings suggest that there are functional domains of Kv7.1 and Kv11.1 with a higher likelihood of hosting LQTS-associated pathologic mutations, whereas WES variants do not.

**Indication for WES testing and variant yield among institutional referrals**

To evaluate whether children referred for WES genetic testing demonstrated pre-WES testing suspicion for LQTS or post-test diagnosis of LQTS, we next evaluated the clinical records of WES referrals from our institution (TCH). A total of 223 unique probands were identified who were referred for WES testing from TCH and who hosted either “likely pathologic” or VUSs in 1 of the LQTS1–3 genes. The median age of this cohort was 5.2 [1.8–8.6] years at the time of genetic testing. Reflective of the WES cohort as a whole, most individuals hosted missense VUSs. Furthermore, the indication for WES testing referral was rarely only cardiac in nature (0.9% [0.11–3.20]) (Figure 7, Supplemental Results, and Supplementary Table 2).

**Clinical evaluation of institutional referrals with variants in LQTS genes**

To determine whether any individual with incidentally identified variants in LQTS-associated genes had clinical suspicion for LQTS, clinical records from institutional referrals were reviewed. The vast majority of children demonstrated no personal or family history of arrhythmic disease, syncope, seizures, or sudden death. Given the link between multiple LQTS-associated variants and severe LQTS, individuals with 4 or more variants within the WES cohort were separately evaluated. Two of these individuals carried borderline prolonged QTc intervals of 467 and 461 ms, respectively; both were noted to have significant congenital anomalies and had congenitally malformed hearts. No individuals were found to have QTc >480 ms, evidence of ventricular arrhythmias, suspected channelopathy, or LQTS (Figure 7 and Supplemental Results). Taken together, these results suggest that incidentally identified LQTS1–3-associated variants are not associated with QT prolongation or LQTS diagnoses. In the absence of a positive family history of LQTS or in the setting of refractory epilepsy, these variants are unlikely to represent disease susceptibility.

**Discussion**

Although early studies of LQTS gene variation in ostensibly healthy individuals have estimated that ~5% of ostensibly healthy individuals will host a rare variant in a LQTS-associated gene, the American College of Medical Genetics and Genomics (ACMG) recommends the reporting of any “known pathologic” variant associated with LQTS. Recently, a curated list of disease-associated genes associated with diseases that present during childhood was
proposed by the BabySeq project. A reflection of rapidly increasing fetal and neonatal WES testing, the authors put forward a list of gene loci associated with clinically actionable, and thus reportable, variants. All major LQTS-associated genes are designated as clinically actionable and, thus, reportable. Although it is clear that broad genetic sequencing has diagnostic value, these studies do not address the large number of incidentally identified rare variants with unclear diagnostic value.

Not unique to LQTS, several studies have demonstrated a significant rate of rare background variation in other channelopathic and cardiomyopathic diseases among healthy individuals. We have found that 6% of ostensibly healthy individuals will host rare variants in either of the 2 genes associated with the channelopathy catecholaminergic polymorphic ventricular tachycardia, and ~9% of individuals undergoing clinical WES testing will host a “likely pathogenic” variant or VUS. Here, we identified a background rate of rare variants in LQTS1–3 genes to be ~13% among individuals composing the GnomAD database. A significantly higher ~38% of individuals hosted a variant in 1 of the LQTS4–17 genes, predominantly localizing to AKAP9 and ANK2. Similarly, incidentally identified variants from the WES cohort were noted to compose a similar prevalence, with the prevalence of LQTS1–3 in pathologic cases of ~45%. This, combined with a high signal-to-noise ratio in the common genes associated with LQTS (KCNQ1 and KCNH2), suggests that incidentally identified variation is likely background genetic noise.

Our data suggest that a pathologic variant in KCNQ1 or KCNH2 connotes a higher chance for a post-test probability of LQTS in the context of pretest clinical suspicion. This reflects the probabilistic approach to diagnosis established by the Bayes theorem, which determines the probability of a future event is related to prior knowledge of the conditions associated with it. As such, identification of an incidentally identified variant in a gene with a low signal-to-noise ratio would hold less diagnostic value. Furthermore, an incidentally identified variant in the absence of a concerning family history or clinical suspicion for LQTS is unlikely to represent monogenic markers of disease even if localizing to an LQTS-associated gene. We acknowledge the significant challenge inherent in variant pathogenicity predication and believe that further study using tools such as machine learning and artificial intelligence have a role in continuing to refine predictive modeling. Because of these factors, care must be given in the interpretation of incidentally identified VUSs.

In addition to the gene locus itself, the diagnostic interpretation of a VUS should involve variant location within the gene sequence, particularly with variants in KCNQ1-encoded Kv7.1 and KCNH2-encoded Kv11.1. Our finding of a variable signal-to-noise ratio at the amino acid level supports a number of previous studies suggesting that the biophysical impact of potassium channel mutations may be largely

---

**Figure 6** Amino acid–level signal-to-noise variant frequency for the 3 major gene products associated with LQTS. Frequency of LQTS (green) and WES (blue) cohort variant frequency vs amino acid position for Kv7.1 (A), Kv11.1 (B), and Nav1.5 (C). AKAP9 = AKAP9-protein binding; CD = cytoplasmic; cNBD = C-terminal nucleotide binding; FGF = FGF13 binding; G = inactivation gate; KCNE1 = KCNE1-protein binding; LC = L-type Ca2+ channel binding; LQTS = long QT syndrome; PAS = Per-Arnt-Sim; SF = selectivity filter; S1–S6 = transmembrane; SSTK = SSTK-interacting protein; TM = transmembrane; TSSK6-activating co-chaperone protein binding domains; WES = whole exome sequencing.
dependent on location. This so-called “intragenic risk” is well established in Kv7.1, whereby LQTS-associated, loss-of-function mutations in the pore domain selectivity filter (S1–S4) are linked to a high risk of torsades de pointes and sudden cardiac death,22,23 and mutations with the PAS domain of Kv11.1 are associated with channel trafficking defects and severe LQTS.24,25 In addition to these known functional domains, we report a significant signal-to-noise peak in Kv11.1 residues 912–930. To our knowledge, this region of Kv11.1 has no identified functional domain or regulatory role. Given the critical functionality of the other domains identified by our analysis, we hypothesize that localization of disease-associated mutations normalized to rare background variation may be a tool for identifying biophysically and physiologically relevant protein domains.

The clinical diagnosis of LQTS can be a challenge, even in the setting of high clinical suspicion for the disease. Although many individuals are diagnosed in childhood, the average diagnosis is often made in the second or third decade of life.7 Because of the referral bias of our institution, the age of genetic testing in the TCH cohort was ~5–6 years. It is notable that ~40% of LQTS probands had a personal history of syncope or a family history of LQTS, whereas this was largely absent in the TCH cohort.7 Although our evaluation for clinical signs, symptoms, or history of LQTS was retrospective in nature, we cannot exclude the possibility of a clinically occult LQTS diagnosis or a diagnosis later in life.

**Conclusion**

Incidentally identified LQTS-associated gene variants are common in children undergoing WES, associated with more one-third of all tests. Based on amino acid–level signal-to-noise analysis, WES-identified variants are indistinguishable from healthy background variation, whereas LQTS case-identified variants localized to clear pathologic “hotspots” within KCNQ1-encoded Kv7.1 and KCNH2-encoded Kv11.1.

**Appendix**

**Supplementary data**

Supplementary data associated with this article can be found in the online version at https://doi.org/10.1016/j.hrthm.2018.02.031.

**References**


