NCJRS
NATIONAL CRIMINAL JUSTICE REFERENCE SERVICE

OFFICE OF JUSTICE PROGRAMS
BJA BJS NIJ OJJDP OVC SMART

The author(s) shown below used Federal funding provided by the U.S. Department of Justice to prepare the following resource:

Document Title:	**Advancing Molecular Diagnostics in Sudden Unexplained Deaths**
Author(s):	**Yingying Tang, M.D., Ph.D., DABMGG**
Document Number:	**300769**
Date Received:	**April 2021**
Award Number:	**2014-DN-BX-K001**

This resource has not been published by the U.S. Department of Justice. This resource is being made publically available through the Office of Justice Programs' National Criminal Justice Reference Service.

Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

Final Summary Overview

NIJ FY 14 Basic Science Research to Support Forensic Science

2014-DN-BX-K001

Advancing Molecular Diagnostics in Sudden Unexplained Deaths

March 28, 2021

Yingying Tang, MD, PhD, DABMGG

Director, Molecular Genetics Laboratory

Office of Chief Medical Examiner

421 East 26th Street

New York, NY, 10016

Tel: 212-323-1340

Fax: 212-323-1540

Email: ytang@ocme.nyc.gov

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

Contents

1. PURPOSE OF THE PROJECT 2
2. PROJECT DESIGN AND METHODS 3
3. PROJECT SUBJECTS 8
4. DATA ANALYSIS FOR VARIANT INTERPRETATION 8
5. PROJECT FINDINGS 9
   A. Evaluation summary of the technical specifications for the panel of 95 cardiac genes 9
   B. Scholarly publications 10
6. IMPLICATIONS FOR CRIMINAL JUSTICE POLICY AND PRACTICE IN THE UNITED STATES 15
7. REFERENCES 16

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

This is the Final Summary Overview on the grant-**awarded project, entitled** "Advancing Molecular Diagnostics in Sudden Unexplained Deaths" (award# 2014-DN-BX-K001). The document addresses the purpose of the project, project design and methods, project subjects, data analysis, project findings, and implications for criminal justice policy and practice in the United States.

1. PURPOSE OF THE PROJECT

The concept of the molecular autopsy has been proposed for more than a decade. However, a systematic molecular evaluation of Sudden Unexplained Death (SUD) cases, for which the causes of death remain undetermined after comprehensive forensic investigation, complete autopsy, and various ancillary laboratory tests (e.g. toxicology and microbiology tests), was not feasible until the advent of next-generation sequencing (NGS) technology. This powerful molecular diagnostic tool is being used in our NIJ FY2011 Basic Research Grant project (2011-DN-BX-K535), which aimed to evaluate cardiac genes in a large SUD cohort. Our published study[1] showed that 40% of SUD cases had identifiable pathogenic/likely pathogenic variants or variants of uncertain significance that warrant further family and functional studies.

In the grant, we proposed two specific aims: 1) validate a molecular testing panel comprising of 95 cardiac arrhythmogenic genes indicated for SUD, using samples routinely collected at New York City Office of Chief Medical Examiner (NYC OCME), such as RNALater® fixed tissue and bloodstain card samples; and 2) evaluate the efficacy of utilizing the cardiac panel for commonly collected, yet technically challenging, postmortem samples, such as formalin-fixed paraffin-embedded (FFPE).

The focus of this grant project is to measure the analytical performance of the testing method (e.g. sensitivity, specificity, reproducibility, and sample input range) and to evaluate a feasible bioinformatic pipeline for high throughput NGS data analysis and genetic variant reporting. This project was carried out within the Molecular Genetics Laboratory, a College of American Pathologists (CAP)-accredited laboratory, in the New York City Office of Chief Medical Examiner.

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

It is noteworthy that this project is the second of four grants which NYC OCME has received from NIJ on the topic of molecular autopsy. Most importantly, each NIJ-awarded grant not only produced scholarly publications based on the funded studies but also served as validation of the positive value of our work and affirmation of the influencer role NYC OCME plays in the field of forensic sciences as it established a model in the United States for the incorporation of state-of-art and high-throughput molecular autopsy in routine death investigations. Specifically, the first grant (2011-DN-BX-K535) allowed NYC OCME to acquire the next-generation sequencing technology, bioinformatics pipeline, and variant interpretation expertise; the second grant (this grant) allowed the validation of cardiac-focused genes testing panel to be implemented routinely in sudden death investigation in NYC OCME; the third grant (2015-DN-BX-K017) enabled the first-ever genetic counseling program to be established in NYC OCME as well as collaborative functional studies of genetic variants to be conducted; the fourth grant (2018-DU-BX-0204) enabled the establishment of comprehensive molecular testing panels for disease-causing genes associated with various underlying genetic conditions associated with sudden death, such as epilepsy, aortopathy, thrombophilia, and non-cardiac channelopathy, adding to the cardiac-focused testing panel implemented through previous grants. **Through NIJ's continuous grant funding support, NYC OCME was able to be an example of high standards to the** nation with high-quality molecular autopsy and nonwavering commitment to forensic science, criminal justice and the surviving families which we serve.

2. PROJECT DESIGN AND METHODS

Genes Tested: The selection of the panel of 95 cardiac genes (table below) was based on literature review. The panel of molecular analysis of 95 cardiac arrhythmogenic genes includes sequence analysis of 95 genes associated with various forms of cardiac channelopathy (e.g. long QT syndrome, Brugada syndrome, and catecholaminergic polymorphic ventricular tachycardia) and cardiomyopathy (e.g. hypertrophic cardiomyopathy, arrhythmogenic right ventricular cardiomyopathy, idiopathic dilated cardiomyopathy, and left ventricular noncompaction cardiomyopathy). Cardiac channelopathy and cardiomyopathy are typically inherited in an autosomal dominant pattern, though a few genes are inherited in an autosomal recessive pattern and some genes are x-linked.

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

Gene	Protein	Ref SEQ (HGMD)	Associated Diseases
ABCC9	a member of the superfamily of ATP-binding cassette (ABC) transporters.	NM_005691.2	AF, Brs, AF, DCM, Cantú syndrome
ACTC1	Actin, alpha, cardiac muscle	NM_005159.4	HCM,RCM, LVNC, DCM
ACTN2	Actinin, alpha 2	NM_001103.2	HCM,DCM, LVNC
AKAP10	a member of the A-kinase anchor protein family.	NM_007202.3	conduction
AKAP9	a member of the AKAP family.	NM_005751.4	LQT11
ANK2	a member of the ankyrin family of proteins	NM_001148.4	LQT4
ANKRD1	Ankyrin repeat domain 1 (cardiac muscle)	NM_014391.2	HCM,DCM
ARHGAP24	Rho GTPase activating protein 24	NM_001025616.2	conduction
BAG3	BCL2-associated athanogene 3	NM_004281.3	RCM, HCM, DCM
CACNA1C	alpha-1 subunit of a voltage-dependent calcium channel.	NM_000719.6	LQT, Brs, Timothy
CACNA2D1	Calcium channel, voltage-dependent, alpha 2/delta subunit 1	NM_000722.2	heart block, Brs, SQT
CACNB2	Calcium channel, voltage-dependent, beta 2 subunit	NM_201590.2	Brs4
CALM1	Calmodulin 1 (phosphorylase kinase, delta)	NM_006888.4	LQT14, CPVT
CALM2	Calmodulin 2 (phosphorylase kinase, delta)	NM_001743.4	LQT15
CALR3	Calreticulin 3	NM_145046.4	HCM
CASQ2	Calsequestrin 2 (cardiac muscle)	NM_001232.3	CPVT, LVNC (AR)
CAV1	Caveolin 1, caveolae protein, 22kDa	NM_001753.4	SIDS, (via connexin 43)
CAV3	Caveolin 3	NM_033337.2	LQT9, HCM, DCM
CRYAB	Crystallin, alpha B	NM_001885.1	DCM
CSRP3	Cysteine and glycine-rich protein 3 (cardiac LIM protein)	NM_003476.4	HCM,DCM
CTF1	Cardiotrophin 1	NM_001330.3	HCM,DCM
DES	Desmin	NM_001927.3	ARVD, RCM, DCM
DPP6	Dipeptidyl-peptidase 6	NM_001936.3	VF
DSC2	Desmocollin 2	NM_024422.3	ARVD, DCM
DSG2	Desmoglein 2	NM_001943.3	ARVD, DCM
DSP	Desmoplakin (DPI,DPII)	NM_004415.2	ARVD, DCM

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

DTNA	Dystrobrevin, alpha	NM_001390.4	LVNC
EMD	Emery-Dreyfuss muscular dystrophy (Emerin)	NM_000117.2	DCM, Emery-Dreifuss muscular dystrophy
FHL2	Four and a half LIM domains 2	NM_201555.1	DCM, (binds to HERG)
GATAD1	GATA zinc finger domain containing 1	NM_021167.4	DCM
GJA1	Gap junction protein alpha 1 (connexin 43)	NM_000165.3	SIDS, (connexin 43)
GJA5	Gap junction protein, alpha 5, 40kDa (connexin 40)	NM_005266.5	AF, Atrial standstill, digenic (GJA5/SCN5A)
GLA	Galactosidase alpha	NM_000169.2	Fabry
GPD1L	Glycerol-3-phosphate dehydrogenase 1-like	NM_015141.3	Brs, SIDS
HCN4	Hyperpolarization activated cyclic nucleotide-gated potassium channel 4	NM_005477.2	Brs, SSS
JPH2	Junctophilin 2	NM_020433.4	ARVD, HCM, DCm
JUP	Junction plakoglobin	NM_002230.2	ARVD, Naxos disease
KCNA5	Potassium voltage-gated channel, shaker-related subfamily, member 5	NM_002234.3	AF
KCND2	Potassium voltage-gated channel, Shal-related subfamily, member 2	NM_012281.2	J-wave syndrome, Sudden death
KCND3	Potassium voltage-gated channel, Shal-related subfamily, member 3	NM_004980.4	Brs, SIDS
KCNE1	Potassium voltage gated channel, Isk related family, member 1	NM_000219.4	LQT5
KCNE1L	KCNE1-like	NM_012282.2	VF, AF
KCNE2	Potassium voltage-gated channel, Isk-related family, member 2	NM_172201.1	LQT6, AF
KCNE3	Potassium voltage-gated channel, Isk-related family, member 3	NM_005472.4	Brs6
KCNE4	Potassium voltage-gated channel, isk-related family, member 4	NM_080671.2	AF
KCNH2	Potassium voltage-gated channel, subfamily H, member 2 (HERG)	NM_000238.3	LQT2, SQT1
KCNJ2	Potassium inwardly-rectifying channel, subfamily J, member 2	NM_000891.2	LQT,SQT, CPVT, AF, Andersen syndrome
KCNJ5	Potassium inwardly-rectifying channel, subfamily J, member 5	NM_000890.3	LQT13
KCNJ8	Potassium inwardly-rectifying channel, subfamily J, member 8	NM_004982.3	Brs, Cantú syndrome, SIDS, VF
KCNQ1	Potassium voltage gated channel, KQT-like subfamily, member 1 (KVLQT1)	NM_000218.2	LQT, SQT, AF
LAMA4	Laminin, alpha 4	NM_002290.4	DCM

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

LAMP2	Lysosomal-associated membrane protein 2	NM_002294.2	HCM, DCM, Danon disease
LDB3	LIM domain binding 3 (Z1 isoform)	NM_001080116.1	ARVD, LVNC, HCM, DCM
LMNA	lamin A/C	NM_170707.3	LVNC, DCM
MYBPC3	Myosin binding protein C, cardiac	NM_000256.3	LVNC, HCM, DCM
MYH6	Myosin, heavy polypeptide 6, cardiac muscle, alpha (cardiomyopathy, hypertrophic 1)	NM_002471.3	SSS, HCM, DCM
MYH7	Myosin, heavy polypeptide 7, cardiac muscle, beta	NM_000257.2	LVNC, RCM, HCM, DCM
MYL2	Myosin, light chain 2, regulatory ventricular	NM_000432.3	HCM10
MYL3	Myosin, light chain, alkali; ventricular and skeletal slow	NM_000258.2	HCM8
MYLK2	Myosin light chain kinase 2, skeletal muscle	NM_033118.3	HCM1
MYOZ2	Myozenin 2	NM_016599.4	HCM16
MYPN	myopalladin	NM_032578.3	HCM,DCM, RCM
NEBL	Nebulette	NM_006393.2	DCM
NEXN	Nexilin (F actin binding protein)	NM_144573.3	HCM20,DCM
NPPA	Natriuretic peptide precursor A	NM_006172.3	AF, Atrial standstill
PKP2	Plakophilin 2	NM_004572.3	ARVD9, Brs, DCM
PLN	Phospholamban	NM_002667.3	HCM18, DCM
PRDM16	PR domain containing 16	NM_022114.3	DCM, LVNC
PRKAG2	Protein kinase, AMP-activated, gamma 2 non-catalytic subunit	NM_016203.3	WPW, HCM6
PTPN11	Protein tyrosine phosphatase, non-receptor type 11	NM_002834.3	noonan, HCM
RANGRF	RAN guanine nucleotide release factor	NM_016492.4	Brs
RBM20	RNA binding motif protein 20	NM_001134363.1	DCM
RYR2	Ryanodine receptor 2 (cardiac)	NM_001035.2	CPVT
SCN10A	Sodium channel, voltage-gated, type X, alpha subunit	NM_006514.2	Brs, AF
SCN1B	Sodium channel, voltage-gated, type 1, beta polypeptide	NM_001037.4	Brs, AF
SCN2B	Sodium channel, voltage-gated, type II, beta	NM_004588.4	LQT, AF, Brs
SCN3B	Sodium channel, voltage-gated, type III, beta	NM_018400.3	Brs, AF
SCN4B	Sodium channel, voltage-gated, type iv, beta	NM_174934.3	LQT10, AF

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

SCN5A	Sodium channel, voltage gated, type V, alpha polypeptide	NM_198056.2	LQT, Brs, AF, DCM, Heart block, SSS,VF
SGCD	Sarcoglycan, delta	NM_000337.5	DCM
SLMAP	Sarcolemma associated protein	NM_007159.2	Brs
SNTA1	Syntrophin, alpha 1 (dystrophin-associated protein A1,	NM_003098.2	LQT12
TAZ	tafazzin	NM_000116.3	LVNC, DCM, Barth syndrome
TCAP	Titin-cap (telethonin)	NM_003673.3	HCM, DCM
TGFB3	Transforming growth factor, beta 3	NM_003239.2	ARVD, Rienhoff syndrome
TMEM43	Transmembrane protein 43	NM_024334.2	ARVD, Emery-Dreifuss muscular dystrophy
TMPO	Thymopoietin	NM_003276.2	DCM
TNNC1	Troponin C, slow	NM_003280.2	HCM, DCM
TNNI3	Troponin I, cardiac muscle isoform	NM_000363.4	HCM, RCM, DCM
TNNT2	Troponin T2, cardiac	NM_001001430.1	LVNC, RCM, HCM, DCM
TPM1	Tropomyosin 1 alpha	NM_001018005.1	LVNC, HCM, DCM
TRDN	Triadin	NM_006073.3	CPVT
TRPM4	transient receptor potential cation channel, subfamily M, member 4	NM_017636.3	heart block, Brs
TTN	Titin	NM_133378.4	ARVD, HCM, DCM
VCL	Vinculin	NM_014000.2	LVNC, HCM, DCM

Testing Methodology: Genomic DNA is isolated from dry bloodstain cards or postmortem tissue samples preserved in RNAlater®. The coding regions and splice sites of the 95 genes are completely sequenced in this test. The test is performed by oligonucleotide-based in-solution target capture (Haloplex Target Enrichment System, Agilent Technologies). Each sample was also indexed by bar-codes during the enrichment process. Following quality and quantity evaluation of the target region enrichment (library validation), bar-code indexed DNA samples were pooled for pair-end sequencing using Illumina Miseq.

Sequencing Data Analysis: Primary sequencing data analysis (generating quality sequencing read and de-multiplexing) is performed on Illumina Miseq. Secondary sequencing analysis (alignments, variant calling, and sequencing quality filtering) is performed using NextGENe (SOFTGENETICS®). Tertiary sequencing analysis

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

(variant annotation and classification) is performed in Geneticist Assistant (SOFTGENETICS®). The genomic reference coordinate is GRCh37/hg19. Sanger sequencing is used to provide sequencing data for regions with no coverage as well as to confirm reportable variants which have less than 100x coverage, less than 0.25 variant frequency and less than 0.25 read balance. This test was developed and its performance characteristics were supported by this grant in the Molecular Genetics Laboratory in the City of New York Office of Chief Medical Examiner.

3. PROJECT SUBJECTS

Control DNAs: genotype known Coriell Control DNA: NA12878, NA12891, and NA12892.

Postmortem Samples: RNAlater® tissue samples (heart, liver, and spleen) and bloodstain card of three postmortem cases; FFPE and RNAlater® tissue samples of 12 postmortem cases. Multiple test runs were performed to determine analytic characteristics.

4. DATA ANALYSIS FOR VARIANT INTERPRETATION

The clinical relevance of sequence variants is evaluated in accordance with the standards and guidelines for the interpretation of sequencing variants established by American College of Medical Genetics and Genomics, and the Association for Molecular Pathology. Searches in disease databases (e.g. Human Gene Mutation Database (HGMD), ClinVar, and appropriate locus-specific databases) and the literature in PubMed are used to determine if a sequence variant has been previously reported with regards to clinical relevance, and if there is evidence from family studies, cohort studies, or functional characterization to support the deleterious or benign nature of a variant. If a variant is previously unreported, the evaluation consists of: 1) examining the inheritance pattern of the gene, variant type (loss of function, missense, or in-frame insertion or deletion) and its location in the functional domain of the protein or the splice site of the gene; 2) estimating the minor allele frequency reported in the large population databases (e.g. The Genome Aggregation Database (gnomAD) with 125,748 exome sequences and 15,708 whole-genome sequences from unrelated individuals sequenced as part of various disease-specific and population genetic studies; Exome Aggregation Consortium (ExAC) with 60,706 unrelated individuals sequenced as part of various disease-specific and population genetic studies, 1000 genome, and NHLBI Exome Sequencing Project (ESP); 3)

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

performing the in silico function prediction reported from computation tools (e.g. Polyphen 2, Provean, and MutationTaster) to a protein; and 4) reviewing relevant autopsy findings, past medical history or family history. After available information of a variant is considered, a variant is classified into one of the five categories: pathogenic, likely pathogenic, variant of uncertain significance (VUS), likely benign, and benign.

5. PROJECT FINDINGS

The project findings include two categories: A) evaluation summary of the technical specifications for the panel of 95 cardiac genes; B) scholarly publications resulted from this grant funded research;

A. Evaluation summary of the technical specifications for the panel of 95 cardiac genes

1) Testing Characteristics: testing characteristics include sensitivity, specificity, accuracy, and reproducibility. Based on multiple test runs data, we have determined that the test sensitivity is 98% to 99% (95% CI), the test specificity is 95% to 98% (95% CI), and average accuracy of the test is 95.12%. The test reproducibility is evaluated by coefficient of variation (CV) which is the ratio of the standard deviation to the mean. The CV for test sensitivity, specificity and accuracy is equal or less than 5%. Since most variants are substitutions the testing characteristics evaluated here are not applicable for other types of variants such as large-scaled insertions or deletions. The difference on low coverage regions is within 15%. Miseq performed well and consistently passed the quality indicators specified by the manufacture. Detailed technical validation data have been provided in the previous progress reports and will not be repeated here due to space limit.

2) Sample Types Evaluation: Heart, liver, spleen, and bloodstained card are most common postmortem sample types received by Molecular Genetics Laboratory, so they were chosen to verify the testing conditions. Among these sample types, gDNA from heart consistently give the best coverage and the coverage is comparable to those from control DNA. There is no difference in variant identification between the four sample types.

3) Bioinformatics Pipeline Evaluation: we chose to evaluate the commercial bioinformatic pipelines from SOFTGENETICS®, as they are commonly used in clinical diagnostic laboratories. Firstly, we focused the evaluation of the NextGENe (SOFTGENETICS®) program, which is chosen to perform the secondary sequencing data analysis, i.e. alignments, variant calling and sequencing quality filtering. Results showed that NextGENe identified all variants

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

known in the control samples with the sequencing coverage above the alignment threshold. This demonstrates that NextGENe could confidently make accurate base calling if the sequencing coverage is good. There was one zygosity discrepancy for two variants in NA12891 and NA12892, respectively. Both variants had low coverage which was below the alignment threshold. If there are variants for which the coverage is below the alignment threshold, they are retested using Sanger sequencing; therefore, those variants won't be missed during data analysis. NextGENe also reports the variants that are not listed in the reference. Those variants which had uneven coverage from forward and reverse direction are confirmed by Sanger sequencing. Secondly, we evaluated the Geneticist Assistant (GA), a tertiary analysis program for variant annotation and classification. The program takes the input VCF files, runs against the databases (e.g. ExAC and gnomAD, 1000 genome, ESP, ClinVar, and dbNSFP) within the GA, and generates a list of annotated variants pending for classification. To evaluate the GA, an input VCF file is compared to its output VCF file, for all three control samples, to determine if any variant in the input file missed in its output file. Results showed that all variants in the input VCF files are included in the output VCF files for all three samples. Variant interpretations are performed by laboratory director(s); therefore, the GA is not evaluated for its accuracy in variant interpretation.

4) Test Limitations: This test may neither detect the variants in non-coding regions that could affect gene expression, nor identify with high accuracy non-substitution type of changes, such as large-scale ins/del variants, or copy number variants encompassing all or a large portion of the gene. Identification of variants may be affected by the sequencing reads in highly homologous regions or repeat rich regions. A low possibility of false negative rate (<2%) is expected due to technical limitation of the Haloplex Target Enrichment System in generating strand biases. A negative result does not exclude the possibility of a deleterious sequence variant in one of the 95 genes that is outside of the regions tested. In addition, a negative test result does not exclude the possibility of a genetic alteration in another gene which was not tested in this panel.

B. Scholarly publications

The following peer-reviewed articles have been published as a result of this grant-funded research:

1) Test Evaluation Using FFPE samples

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

We published the study of evaluating test results using FFPE samples (PMID: 32155531), entitled "Using postmortem formalin fixed paraffin-embedded tissues for molecular testing of sudden cardiac death: A cautionary tale of utility and limitations"[2]. For archived cases of previously young healthy individuals where cause of sudden death remains undetermined, formalin fixed paraffin-embedded tissues (FFPE) samples are often the only biological resource available for molecular testing. We aim to ascertain the validity of postmortem molecular analysis of 95 cardiac genes using the FFPE samples routinely processed in the offices of medical examiners - typical fixation time in formalin ranges from days to months. The study was conducted in the College of American Pathologists accredited Molecular Genetics Laboratory within the City of New York Office of Chief Medical Examiner. Twelve cases, with FFPE samples and corresponding non-formalin fixed samples (RNAlater-preserved tissues or bloodstain card), were chosen for testing results comparison. The methods of extracting DNA from FFPE samples using Covaris, Qiagen, and Promega products showed comparable results. The quality of the extracted DNA, the target-enriched DNA libraries of 95 cardiac genes using HaloPlex Target Enrichment system by Agilent Technologies, and sequencing results using Illumina Miseq instrument were evaluated. Compared to the sequencing results of the nonfixed samples, the FFPE samples were categorized into three groups: 1) Group 1 samples fixed in formalin 2-6 days, had greater than 55 % sequencing regions ⩾30x and 94%-100% variant concordance. 2) Group 2 samples fixed in formalin for 8 days, showed intra-sample sequencing variations: the surface tissues showed 25%-27% extra variants (false positive) and 8.1%-9.7% missing variants (false negative), whereas the repeated core tissues showed reduced extra variants to 1.6 % and the false negative error was unchanged. 3) Group 3 samples fixed in formalin 29-136 days, had 2-55 % sequencing regions ⩾30x, up to 52.2 % missed variants and up to 6.3 % extra variants. All reportable variants (pathogenic, likely pathogenic or variant of uncertain significance) identified in the nonfixed samples were also identified in FFPE, albeit three variants had low confidence variant calling. In summary, our study showed that postmortem molecular diagnostic testing using FFPE samples routinely processed by the medical examiners should be cautioned, as they are replete with false positive and negative results, particularly

This resource was prepared by the author(s) using Federal funds provided by the U.S.
Department of Justice. Opinions or points of view expressed are those of the author(s) and do not
necessarily reflect the official position or policies of the U.S. Department of Justice.

when sample fixation time is longer than 8 days. Saving non-formalin fixed samples for high fidelity molecular analysis is strongly encouraged.

2) Case Studies and Cohort Report:

a. In one case study (PMID: 30582040, partially supported by this grant), entitled "Molecular autopsy: using the discovery of a novel de novo pathogenic variant in the *KCNH2* gene to inform healthcare of surviving family"[3], forensic investigation and molecular autopsy were performed on an 18-year-old female who died suddenly and unexpectedly. Co-segregation family study of the first-degree relatives and functional characterization of the variant were conducted. We identified a novel nonsense variant, NP_000229.1:p.Gln1068Ter, in the long QT syndrome type II gene, *KCNH2*, in the decedent. This finding correlated with her ante-mortem electrocardiogram findings. Patch clamp functional studies using transfected COS-7 cells show that hERG-ΔQ1068 has a mixed phenotype, with both gain- (negative voltage shift of steady-state activation curve, the positive shift of the steady-state inactivation curve, and accelerated activation) and loss-of function (reduced current density, reduced surface expression and accelerated deactivation) hallmarks. Loss of cumulative activation during rapid pacing demonstrates that the loss-of-function phenotype predominates. The wild-type channel did not rescue the hERG-ΔQ1068 defects, demonstrating haploinsufficiency of the heterozygous state. Targeted variant testing in the family showed that the variant in *KCNH2* arose de novo, which eliminated the need for exhaustive genome testing and annual cardiac follow-up for the parents and four siblings. Molecular testing enables accurate determination of natural causes of death and precision care of the surviving family members in a time and cost-saving manner. We advocate for molecular autopsy being included under the healthcare coverage in US.

b. In one case series study (PMID: 30282064, partially supported by this grant), entitled "Phenotypic variations in carriers of predicted protein-truncating genetic variants in *MYBPC3*: an autopsy-based case series"[4], we characterize predicted protein-truncating variants (PTVs) in *MYBPC3*, the gene most commonly associated with hypertrophic cardiomyopathy (HCM), found in a series of autopsied HCM cases after sudden unexpected cardiac death. All cases underwent death scene investigation, gross and microscopic autopsies, toxicological testing, a

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

review of medical records, and a molecular analysis of 95 cardiac genes. We found four pathogenic PTVs in *MYBPC3* among male decedents. All variants were previously submitted to ClinVar without phenotype details. Two PTVs were located in the cardiac-specific myosin S2-binding (M) motif at the N-terminus of the *MYBPC3*-encoded cMyBP-C protein, and two PTVs were in the non-cardiac-specific C-terminus of the protein. The carriers of two cardiac-specific M-motif PTVs died at 38-years-old; their heart weight (HW, g) and body mass index (BMI, kg/m2) ratio were 34.90 (890/25.5) and 23.56 (980/41.6), respectively. In contrast, the carriers of two non-cardiac-specific C-terminal PTVs died at age 57 and 67 years, respectively; their HW and BMI ratio were 14.71 (450/30.6) and 13.98 (600/42.9), respectively. A detailed three-generation family study was conducted in one case. This study showed age-at-death variations among *MYBPC3* PTVs carriers in adult males.

c. In a cohort study (PMID: 29247119, partially supported by this grant), entitled "Applying High-Resolution Variant Classification to Cardiac Arrhythmogenic Gene Testing in a Demographically Diverse Cohort of Sudden Unexplained Deaths"[1], we reported an evaluation of cardiac channelopathy and cardiomyopathy genes in a large, demographically diverse sudden unexplained death cohort that underwent thorough investigation in the United States' largest medical examiner's office. The cohort has 296 decedents: 147 Blacks, 64 Hispanics, 49 Whites, 22 Asians, and 14 mixed ethnicities; 142 infants (1 to 11 months), 39 children (1 to 17 years), 74 young adults (18 to 34 years), and 41 adults (35 to 55 years). Eighty-nine cardiac disease genes were evaluated. Using a high-resolution variant interpretation workflow, we classified 17 variants as pathogenic or likely pathogenic (2 of which were incidental findings and excluded in testing yield analysis), 46 novel variants of uncertain significance, and 130 variants of uncertain significance. Nine pathogenic or likely pathogenic variants in ClinVar were reclassified to likely benign and excluded in testing yield analysis. The yields of positive cases by ethnicity and age were 21.4% in mixed ethnicities, 10.2% Whites, 4.5% Asians, 3.1% Hispanics, and 2% Blacks; 7.7% children, 7.3% in adults, 5.4% young adults, and 2.8% infants. The percentages of uncertain cases with variants of uncertain significance by ethnicity were 45.5% in Asians, 45.3% Hispanics, 44.20% Blacks, 36.7% Whites, and 14.3% in mixed ethnicities. In conclusion, High-resolution variant interpretation provides diagnostic accuracy and healthcare efficiency. Under-represented populations warrant greater inclusion in future studies.

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

d. A cohort study ((PMID: 31436011, partially supported by this grant), was entitled "Lessons learned from testing cardiac channelopathy and cardiomyopathy genes in individuals who died suddenly: A two - year prospective study in a large medical examiner's office with an in - house molecular genetics laboratory and genetic counseling services"[5]. This is a comprehensive review and analysis of 254 cases tested consecutively in the in - house College of American Pathologist - accredited molecular genetics laboratory within the New York City Office of Chief Medical Examiner between October 2015 and February 2018, using a multigene cardiac panel composed of 95 genes associated with cardiac channelopathy and cardiomyopathy. Demographics, autopsy findings, medical history, and postmortem genetic testing results were collected for each case. The majority of decedents were adults (>25 years old, 52.7%), followed by infants (<12 months, 25.6%), young adults (19–25 years old, 11.4%), and children (1–18 years old, 10.2%). There were more males (64.2%) than females (35.8%). The racial/ethnic composition of decedents was 40.2% Black, 29.9% Hispanic, 22.4% White, 5.1% Asian/Pacific Islander, and 2.8% mixed/unspecified. Overall, 45.7% of decedents had a negative autopsy, and the remaining had one to four cardiac findings (cardiac hypertrophy, dilation, atherosclerosis, and fatty change). Twenty - seven pathogenic/likely pathogenic variants (P/LP) and 99 variants of uncertain significance (VUS) were identified in 10.6% and 39% of decedents respectively. P/LP and VUS were found in 51 cardiac genes of the total 95 genes, where MYBPC3, TTN (predicted truncating variants), KCNH2, RYR2 and DSP genes had more than two P/LP variants identified. Among the 73 decedents who were suspected of having cardiac arrhythmia or cardiomyopathy, 20.3% had P/LP variants and 47.9% had VUS; among 23 decedents who had hypertensive cardiovascular diseases and 20 decedents with a history of substance use, 13% and 30% had P/LP variants, respectively. There were 26 referrals from medical examiners for genetic counseling and the outcomes are discussed. The study demonstrates characteristics of the diverse population typically seen by medical examiners in an urban center and our results support a broader implementation of molecular testing in sudden death.

3) Review Articles and Book Chapter:

Two invited review articles were published (PMID: 28449775 and PMID: 30535908), entitled "Holistic Approach to Determine Cause of Autopsy-Negative Sudden Natural Death"[6] and "Genetic testing in sudden unexpected natural

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

death in the young: New York City Office of Chief Medical Examiner's experience and perspective"[7], respectively. In addition, one chapter "The Molecular Autopsy" was published in a book "Medical Legal Handbook" (Hammers and Fitzsimmons editors).

6. IMPLICATIONS FOR CRIMINAL JUSTICE POLICY AND PRACTICE IN THE UNITED STATES

This grant-funded project has directly improved how sudden unexpected death are investigated by NYC OCME as the medical examiners add molecular analysis of 95 cardiac disease genes in their toolboxes. The medical examiner in NYC routinely has jurisdiction over SUD cases. Generally, half of the deaths coming to the attention of the medical examiner will be due to natural causes. In a fraction of cases, a definitive cause of death may not be identified following an autopsy, various laboratory tests (e.g. toxicological, microbiological tests, etc.) and medical records review. Molecular diagnostic testing has become a powerful additional tool, allowing medical examiners to determine the cause of death in cases that would have previously remained undetermined. Through our casework and publications, we have demonstrated that timely and accurate diagnosis of medically important diseases in a decedent has significant impact on certifying death and on the surviving relatives of decedents by allowing them the opportunity to seek treatment for certain hereditary diseases. Furthermore, we have shown that that the quality of testing FFPE samples routinely collected in the ME's offices provide suboptimal results; therefore, it is important to save non-formalin fixed samples, such as *RNALater* tissue samples or bloodstain cards for accurate diagnosis through molecular analysis.

The recommendations made by the National Research Council of NAS in their 2009 report, entitled "Strengthening Forensic Science in the United States: A Path Forward", specifically stated "Investigations of unexplained sudden deaths, especially in young people and infants, would benefit from greater access to molecular diagnostics". The Council further pointed out that in addition to helping in medicolegal investigations, "Some molecular diseases are heritable, and it could be argued that the ME/C has a duty to identify these diseases and alert families about their presence." This research project has provided data to support those recommendations.

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

This project has impacted the criminal justice community in the United States and beyond through our scholarly publications, demonstrated the utility of routine molecular diagnostic testing in the setting of a medical examiner's office. Our publications in forensic science journal, forensic pathology journal, medical genetics journal, and health science journal have reached the audience beyond medical examiners' community. Multi-disciplinary experts have been brought together to advance study in sudden death through molecular autopsy work supported by this grant.

7. REFERENCES

1. Lin Y, Williams N, Wang D, et al. Applying High-Resolution Variant Classification to Cardiac Arrhythmogenic Gene Testing in a Demographically Diverse Cohort of Sudden Unexplained Deaths. Circ Cardiovasc Genet 2017;10.
2. Lin Y, Gryazeva T, Wang D, et al. Using postmortem formalin fixed paraffin-embedded tissues for molecular testing of sudden cardiac death: A cautionary tale of utility and limitations. Forensic Sci Int 2020;308:110177.
3. Dong J, Williams N, Cerrone M, et al. Molecular autopsy: using the discovery of a novel de novo pathogenic variant in the KCNH2 gene to inform healthcare of surviving family. Heliyon 2018;4:e01015.
4. Williams N, Marion R, McDonald TV, et al. Phenotypic variations in carriers of predicted protein-truncating genetic variants in MYBPC3: an autopsy-based case series. Cardiovasc Pathol 2018;37:30-33.
5. Williams N, Manderski E, Stewart S, Bao R, Tang Y. Lessons learned from testing cardiac channelopathy and cardiomyopathy genes in individuals who died suddenly: A two-year prospective study in a large medical examiner's office with an in-house molecular genetics laboratory and genetic counseling services. J Genet Couns 2020;29:293-302.
6. Sampson BA, Tang Y. Holistic Approach to Determine Cause of Autopsy-Negative Sudden Natural Death. J Am Coll Cardiol 2017;69:2146-2148.
7. Tang Y, Williams N, Sampson BA. Genetic testing in sudden unexpected natural death in the young: New York City Office of Chief Medical Examiner's experience and perspective. Forensic Sci Med Pathol 2019;15:481-484.

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.