

Research paper

Technical in-depth comparison of two massive parallel DNA-sequencing methods for formalin-fixed paraffin-embedded tissue from victims of sudden cardiac death

Emma Adolfsson^{a,b,*}, Alvida Qvick^{a,b}, Henrik Gréen^{c,d}, Daniel Kling^d, Cecilia Gunnarsson^{e,f}, Jon Jonasson^{a,e}, Anna Gréen^{a,b}

^a Department of Laboratory Medicine, Örebro University Hospital, Sweden

^b Faculty of Medicine and Health, Örebro University, Örebro Sweden

^c Division of Drug Research, Department of Biomedical and Clinical Sciences, Linköping University, Linköping, Sweden

^d Department of Forensic Genetics and Forensic Toxicology, National Board of Forensic Medicine, Linköping, Sweden

^e Department of Clinical Genetics and Department of Biomedical and Clinical Sciences, Linköping University, Linköping, Sweden

^f Centre for Rare Diseases in South East Region of Sweden, Linköping University, Linköping, Sweden

ARTICLE INFO

Keywords:

Paraffin embedding

Tissue fixation

Death, Sudden, Cardiac

Sequence analysis, DNA

DNA mutational analysis/methods

Massive parallel sequencing, MPS

ABSTRACT

Sudden cardiac death (SCD) is a tragic and traumatic event. SCD is often associated with hereditary genetic disease and in such cases, sequencing of stored formalin fixed paraffin embedded (FFPE) tissue is often crucial in trying to find a causal genetic variant. This study was designed to compare two massive parallel sequencing assays for differences in sensitivity and precision regarding variants related to SCD in FFPE material. From eight cases of SCD where DNA from blood had been sequenced using HaloPlex, corresponding FFPE samples were collected six years later. DNA from FFPE samples were amplified using HaloPlex HS, sequenced on MiSeq, representing the first method, as well as amplified using modified Twist and sequenced on NextSeq, representing the second method. Molecular barcodes were included to distinguish artefacts from true variants. In both approaches, read coverage, uniformity and variant detection were compared using genomic DNA isolated from blood and corresponding FFPE tissue, respectively.

In terms of coverage uniformity, Twist performed better than HaloPlex HS for FFPE samples. Despite higher overall coverage, amplicon-based HaloPlex technologies, both for blood and FFPE tissue, suffered from design and/or performance issues resulting in genes lacking complete coverage. Although Twist had considerably lower overall mean coverage, high uniformity resulted in equal or higher fraction of genes covered at $\geq 20\times$. By comparing variants found in the matched samples in a pre-defined cardiagnostic gene panel, HaloPlex HS for FFPE material resulted in high sensitivity, 98.0% (range 96.6–100%), and high precision, 99.9% (range 99.5–100%) for moderately fragmented samples, but suffered from reduced sensitivity (range 74.2–91.1%) in more severely fragmented samples due to lack of coverage. Twist had high sensitivity, 97.8% (range 96.8–98.7%) and high precision, 99.9% (range 99.3–100%) in all analyzed samples, including the severely fragmented samples.

1. Introduction

Sudden cardiac death (SCD) is an event defined as a sudden and unexpected death occurring within an hour of the onset of symptoms, or occurring in people found dead within 24 h of being asymptomatic, who presumably died due to a cardiac arrhythmia or hemodynamic catastrophe [1]. The overall incidence is unknown, but in North America and

Europe the annual incidence of SCD ranges between 50 and 100 per 100,000 individuals per year in the general population [2] and for the younger population of 1–40 years, the incidence is estimated to 1.3–8.5 per 100,000 per year [3,4]. Standard post-mortem autopsies to determine cause of death may be complemented by molecular autopsies to determine a probable cause of death through DNA analyzes. Commonly, the molecular investigation is performed on blood or fresh frozen tissue,

* Correspondence to: Department of Laboratory Medicine, Örebro University, Södra Grevrosengatan, 701 85 Örebro, Sweden.

E-mail address: emma.adolfsson@regionorebrolan.se (E. Adolfsson).

<https://doi.org/10.1016/j.fsigen.2021.102522>

Received 15 December 2020; Received in revised form 24 March 2021; Accepted 19 April 2021

Available online 26 April 2021

1872-4973/© 2021 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

where high quality DNA can easily be extracted. For suspected SCD, molecular autopsy from blood samples may determine the underlying cause of death in ~ 30% of cases [5].

SCD causes include inherited structural, functional and/or cardiac abnormalities [6,7]. The most common cause for SCD in the young is ventricular tachyarrhythmia. The disorders have common features in regard to genetics: they are familial rather than sporadic; they are autosomal dominant diseases, and the vast majority show incomplete penetrance. Furthermore, they show marked genetic heterogeneity with multiple types of variants [8]. Other causes of SCD are vascular disorders, especially aortic diseases and familial hypercholesterolemia (FH). To date, more than 100 genes have been associated with inherited cardiac disease and SCD [9]. Previous studies on the genetics of SCD, using targeted sequencing with gene panels of 70–192 genes associated with cardiac disease, showed that potentially pathogenic variants were found in 13–50% of the cases [3,10,11]. Clinical screening of relatives identifies an inherited genetic condition in 22–53% of the families, and targeted genetic DNA screening identifies pathological variants in genes coding for cardiac ion channels in up to 35% [12–14].

Genetic testing may be delayed due to a traumatized family and/or lack of information on the possibility of a genetic cause of the SCD. In yet other cases, patients seek genetic counseling, having relatives who died of SCD, often several years ago. Generally, long time storage of blood samples is not practiced in most forensic and pathology departments, entailing that if a DNA analysis is requested, the only source for DNA analysis may be FFPE material, which is usually saved after histopathological examination of tissues.

FFPE tissue is challenging as a source for DNA extraction and subsequent sequencing. Fixation of the tissue in buffered formalin preserves morphology and enables storage at ambient room temperature. However, the formalin induces damage to the DNA molecules and the fixation process results in shearing of nucleic acids. In turn, the fragmented

DNA directly influences the amount of templates available for PCR reactions used in subsequent library generation [15]. Formalin also causes crosslinks between DNA, proteins and histones, as well as deamination of cytosine resulting in sequencing errors such as C>T/G>A or A>G/T>C variants [16]. All of this leads to sequence artefacts that can be falsely interpreted as clinically important disease-causing variants.

In this study, the aim was to compare two different massive parallel sequencing assays for SCD using FFPE tissue as starting material. The first protocol employs improved amplicon-based HaloPlex High Sensitivity (HS) design for targeted amplification of selected genes, while the second protocol, based on Twist library preparation, utilizes hybridization capture-based amplification of the whole exome and subsequent bioinformatic filtering with virtual ad hoc gene panels. To avoid false positive variants, unique molecular identifiers (UMI) were used to tag individual DNA templates, permitting sequencing and PCR errors in high coverage NGS data to be accounted for. Twist's double stranded probes enable sequencing of both strands, which is an advantage. The methods were validated and compared against HaloPlex amplification of genomic DNA extracted from blood samples.

The current clinical protocol at the forensic department in Linköping, Sweden, uses HaloPlex and a cardiodiagnostic gene panel (hereafter referred to as Cardio Diagnostic Gene Panel, CDGP) containing 81 genes linked to 10 diseases, see Fig. 1. This method is based on sequencing of relatively long amplicons, which is rarely compatible with the fragmented DNA extracted from FFPE material. Moreover, since for each location, the sequencing reads represent only one of the two strands in the DNA double helix and UMIs are not utilized, the FFPE artefacts cannot be removed bioinformatically. Its application is therefore limited to cases where blood or fresh tissue samples are available and therefore there is a clinical need for genetic analysis compatible with DNA from FFPE.

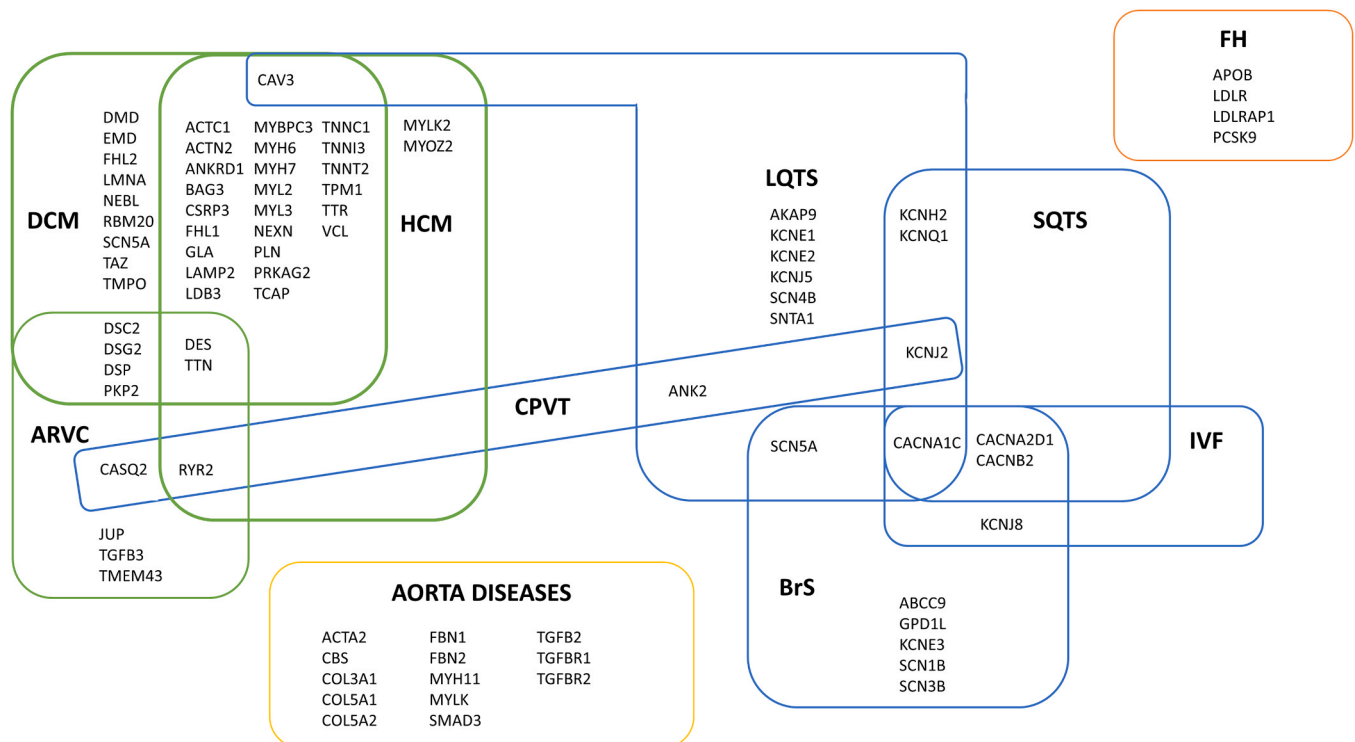


Fig. 1. Genes included in the cardiodiagnostic gene panel ($n = 81$) grouped by associated disease ($n = 10$). Diseases covered by the CDGP are aorta diseases (multigene panels including Marfan syndrome, Loeys-Dietz syndrome, familial thoracic aortic aneurysm disease), arrhythmogenic right ventricular cardiomyopathy (ARVC), Brugada syndrome (BrS), catecholaminergic polymorphic ventricular tachycardia (CPVT), dilated cardiomyopathy and conduction defects (DCM), familial hypercholesterolemia (FH), hypertrophic cardiomyopathy (HCM), idiopathic ventricular fibrillation (IVF), long QT syndrome (LQTS), and short QT syndrome (SQTs).

2. Materials and methods

The work described herein has been carried out in accordance with the code of ethics of the World Medical Association declaration of Helsinki. The research was approved by the regional ethics committee in Linköping, ethic permission nr 2016-389/61.

2.1. Subjects/specimen selection

Eight consecutive molecular autopsy cases were selected from two forensic medical departments. FFPE heart tissue was chosen for this study. The set of blood samples from these cases had previously been analyzed six years earlier in the forensic routine with HaloPlex using the CDGP containing 81 genes, as previously described for ARVC genes [17]. The study plan was to isolate DNA from the FFPE cardiac tissue samples for sequencing with HaloPlex HS and Twist, respectively, to compare the results of the two methods with results from previous fresh blood sequencing as a true result. Genomic standards NA12877 and NA12878 [18] were included as validation controls.

2.2. Design

2.2.1. HaloPlex HS for FFPE

HaloPlex HS for FFPE experiments were performed at department of department of Clinical Genetics, Linköping University Hospital. A HaloPlex design for the CDGP was made with Agilent's SureDesign (Agilent, Santa Clara, United States), using High Sensitivity and FFPE settings for inclusion of shorter fragments to avoid dropouts, with targeting of both strands and incorporation of 10–16 bp UMIs. (See [supplementary Table 1](#) for genes and transcripts).

2.2.2. Modified Twist with UMIs and Westburg library preparation

Twist whole exome sequencing of FFPE-samples was performed at the Department of Laboratory Medicine, Örebro University Hospital. This modified Twist protocol uses 120 nucleotides whole exome target double stranded DNA probes to detect a 33 Mb protein-coding region. Capture was performed with Twist core exome based on CCDS accessions, with the addition of spiked-in RefSeq areas. To account for variation in input DNA quality, the protocol was adjusted using library preparation kits with options to include/exclude further fragmentation of DNA. UMIs were included to enhance the suitability for FFPE samples. After whole exome library preparation and sequencing, the CDGP including exons with 10 bp padding of the 81 genes was added as an ad hoc bioinformatic filter to detect genetic variants.

2.3. DNA extraction and quality control

DNA was extracted using QIAamp DNA FFPE Tissue Kit (Qiagen, Hilden, Germany). Concentration was determined using Qubit 1.8 (Invitrogen, Carlsbad, USA) and the Quant-it dsDNA BR Assay Kit (Life Technologies, Carlsbad, USA). Extent of fragmentation was assessed on Agilent's 4200 Tape Station system using Genomic DNA Screen Tape (Agilent, Santa Clara, United States) to obtain average fragment length and DNA integrity number (DIN).

2.4. Library preparation

2.4.1. HaloPlex HS for FFPE tissue

Library preparations were performed according to the manufacturer's protocol. DNA input was determined based on the extent of fragmentation in the FFPE-samples. 250 ng of DNA was used for samples with average fragment length > 2000 bp. 500 ng of DNA was used for samples with average fragment length 1000–2000 bp. All available DNA was used (656–2125 ng) for samples with fragment length < 1000 bp. For the genomic standards, 50 ng of DNA was used.

All incubation steps and PCR reactions were performed on a 2720

Thermal cycler (Applied Biosystems, Foster City, United States). Briefly, DNA from all samples was cleaved by eight different enzyme mixes at 37 °C for 30 min. Correct restriction patterns were assured by analyzing samples and the ECD control using 2100 Bioanalyser with High Sensitivity DNA Kit (Agilent, Santa Clara, United States) according to manufacturer's instructions.

After fragmentation, HaloPlex gene panel probes and sample indexes were hybridized to pooled fragmented DNA by incubating at 95 °C for 5 min, followed by 16 h hybridization at 58 °C. The hybridized pool was purified using AmPure XP beads (Beckman Coulter, Brea, United States). Fragments were circularized at 55 °C for 10 min, then captured using Dynabeads™ MyOne™ Streptavidin T1 (Life Technologies, Carlsbad, United States). PCR amplification was performed for 22 (moderately fragmented samples) or 24 cycles (severely fragmented samples) followed by library purification using AmPure XP beads. Library length was verified using Bioanalyzer and DNA High Sensitivity Assay. If artefact peaks were discovered, an additional washing step (one or two times as required) were performed. Finally, library concentrations were determined using Qubit.

2.4.2. Twist

For FFPE samples and for genomic standards, 250 ng and 50 ng of DNA were used to generate libraries, respectively. Degree of fragmentation was the decision point for choosing to include or exclude a fragmentation step in the library preparation method. All incubation steps and PCR reactions were performed in a Veriti™ 96-Well Thermal Cycler (Applied Biosystems, Foster City, USA).

2.4.2.1. Westburg DNA library preparation kits. FFPE samples with average fragment length > 1000 bp were selected for library preparation using Westburg library preparation kit. Briefly, DNA was enzymatically fragmented at 32 °C for 4 min, followed by DNA end-repair and dA-tailing at 65 °C for 30 min using Westburg NGS DNA Library Prep Kit (Westburg, Leusden, the Netherlands).

FFPE samples with average fragment length < 1000 bp were selected for library preparation using Westburg NoFrag Library Prep Kit. DNA was end-repaired and dA-tailed by incubating the samples for 30 min at 20 °C, followed by 30 min at 65 °C.

2.4.2.2. Adapter ligation using UMIs. Duplex adapters containing UMIs (xGEN Duplex Seq Adapters, Integrated DNA Technologies, Inc., Coralville, USA) were ligated to the fragments by incubating at 20 °C for 30 min. Ligated fragments were amplified in a pre-hybridization PCR (10–12 cycles) using KAPA HiFi HotStart Ready Mix (KAPA Biosystems, Wilmington, USA) and 8 nt long IDT duplex indexing primers (Integrated DNA Technologies, Inc., Coralville, USA). Libraries were purified using DNA purification beads (Twist Bioscience, San Francisco, USA). Library concentration was measured using Qubit 2.0 Fluorometer and Qubit™ dsDNA HS Assay Kit and fragment length using 4200 TapeStation System and D1000 Screen Tape.

2.4.2.3. Twist whole exome library preparation. Exome libraries were generated using Twist Human Core Exome Multiplex Hybridization Kit (Twist Bioscience, San Francisco, United States of America) in a modified protocol. Briefly, eight individual libraries (1500 ng of DNA) were pooled. The pool was hybridized at 70 °C for 16 h using whole exome probes with the addition of spiked-in RefSeq Human Panel (Twist Bioscience, San Francisco, United States of America). The exome library was amplified for 10 PCR-cycles using KAPA HiFi HotStart Ready Mix and xGEN Library Amplification Primer (xGEN Duplex Seq Adapters, Integrated DNA Technologies Inc., Coralville, USA). DNA concentration of the pool was determined using Qubit 2.0 Fluorometer and Qubit™ dsDNA HS Assay Kit and analyzed on TapeStation using D1000 Screen Tape.

2.5. Sequencing

HaloPlex HS sequencing was done using MiSeq (Illumina, San Diego, USA). Sample sheets were edited to include HaloPlex indexes and molecular barcodes. Either four pooled libraries were sequenced at 6 pM on MiSeq Reagent v2 300 cycle kit, or six pooled libraries at 8 pM on MiSeq Reagent v3 600 cycle kit. The MiSeq instrument was configured to collect fastq files for two index reads, as well as 151 cycles of paired end sequencing.

Twist sequencing of the eight pooled samples was performed on NextSeq (Illumina, San Diego, USA) on a high-output kit (300 cycles) using 1 pM of the whole exome pool as input following manufacturer's instructions.

2.6. Bioinformatic pipelines

All genomic positions and variants denoted in this manuscript uses the Genome Reference Consortium Human Build 37 (GRCh37) hg19 coordinates.

2.6.1. HaloPlex HS

An in-house bioinformatic pipeline was designed. The procedure is stepwise, I) Trimming adapters and barcodes from fastq-files and synchronizing reads for paired end alignment II) All low-quality base calls (Phred score ≤ 14) were substituted with 'N' III) Reads were aligned to hg19 reference genome using bwa mem [19]. Properly mapped read pairs were combined, and barcodes restituted. IV) Bam-files were created and indexed, and reads were tabulated with samtools mpileup [20]. V) All relevant regions (CDGP +/- 10 bp flanking sequences) were extracted. Variants were detected using an in-house script detailed in previous publication [17] and vcf and coverage files were created.

2.6.2. Twist

Fastq files were generated from the NextSeq run using bcl2fastq2 Conversion Software v2.20 set to trim TruSeq adapters, and to include UMIs in the read header as well as trimming the first five cycles in each read to exclude UMIs from read sequences. Reads were also trimmed 3' by 5 bp to remove UMI remnants. Fastq files were analyzed using Bcbio, with the aligner bwa-mem [19] and the callers GATK version v4.1.8.1 [21], Samtools version 1.9 [20] and Freebayes version v1.1.0-46-g8d2b3a0-dirty [22]. Duplicates were marked. Ensembl Variants from ensembl files from the three callers were used for variant comparison using hg38 as reference. Multisample vcf files were split into individual sample vcf files using bcf-tools. These variant and bam files were lifted from hg38 coordinates to hg19 coordinates using CrossMap [23] to enable variant comparisons. Variants had to be called by two out of three variant callers in order to be considered "true" variants.

3. Data analysis and statistics

3.1. Sequencing metrics

Sequencing data were analyzed using MultiQC [24] to include reads per sample, mapped reads, reads on target, and duplicates. Sequencing data was also analyzed in R using custom scripts to generate summary statistics such as sequence length, GC-content and nucleotide balance. In particular, a special focus was given to the CDGP subset regions, where data from both Twist and HaloPlex protocols are available.

3.2. Read coverage and uniformity in the CDGP

Sequencing data were analyzed using in-house scripts to determine coverage and uniformity in each gene of the CDGP. Measures of coverage included overall coverage per gene, percentage of bases covered > 20X per gene, percentage of bases covered 0X per gene, and exceedance coverage, expressed as percentage of bases covered 100%

between 0 and 50X. Measures of uniformity included fold80/fold 90 and percentage of bases covered at > 0.2X of the calculated mean coverage.

3.3. Variant comparison – validation samples

CLC Genomics workbench 20.0 (<https://digitalinsights.qiagen.com>) was used for all variant comparisons employing the hg19 build as reference genome. Platinum genomes NA12877/NA12878 were used to verify the setup of HaloPlex, HaloPlex HS and Twist by comparing the consensus variants found after amplification/sequencing/bioinformatics and filtering to the confident regions of the CDGP. Variants not overlapping between assays for corresponding samples were analyzed in Integrative Genomic Viewer [25] where coverage in the variant position/s, allele frequencies and number of UMIs present in the variant was evaluated.

For intra-sample comparisons, variants obtained using HaloPlex amplification of blood samples were considered true, since this method has previously been validated using Sanger data [17]. Variants from each blood sample amplified using HaloPlex were compared to the variants detected using HaloPlex HS and Twist, respectively, of the corresponding FFPE DNA. Additional variants found in FFPE reads were denoted extra variants and variants not found in the FFPE reads were denoted missed variants.

Sensitivity [shared variants / (shared variants + missed variants)] and precision/positive predictive value (PPV) [shared variants / (shared variants + extra variants)] were calculated for each method.

4. Results

4.1. Sequencing metrics

4.1.1. HaloPlex HS

Libraries were successfully generated and sequenced on MiSeq from all FFPE samples. The average read length ranged 73–101 bp, with shorter reads lengths from poorer quality samples. Reads per sample ranged from 2.8 to 8.6 Mb in CDGP, despite equal input of all libraries in the sequencing reaction, with a tendency for more reads from higher quality samples. Percentage of duplicate were high; ranging from 14.4% to 49.2% without correlation to sample quality. The overall 20X coverage was higher for moderately fragmented samples (98.55–99.44%) compared to severely fragmented FFPE samples (76.19–90.44%). Same pattern was seen for average depth; 749–1565 X vs. 408–771 X, respectively (See Table 1 and Fig. 2). Skewed nucleotide balance and increased GC-content were observed in severely fragmented samples, see Supplementary Fig. 1 and Supplementary Fig. 2.

4.1.2. Modified Twist

Generating libraries from FFPE-samples using Twist required optimization of the protocol for each individual sample. Based on degree of fragmentation and DIN-score, the protocol was adjusted to include or exclude fragmentation, i.e. to use Westburg NGS kit for moderate fragmented samples or Westburg NoFrag kit for severely fragmented samples with average fragment length < 1000 bp. Fragmentation time was reduced from proposed 22–4 min for FFPE samples to achieve library lengths of the target length 375–425 bp. Using this strategy, libraries were successfully generated and sequenced on NextSeq.

Reads per sample ranged from 0.4 to 0.8 millions in CDGP. Percentage of duplicates ranged from 5.6% to 10%. Sequencing metrics showed that > 99.7% of reads were mapped, with > 60% reads on target (data not shown). Sequence reads were significantly longer than for HaloPlex HS (131 ± 8 bp vs 91 ± 10 bp, $p < 0.05$). Average sequence depth ranged from 57 to 171X, and coverage 20X ranged from 85.8% to 99.5%. Severely fragmented samples remained covered to a high degree. See Table 1 and Fig. 2.

Detailed data analysis did not show correlation between number of reads and initial DNA fragment length, DIN-score or mean insert size

Table 1

HaloPlex HS and Twist sequencing metrics. The table displays library length, average read length, reads in the cardiodiagnostic gene panel (CDGP), duplicates, average sequence depth and overall coverage in CDGP.

Sample fragmentation	Library length (bp)		Average read length (bp)		Reads per sample CDGP (M)		Duplicates (%)		Average depth (X)		Coverage 20X CDGP (%)	
	HaloPlex HS	Twist	HaloPlex HS	Twist	HaloPlex HS	Twist	HaloPlex HS	Twist	HaloPlex HS	Twist	HaloPlex HS	Twist
Moderate 1	277	362	101	137	8.6	0.8	28.2	8.9	1565	171	99.4	99.5
Moderate 2	273	375	100	136	5.8	0.7	43.8	5.6	1006	152	99.1	99.4
Moderate 3	273	343	100	135	4.1	0.8	49.2	7.6	748	146	98.7	99.5
Moderate 4	292	345	95	135	6.2	0.8	34.1	6.6	1074	161	98.8	99.8
Moderate 5	308	475	97	138	8.5	0.4	34.3	10	1506	78	99.4	98.7
Severe 1	257	269	83	129	4.8	0.5	14.4	9.6	771	88	90.2	98.2
Severe 2	229	255	81	124	4.7	0.5	32.7	7.8	738	81	90.4	97.7
Severe 3	271	214	73	116	2.8	0.4	42.0	8.0	408	57	76.2	85.8

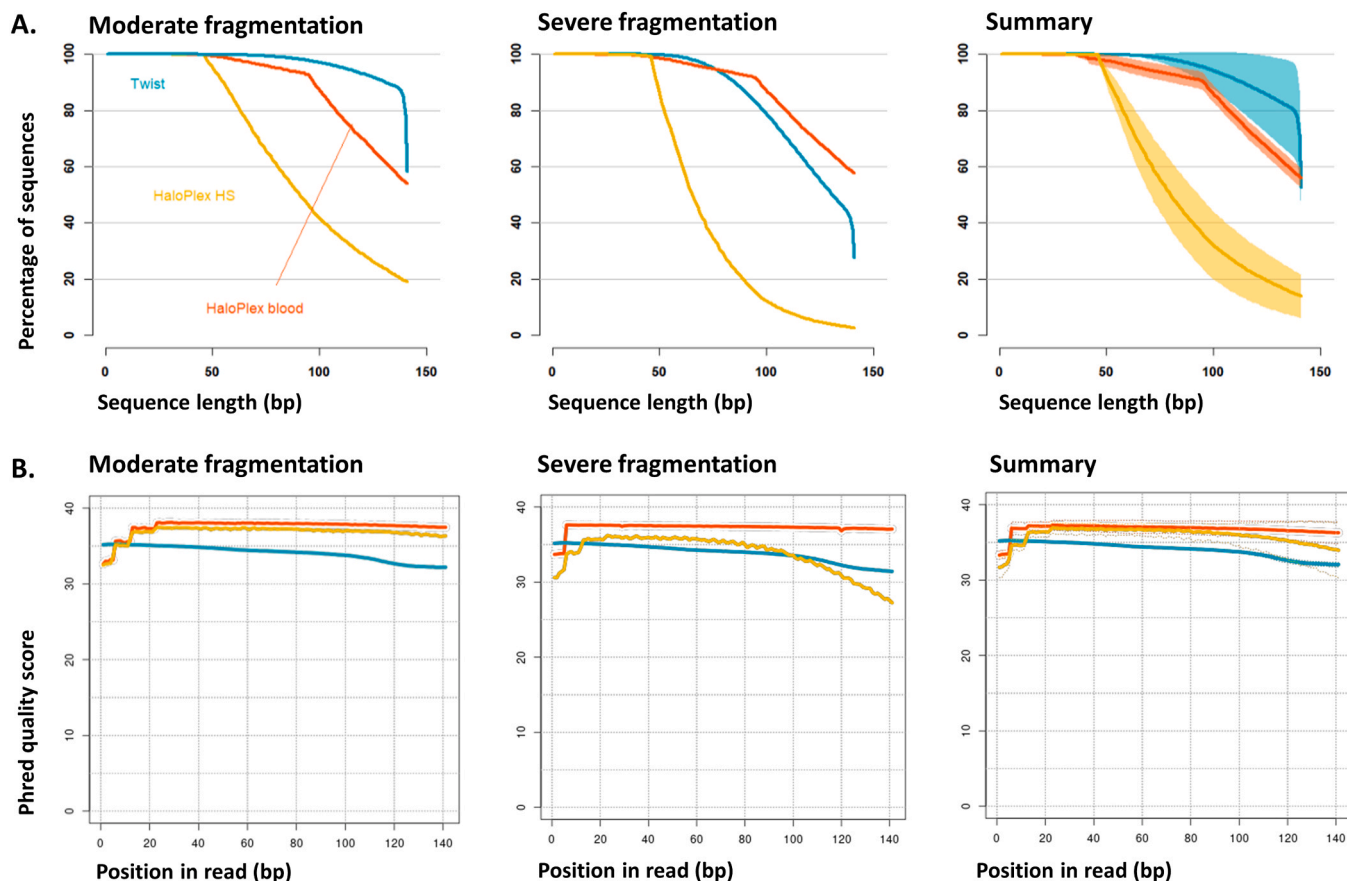


Fig. 2. Raw sequencing metrics. A: Sequence length distribution, given as the percentage of adapter-trimmed reads that reach a specified length. The graph illustrates results from Twist (blue), HaloPlex HS (yellow) and HaloPlex blood (red). Left: FFPE sample with moderate fragmentation. Middle: FFPE sample with severe fragmentation. Right: average and one SD confidence bands for all the eight validation samples in this study (ranging from moderate to severe fragmentation). B: Quality per read position, given as the Phred quality score of adapter trimmed-reads. Twist (blue), HaloPlex HS (yellow) and HaloPlex blood (red). Left: FFPE sample with moderate fragmentation. Middle: FFPE sample with severe fragmentation. Right: average and one SD confidence bands for all eight validation samples in this study. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

(data not shown). For severely fragmented samples, GC-content increased and nucleotide imbalances with increasing percentage of G/C in relation to A/T were observed compared to samples with moderate fragmented DNA, but to a lesser degree for HaloPlex HS. See [Supplementary Fig. 1](#) and [Supplementary Fig. 2](#).

4.2. Read coverage and uniformity in the CDGP

Average sequence depth was higher in HaloPlex HS than Twist (408–1565X, vs 57–171X, $p < 0.001$). Using $\text{Pr} > 20\text{X}$ as threshold for sufficient coverage for hereditary disease [26] Twist had higher fraction

of bases covered above threshold compared to HaloPlex HS for FFPE, while HaloPlex HS had higher fraction of bases lacking coverage (Pr0X). Read coverage uniformity was assessed using fold-80 and fold-90 base penalty metrics, defined as the fold change of non-zero reads coverage needed to bring 80% or 90% of the bases to the observed mean coverage [27]. For Twist, lower values were obtained indicating less variability and hence higher uniformity. Fold90 showed that HaloPlex for blood samples had difficulties with uniformity in certain regions. Percentage of bases covered at $> 0.2\text{X}$ of the calculated mean coverage was higher and more uniform for Twist. HaloPlex for blood was included for comparisons. See [Fig. 3](#).

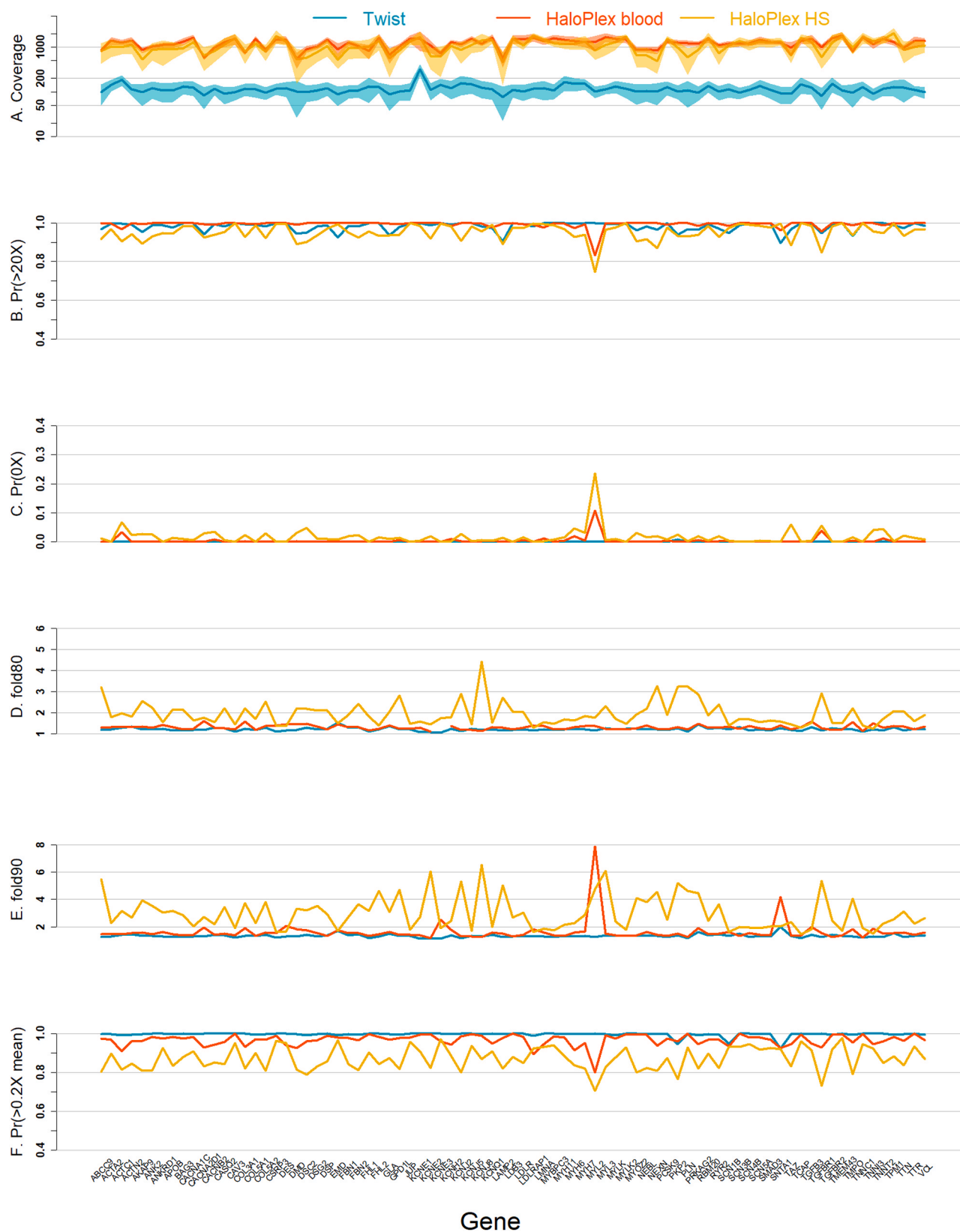


Fig. 3. Summary statistics for each gene in the cardiodiagnostic gene panel (CDGP) (x-axis). From top to bottom, A) aligned average coverage with a one SD confidence band, B) Fraction of bases covered to at least 20X, C) Fraction of bases not covered (0X), D) fold80, a measure of uniformity calculated as the fold change to raise 80% of the bases to the mean coverage, E) fold90, calculated as the fold change to raise 90% of the bases to the mean coverage, F) Fraction of bases above 0.2 times the mean coverage.

Since fold-80/fold-90 ignores base positions with zero coverage, another measure of uniformity was applied; the fraction of bases in each gene covered at least 20% of the observed mean coverage for that gene. A similar pattern as with fold-80/fold-90 was observed, where Twist displayed a less variable and more uniform coverage, whereas HaloPlex technologies resulted in variable coverage, with specific genes having less uniformity within a gene. Detailed investigation of genes not achieving $\geq 20X$ coverage in all bases within the gene was performed. See Fig. 4.

The HaloPlex method on fresh blood lacked complete coverage in 15 genes; *ACTC1*, *AKAP9*, *COL5A1*, *FBN2*, *KCNH2*, *KCNQ1*, *MYH6*, *MYL2*, *MYLK*, *PCSK9*, *RYR2*, *SMAD3*, *TGFBF1*, *TNNT2*, and *TTN*. For HaloPlex HS for FFPE, additional genes lacked complete coverage (*ANK2*, *LAMP2*, *MYL7* and *TMP*). For Twist, the design lacked coverage in exon 6 of the *PKP2* gene region resulting in 94.8% coverage $20X$ in the *PKP2* gene. (See [supplementary Table 2](#), [Table 3](#), [Table 4](#)).

4.3. Variant comparisons – validation samples

Initially, HaloPlex HS and Twist were validated using platinum genomes samples NA12877 and NA12878 with known reference variants as well as defined regions of high confidence [18]. Variant detection using both technologies showed high concordance with reference variants in NA12877 and NA12878 (data not shown).

4.3.1. HaloPlex HS

In total, HaloPlex HS detected 1226 variants in the FFPE samples, compared to HaloPlex sequencing of matched blood, which detected 1300 variants. The two methods shared 1215 variants. Missed variants ($n = 85$) were identified, the majority were due to lack of coverage in HaloPlex HS ($n = 69$), resulting in a false negative rate of 6.5% (0–25.8%). Extra variants ($n = 11$) were identified with a false positive rate of 0.8% (0–2.0%). Sensitivity of HaloPlex HS for FFPE was 93.5% (74.2–98.1%), moderately fragmented samples having higher sensitivity compared to the severely fragmented samples. PPV was 99.8% (99.1–100%). See [Fig. 5A](#) and [Table 2](#).

Using IGV [25] the missed and/or extra variants detected with HaloPlex HS were analyzed. Two false positives and four true negatives were identified. The false positives were a nine bp deletion located in a repeat region in chr2:21266775, and a SNV (G/C) located in chr2:17946330. The false negative variants were SNVs located in chr1:237957161 (A/G), chr1:2377754200 (C/T) and chr2:179598034 (C/T). These variants displayed skewed allele ratios in HaloPlex HS and were hence not called. The skewed allele ratios were only noted with HaloPlex HS, not with Twist. (See [Supplemental Table 5](#) for further details.).

4.3.2. Twist

In total, Twist detected 1295 variants in the FFPE samples, compared to HaloPlex sequencing of matched blood, which detected 1300 variants. The two methods shared 1271 variants. Missed variants ($n = 29$) resulted in false negative rate of 2.2% (range 1.3–3.4%). Extra variants ($n = 24$) were detected resulted in false positive rate of 1.8% (range 1.3–3.4%). Sensitivity for Twist was 97.8% (range 96.8–98.7%). PPV for Twist was 99.9% (range 99.3–100%). See [Fig. 5B](#) and [Table 2](#). In contrast to HaloPlex HS, sensitivity was even for all analyzed samples. See [Fig. 6](#).

Using IGV, three unique false negative variants were discovered that were not present in FFPE samples, only in the corresponding blood. One variant was a SNV located in chr2:179458591 (C/T) present in only one blood sample. One variant was a 2 bp deletion located in chr14:23858272–23858273 (GG/-) present in two samples. The last variant was a SNV chr21:35821821 (T/C) present in five of the eight samples. This variant was present in the bam-files for Twist, but only in reads with MAPQ = 0, indicating that there may be mapping ambiguities, possibly due to a pseudogene. We found only one false positive

variant with Twist. This was a 3 bp deletion (GCT/-) located in chr9:137534099–137534101, which presented in only one of the matched samples. (See [Supplementary Table 6](#) for further details.).

5. Discussion

5.1. Methodological differences and their attribution to the outcome

We suggest that the observed differences in read coverage and uniformity in the present study can be explained in terms of performance of HaloPlex and Twist technologies. In general, amplicon methods have been described as having higher on-target rates, whereas hybridization capture-based approaches have been shown to have lower on-target rate, but better uniformity [28]. Limited data indicate that use of the Twist protocol results in more uniform coverage and may be suitable for highly degraded DNA [29]. Indeed, this is observed in this study. Although the average sequence depth is higher for HaloPlex technologies than for Twist ([Fig. 4](#)), the fold80/fold90 measurements show a more uniform coverage for Twist ([Fig. 3](#)), resulting in higher $20X$ coverage in the genes, and higher sensitivity in the more fragmented samples compared to HaloPlex HS ([Fig. 6](#)).

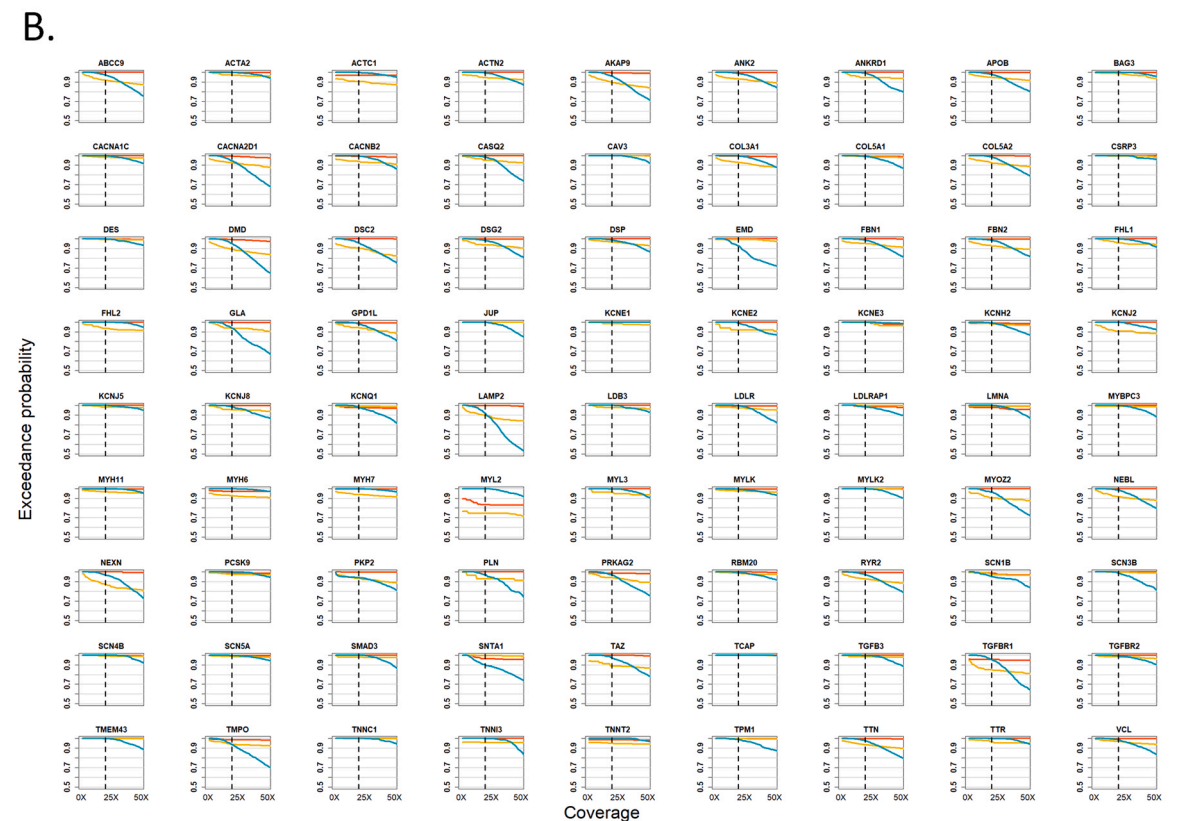
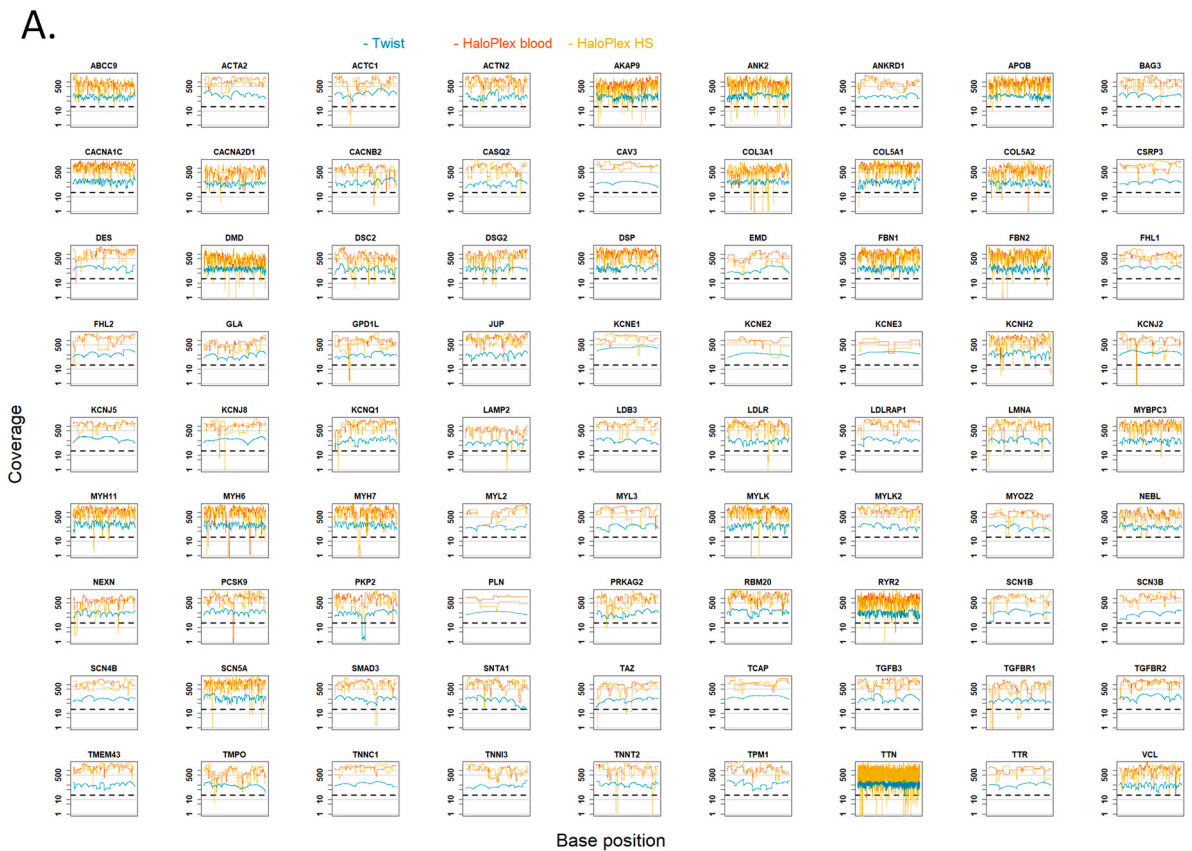
An advantage of the Twist protocol is the possibility to exclude enzymatic fragmentation and only do end repair and poly-A ligation, making it easier to generate libraries from severely fragmented samples, compared to HaloPlex HS. In this study, the DNA extracted from FFPE tissue ranged from moderately fragmented ~ 3000 bp in average length, to severely fragmented samples as short as ~ 300 bp ([Table 2](#)). One major observation was the considerable reduction in sensitivity for severely fragmented DNA samples using the HaloPlex HS protocol. The same was not observed using the Twist protocol ([Table 2](#), [Fig. 6](#)). This clearly gives the Twist technology an edge, since the clinical laboratory must rely on a robust method for all types of FFPE samples, regardless of the degree of fragmentation. However, some of the lost coverage using HaloPlex HS method on extensively fragmented samples might be mitigated through increasing library input, and thereby the amount of reads.

Twist requires less DNA input to generate reliable results. The ability to use lower amount of DNA is an advantage since extracting large amounts of DNA from FFPE tissue is difficult. A potential disadvantage of Twist is that samples are pooled at the hybridization step, prohibiting re-sequencing of a single sample without repeating the hybridization step. Twist protocols are also labor intensive and time consuming.

Focusing on coverage data, Twist lacked coverage in the *PKP2* gene. Per design, probes were lacking in exon 6 of the gene. This resulted in complete absence of coverage for this particular region in all samples (see [supplemental Fig. 3](#)). Pathogenic/VUS variants correlated to cardiomyopathy are located in this region according to ClinVar [30]. In comparison, using HaloPlex technologies, we observed lack of coverage in several regions of interest ([supplemental Table 3](#)). For example, neither HaloPlex nor HaloPlex HS completely covers *KCNQ1*, a clinically relevant gene containing several pathogenic variants linked to LQTS and other cardio-related diseases.

5.2. Variant detection - comparisons

For both technologies, the concordance between variants found in whole blood and the variants found in the corresponding FFPE samples was high ($>99.999\%$). However, for HaloPlex HS, lack of coverage resulted in increasing number of missed variants as the quality of DNA dropped ([Table 2](#)). The most fragmented sample contained only 116 variants, as opposed to the matched blood which contained 155 variants. In this sample, 40 variants were undetected in the FFPE-DNA, resulting in false negative rate of 25.8% and a sensitivity of just 74.2%. This phenomenon was not seen with Twist where less variants were undetected, and hence sensitivity for the severely fragmented samples was higher (97.7–98.7% for Twist compared to 74.2–91.1% for



(caption on next page)

Fig. 4. Coverage graphs. A. Per gene coverage (based on aligned reads) in the cardiognostic gene panel (CDGP) for each arbitrary base covered in the panel (x-axis). For comparative reasons, the y-axis (coverage) is on log scale. 20X coverage is marked by a dashed line. For each gene, the average coverage for HaloPlex HS (yellow line), HaloPlex blood (red line) and Twist (blue line) are shown. Values are based on all samples, n = 8. HaloPlex and HaloPlex HS for FFPE results in much higher average coverage. However, as shown in the figure, HaloPlex HS has several regions where coverage fails to exceed the 20X threshold, indicating design and/or performance issues. This also happens to HaloPlex and Twist, but to a lesser extent. B. Per gene exceedance probability graphs for the cardiognostic gene panel (CDGP). The graphs show fraction (y-axis) of bases covered 0–50X (x-axis). Values are based on all samples, n = 8. Ideally, each gene is covered 100% at least up to the clinically important threshold value 20X, marked by a dashed line. Twist, in blue, displays rapid drop in fraction of bases covered after 20X, compared to HaloPlex and HaloPlex HS. There are also genes where coverage drops before 20X, for example *EMD* and *SNTA1*. With HaloPlex HS for FFPE, on the other hand, several genes never reach coverage 100%, even at 0–20X. Examples include *FHL2*, *JUP*, *MYL2* and *TGFBRI*. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

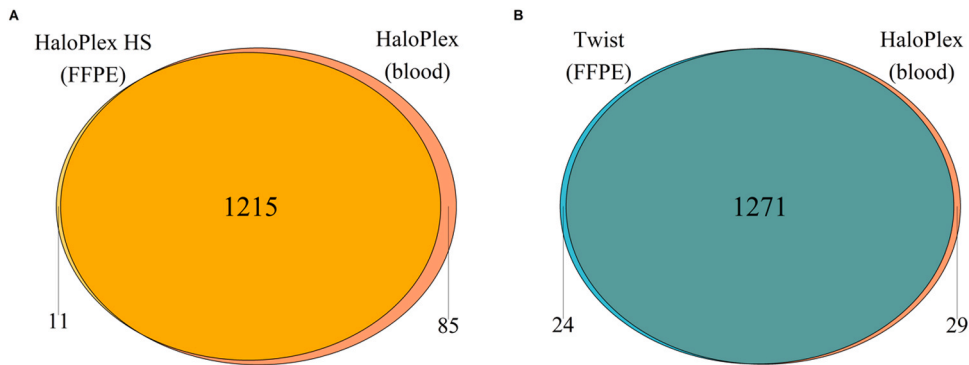


Fig. 5. Variant comparison. Variants detected in HaloPlex blood, HaloPlex HS and Twist in the cardiognostic gene panel (CDGP) and their overlap. A: HaloPlex for blood and HaloPlex HS for FFPE samples variant comparison. B: HaloPlex for blood and Twist, variant comparison.

Table 2
Comparison between HaloPlex HS and Twist, variant detection (n = 8). Both methods are compared against the golden standard HaloPlex for matched blood samples. Twist has lower fraction of missed variants (2.2% vs 6.5%), slightly higher fraction of extra variants (1.8% vs 0.8%), higher sensitivity (97.8% vs 93.5%) and specificity/PPV (99.9% vs 99.8%).

Sample	Fragment size (kb)	Missed variants (%)		Extra variants (%)		Sensitivity (%)		Positive predictive value (%)	
		HaloPlex HS	Twist	HaloPlex HS	Twist	HaloPlex HS	Twist	HaloPlex HS	Twist
Moderate 1	3.3	0	3.2	1.1	1.6	100	96.8	100	100
Moderate 2	2.3	3.4	2.7	2.0	3.4	96.6	97.3	100	99.3
Moderate 3	2.0	2.6	1.5	1.0	1.5	97.4	98.5	99.5	100
Moderate 4	2.2	2.1	3.4	0.7	2.1	97.9	96.6	100	100
Moderate 5	2.5	1.9	1.9	0	1.3	98.1	98.1	100	100
Severe 1	0.8	9.8	2.3	0.6	1.7	90.2	97.7	100	100
Severe 2	0.5	8.9	1.5	0.7	1.5	91.1	98.5	100	100
Severe 3	0.3	25.8	1.3	0.6	1.9	74.2	98.7	99.1	100
Average	1.7	6.5	2.2	0.8	1.8	93.5	97.8	99.8	99.9

HaloPlex HS).

Lack of coverage in the HaloPlex for blood golden standard method resulted in variants assigned as extra variants in Twist (n = 22) or HaloPlex HS (n = 6) and is reflected in the false positive rate and PPV, which then would give the impression of Twist being less specific than HaloPlex HS. We therefore excluded the extra variants that were due to lack-of-coverage in the HaloPlex reference method when calculating PPV (see [Supplementary Table 5](#) and [Table 6](#)).

5.3. FFPE artefacts - our findings in relation to others

A study by Lin et al. [31] compared variants found in FFPE cardiac tissue samples against paired whole blood samples, using HaloPlex amplification and MiSeq sequencing technology. They found both false negative and false positive variants compared to blood, but also intra-sample variation in the FFPE material, which they attributed to depth of sampling in the tissue block. The most common sequencing errors were A:T > G:C nucleotide change. In comparison, Baudhuin et al. [32] achieved 100% concordant accuracy between FFPE material, dried blood spots and whole blood using SureSelect for amplification and MiSeq sequencing technology. Bhagwate et al. [33] investigated

concordance between FFPE and fresh frozen breast tissue material. By adding UMIs and bioinformatic variant detection and filtering strategies such as simple filtering criterion to exclude variants below 5% alternative allele frequency, they achieved sensitive and confident variant calling from FFPE material although the number of variants were higher compared to the fresh frozen samples.

In this study, no false positive variants were attributed to FFPE artefacts. We employed UMI as a strategy to avoid this, by inclusion of molecular barcodes in both HaloPlex HS and Twist. Also, with Twist, three separate calling tools were used and applied a rule stating that a variant should only be called if detected by at least two separate callers. We conclude these strategies to be successful, but careful examination of each detected variant is needed in clinical practice to ensure that the variant is present in several amplicons (HaloPlex) and originates from several unique DNA molecules to discriminate a true variant from a FFPE artefact when performing variant interpretation from FFPE-DNA.

5.4. Clinical considerations

The choice between a targeted sequencing approach using a gene panel and a whole exome sequencing approach with post-sequence ad

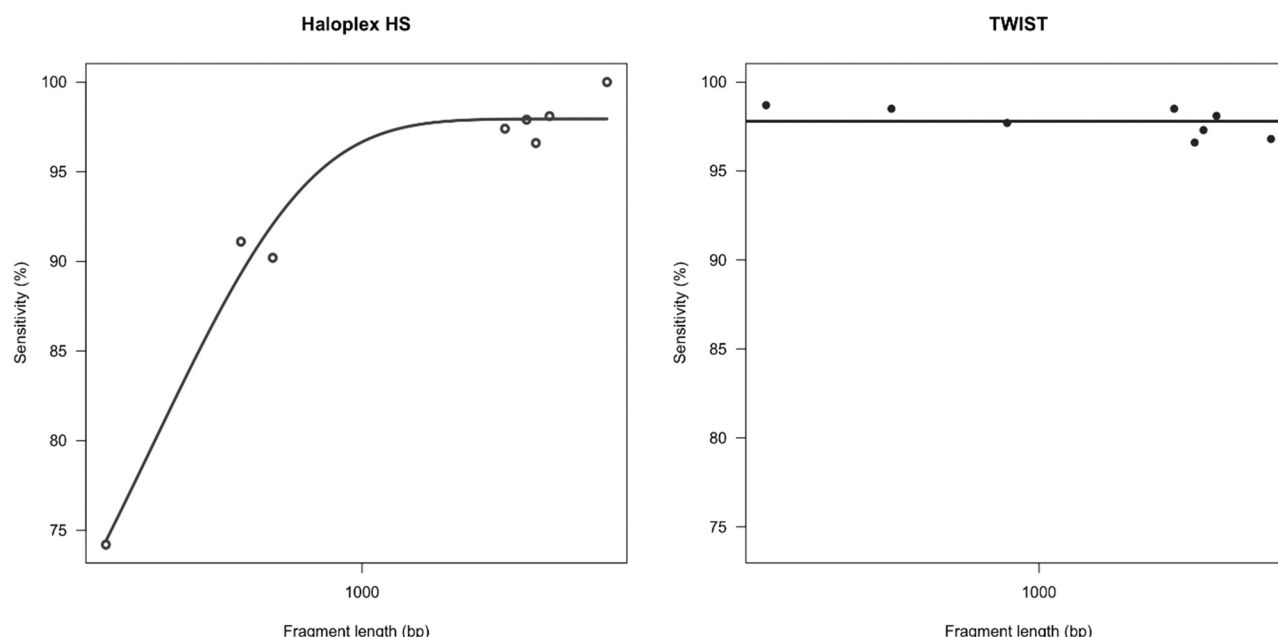


Fig. 6. Sensitivity graph for HaloPlex HS (open circles) and Twist (closed circles) versus average fragment length (log scale) in each FFPE sample. In the HaloPlex HS graph, the solid line represents a fitted logistic curve. In the Twist graph, the solid line represents the mean sensitivity from the Twist analyzes.

hoc filtering may be a cost-dependent decision. The considerations range from coverage and cost, to the possibility of including more genes as new knowledge emerges. Each clinical laboratory must make a choice here.

The strategy of using gene panels have several advantages over WES. The most common genes associated with SCD can be investigated to a lower cost, with possibly shorter turnaround time. HaloPlex have the great advantage that even a single sample can be prepared and sequenced at a time to a comparatively low expense, by using smaller flow cells. Twist, on the other hand, has the advantage of requiring less sequencing for optimal coverage but the disadvantage of requiring eight samples simultaneously for lowest cost of library preparation. Another option is to take advantage of the uniform coverage obtained with Twist technology, but use a customized Twist panel based on, for example the CDGP panel used in this study, and excluding all other probes in the Twist design used here. Twist yielded roughly 1/10 the number of reads in the CDGP region compared to HaloPlex HS, while still providing equal or better results in terms of coverage statistics. Assuming same on-target rate and duplicate percentage for the customized probes design as for core exome, an estimated 1.5 million reads would be needed to obtain same coverage as seen with Twist WES. As a consequence, instead of sequencing 8 exome FFPE samples on a NextSeq high-output kit - theoretically yielding 800 M paired-end reads-, ~ 500 custom panel FFPE samples could be sequenced. This corresponds to a reduction in sequencing cost of ~ 98%. Compared to HaloPlex HS, 16–20 samples could be sequenced on a MiSeq v2 300 cycle kit (theoretically yielding 24–30 million paired-end reads) using the customized Twist panel, i.e. 4–5 times the number of samples sequenced when using HaloPlex HS. Sequencing cost compared to HaloPlex HS would therefore be reduced by 75–80%.

5.5. Limitations of the study

The massive parallel sequencing experiments described in this study were performed over a period of six years at two different laboratories using different sequencing instruments and different pipelines with different callers for the analysis, but the same DNA. At each site, pipelines used for bioinformatic analysis were optimized for their respective protocol. Although the bioinformatics has not been uniform, the high

concordance of variant calls means that the pipelines themselves are not expected to impact the validity of the observations or the conclusions of this study.

6. Conclusion

SCD is a traumatic event for relatives of the victim, and finding a potential underlying genetic cause is very important for the relatives since genetic testing of the family may prevent further deaths [34]. Since the only remaining material may be FFPE samples from autopsy, genetic testing must be compatible with this material. In this study, two approaches of DNA sequencing for detecting genetic variants causing SCD were contrasted. Despite significant challenges with NGS data from FFPE, a combination of strategies (inclusion of UMIs, optimization of library preparation, bioinformatics variant detection and filtering strategies) allowed us to get sensitive and confident variant calling from both NGS approaches. However, it should also be added that we have from our own experience encountered FFPE samples that have been so degraded that they could not be sequenced with either the Twist or HaloPlex HS methods described here.

In conclusion, this study provides methodological insights to sequencing of FFPE samples for detection of clinical variants relevant to SCD using Twist technology and HaloPlex HS. We show that although using a lower median coverage for the capture-based Twist technology the sequencing results in a more uniform coverage, higher fraction of the target region above 20X, a high sensitivity and maintained specificity for moderately and severely fragmented FFPE samples, as compared to HaloPlex HS. Also in comparison, the amplicon-based HaloPlex HS method resulted in a higher number of missed variants i.e. lowered sensitivity due to insufficient coverage in especially the more severely fragmented samples. In summary, the Twist technology resulted in more variants being identified in the FFPE samples (also compared to the HaloPlex sequencing of DNA from fresh tissue), and that genetic variants could be identified in cases with severely fragmented DNA, making it a suitable technology for sequencing of FFPE samples.

Funding

This work was supported by ALF funding Region Örebro County, and

Örebro County Council Research committee.

Declaration of competing interest

None.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at [doi:10.1016/j.fsigen.2021.102522](https://doi.org/10.1016/j.fsigen.2021.102522).

References

- [1] E.J. Benjamin, S.S. Virani, C.W. Callaway, A.M. Chamberlain, A.R. Chang, S. Cheng, S.E. Chiuve, M. Cushman, F.N. Delling, R. Deo, S.D. de Ferranti, J. F. Ferguson, M. Fornage, C. Gillespie, C.R. Isasi, M.C. Jiménez, L.C. Jordan, S. E. Judd, D. Lackland, J.H. Lichtman, L. Lisabeth, S. Liu, C.T. Longenecker, P. L. Lutsey, J.S. Mackey, D.B. Matchar, K. Matsushita, M.E. Mussolino, K. Nasir, M. O'Flaherty, L.P. Palaniappan, A. Pandey, D.K. Pandey, M.J. Reeves, M. D. Ritchey, C.J. Rodriguez, G.A. Roth, W.D. Rosamond, U. Sampson, G.M. Satou, S. H. Shah, N.L. Spartano, D.L. Tirschwell, C.W. Tsao, J.H. Voeks, J.Z. Willey, J. T. Wilkins, J.H. Wu, H.M. Alger, S.S. Wong, P. Muntner, S. American Heart Association Council on Epidemiology and Prevention Statistics Committee and Stroke Statistics, Heart disease and stroke statistics-2018 update: a report from the American Heart Association, *Circulation* 137 (12) (2018) e67–e492, <https://doi.org/10.1161/cir.0000000000000558>.
- [2] G.I. Fishman, S.S. Chugh, J.P. Dimarco, C.M. Albert, M.E. Anderson, R.O. Bonow, A.E. Buxton, P.S. Chen, M. Estes, J. Jouven, R. Kwong, D.A. Lathrop, A. M. Mascette, J.M. Nerbonne, B. O'Rourke, R.L. Page, D.M. Roden, D.S. Rosenbaum, N. Sotoodehnia, N.A. Trayanova, Z.J. Zheng, Sudden cardiac death prediction and prevention: report from a National Heart, Lung, and Blood Institute and Heart Rhythm Society Workshop, *Circulation* 122 (22) (2010) 2335–2348, <https://doi.org/10.1161/CIRCULATIONAHA.110.976092>.
- [3] M.K. Larsen, S.L. Christiansen, C.L. Hertz, R. Frank-Hansen, H.K. Jensen, J. Banner, N. Morling, Targeted molecular genetic testing in young sudden cardiac death victims from Western Denmark, *Int. J. Leg. Med.* 134 (1) (2020) 111–121, <https://doi.org/10.1007/s00414-019-02179-x>.
- [4] J. Isbister, C. Semsarian, Sudden cardiac death: an update, *Intern. Med. J.* 49 (7) (2019) 826–833, <https://doi.org/10.1111/imj.14359>.
- [5] D.J. Tester, M.J. Ackerman, The role of molecular autopsy in unexplained sudden cardiac death, *Curr. Opin. Cardiol.* 21 (3) (2006) 166–172, <https://doi.org/10.1097/01.hco.0000221576.33501.83>.
- [6] M.-A. Jazayeri, M.P. Emert, Sudden cardiac death: who is at risk? *Med. Clin. N. Am.* 103 (5) (2019) 913–930, <https://doi.org/10.1016/j.mcna.2019.04.006>.
- [7] S.M. Al-Khatib, W.G. Stevenson, M.J. Ackerman, W.J. Bryant, D.J. Callans, A. B. Curtis, B.J. Deal, T. Dickfeld, M.E. Field, G.C. Fonarow, A.M. Gillis, C.B. Granger, S.C. Hammill, M.A. Hlatky, J.A. Joglar, G.N. Kay, D.D. Matlock, R.J. Myerburg, R. L. Page, 2017 AHA/ACC/HRS guideline for management of patients with ventricular arrhythmias and the prevention of sudden cardiac death: a report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines and the Heart Rhythm Society, *J. Am. Coll. Cardiol.* 72 (14) (2018) e91–e220, <https://doi.org/10.1016/j.jacc.2017.10.054>.
- [8] A. Barsheshet, A. Brenyo, A.J. Moss, I. Goldenberg, Genetics of sudden cardiac death, *Curr. Cardiol. Rep.* 13 (5) (2011) 364–376, <https://doi.org/10.1007/s11886-011-0209-y>.
- [9] S.L. Christiansen, C.L. Hertz, L. Ferrero-Miliani, M. Dahl, P.E. Weeke, LuCamp, G. L. Ottesen, R. Frank-Hansen, H. Bundgaard, N. Morling, Genetic investigation of 100 heart genes in sudden unexplained death victims in a forensic setting, *Eur. J. Hum. Genet.* 24 (12) (2016) 1797–1802, <https://doi.org/10.1038/ejhg.2016.118>.
- [10] Y. Hata, K. Kinoshita, K. Mizumaki, Y. Yamaguchi, K. Hirono, F. Ichida, A. Takasaki, H. Mori, N. Nishida, Postmortem genetic analysis of sudden unexplained death syndrome under 50 years of age: a next-generation sequencing study, *Heart Rhythm* 13 (7) (2016) 1544–1551, <https://doi.org/10.1016/j.hrthm.2016.03.038>.
- [11] C.L. Hertz, S.L. Christiansen, L. Ferrero-Miliani, M. Dahl, P.E. Weeke, LuCamp, G. L. Ottesen, R. Frank-Hansen, H. Bundgaard, N. Morling, Next-generation sequencing of 100 candidate genes in young victims of suspected sudden cardiac death with structural abnormalities of the heart, *Int. J. Leg. Med.* 130 (1) (2016) 91–102, <https://doi.org/10.1007/s00414-015-1261-8>.
- [12] L.M. Nunn, L.R. Lopes, P. Syrris, C. Murphy, V. Plagnol, E. Firman, C. Dalageorgou, E. Zorio, D. Domingo, V. Murday, I. Findlay, A. Duncan, G. Carr-White, L. Robert, T. Bueser, C. Langman, S.P. Fynn, M. Goddard, A. White, H. Bundgaard, L. Ferrero-Miliani, N. Wheeldon, S.K. Suvana, A. O'Beirne, M.D. Lowe, W.J. McKenna, P. M. Elliott, P.D. Lambiase, Diagnostic yield of molecular autopsy in patients with sudden arrhythmic death syndrome using targeted exome sequencing, *Eurospace* 18 (6) (2016) 888–896, <https://doi.org/10.1093/eurospace/euv285>.
- [13] H.L. Tan, N. Hofman, I.M. van Langen, A.C. van der Wal, A.A. Wilde, Sudden unexplained death: heritability and diagnostic yield of cardiologic and genetic examination in surviving relatives, *Circulation* 112 (2) (2005) 207–213, <https://doi.org/10.1161/circulationaha.104.522581>.
- [14] E.R. Behr, C. Dalageorgou, M. Christiansen, P. Syrris, S. Hughes, M.T. Tome Esteban, E. Rowland, S. Jeffery, W.J. McKenna, Sudden arrhythmic death syndrome: familial evaluation identifies inheritable heart disease in the majority of families, *Eur. Heart J.* 29 (13) (2008) 1670–1680, <https://doi.org/10.1093/eurheartj/ehn219>.
- [15] H. Do, A. Dobrovic, Sequence artifacts in DNA from formalin-fixed tissues: causes and strategies for minimization, *Clin. Chem.* 61 (1) (2020) 64–71, <https://doi.org/10.1373/clinchem.2014.223040>.
- [16] S.J. McDonough, A. Bhagwate, Z. Sun, C. Wang, M. Zschunke, J.A. Gorman, K. J. Kopp, J.M. Cunningham, Use of FFPE-derived DNA in next generation sequencing: DNA extraction methods, *PLoS One* 14 (4) (2019) e0211400, <https://doi.org/10.1371/journal.pone.0211400>.
- [17] A. Gréen, H. Gréen, M. Rehnberg, A. Svensson, C. Gunnarsson, J. Jonasson, Assessment of HaloPlex amplification for sequence capture and massively parallel sequencing of arrhythmogenic right ventricular cardiomyopathy-associated genes, *J. Mol. Diagn.* 17 (1) (2015) 31–42, <https://doi.org/10.1016/j.jmoldx.2014.09.006>.
- [18] M.A. Eberle, E. Frittilas, P. Krusche, M. Källberg, B.L. Moore, M.A. Bekritsky, Z. Iqbal, H.Y. Chuang, S.J. Humphray, A.L. Halpern, S. Kruglyak, E.H. Margulies, G. McVean, D.R. Bentley, A reference data set of 5.4 million phased human variants validated by genetic inheritance from sequencing a three-generation 17-member pedigree, *Genome Res.* 27 (1) (2017) 157–164, <https://doi.org/10.1101/gr.210500.116>.
- [19] H. Li, Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. 2013: arXiv:1303.3997v1 [q-bio.GN]. <https://github.com/lh3/bwa>.
- [20] H. Li, B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, R. Durbin, Genome Project Data Processing S., The Sequence Alignment/Map format and SAMtools, *Bioinformatics* 25 (16) (2009) 2078–2079, <https://doi.org/10.1093/bioinformatics/btp352>.
- [21] G.A. Van der Auwera, M.O. Carneiro, C. Hartl, R. Poplin, G. Del Angel, A. Levy-Moonshine, T. Jordan, K. Shakir, D. Roazen, J. Thibault, E. Banks, K.V. Garimella, D. Altschuler, S. Gabriel, M.A. DePristo, From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline, *Curr. Protoc. Bioinforma.* 43 (1110) (2013) 11.10.1–11.10.33, <https://doi.org/10.1002/0471250953.bi1110s43>.
- [22] G.M. Erik Garrison, Haplotype-based variant detection from short-read sequencing. 2012, arXiv:1207.3907v2 [q-bio.GN].
- [23] H. Zhao, Z. Sun, J. Wang, H. Huang, J.P. Kocher, L. Wang, CrossMap: a versatile tool for coordinate conversion between genome assemblies, *Bioinformatics* 30 (7) (2014) 1006–1007, <https://doi.org/10.1093/bioinformatics/btt730>.
- [24] M.M. Philip Ewels, Sverker Lundin, Max Källér, MultiQC: summarize analysis results for multiple tools and samples in a single report, *Bioinformatics* 32 (2016) 3047–3048, <https://doi.org/10.1093/bioinformatics/btw354>.
- [25] J.T. Robinson, H. Thorvaldsdóttir, W. Winkler, M. Guttman, E.S. Lander, G. Getz, J.P. Mesirov, Integrative genomics viewer, *Nat. Biotechnol.* 29 (1) (2011) 24–26, <https://doi.org/10.1038/nbt.1754>.
- [26] S.W. Kong, I.H. Lee, X. Liu, J.N. Hirschhorn, K.D. Mandl, Measuring coverage and accuracy of whole-exome sequencing in clinical context, *Genet. Med.* 20 (12) (2018) 1617–1626, <https://doi.org/10.1038/gim.2018.51>.
- [27] E.S. Lander, M.S. Waterman, Genomic mapping by fingerprinting random clones: a mathematical analysis, *Genomics* 2 (3) (1988) 231–239, [https://doi.org/10.1016/0888-7543\(88\)90007-9](https://doi.org/10.1016/0888-7543(88)90007-9).
- [28] E. Samorodnitsky, B.M. Jewell, R. Hagopian, J. Miya, M.R. Wing, E. Lyon, S. Damodaran, D. Bhatt, J.W. Reeser, J. Datta, S. Roychowdhury, Evaluation of hybridization capture versus amplicon-based methods for whole-exome sequencing, *Hum. Mutat.* 36 (9) (2015) 903–914, <https://doi.org/10.1002/humu.22825>.
- [29] E. Leproust, New double-stranded twist on NGS adds more depth to each run, *Genet. Eng. Biotechnol. News* 39 (11) (2019) 60–62, <https://doi.org/10.1089/gen.39.11.19>.
- [30] M.J. Landrum, J.M. Lee, M. Benson, G.R. Brown, C. Chao, S. Chitipiralla, B. Gu, J. Hart, D. Hoffman, W. Jang, K. Karapetyan, K. Katz, C. Liu, Z. Maddipati, A. Malheiro, K. McDaniel, M. Ovetsky, G. Riley, G. Zhou, J.B. Holmes, B. L. Kattman, D.R. Maglott, ClinVar: improving access to variant interpretations and supporting evidence, *Nucleic Acids Res.* 46 (D1) (2018) D1062–D1067, <https://doi.org/10.1093/nar/gkx1153>.
- [31] Y. Lin, T. Gryazeva, D. Wang, B. Zhou, S.Y. Um, L.S. Eng, K. Ruiter, L. Rojas, N. Williams, B.A. Sampson, Y. Tang, Using postmortem formalin fixed paraffin-embedded tissues for molecular testing of sudden cardiac death: a cautionary tale of utility and limitations, *Forensic Sci. Int.* 308 (2020), 110177, <https://doi.org/10.1016/j.forsciint.2020.110177>.
- [32] L.M. Baudhuin, C. Leduc, L.J. Train, R. Avula, M.L. Kluge, K.E. Kotzer, P.T. Lin, M. J. Ackerman, J.J. Maleszewski, Technical advances for the clinical genomic evaluation of sudden cardiac death: verification of next-generation sequencing panels for hereditary cardiovascular conditions using formalin-fixed paraffin-embedded tissues and dried blood spots, *Circ. Cardiovasc. Genet.* 10 (6) (2017), <https://doi.org/10.1161/circgenetics.117.001844>.
- [33] A.V. Bhagwate, Y. Liu, S.J. Winham, S.J. McDonough, M.L. Stallings-Mann, E. P. Heinzen, J.I. Davila, R.A. Vierkant, T.L. Hoskin, M. Frost, J.M. Carter, D. C. Radisky, J.M. Cunningham, A.C. Degnim, C. Wang, Bioinformatics and DNA-extraction strategies to reliably detect genetic variants from FFPE breast tissue samples, *BMC Genom.* 20 (1) (2019) 689, <https://doi.org/10.1186/s12864-019-6056-8>.
- [34] W. Aase, S. Eva-Len, Plötslig hjärtdöd bland unga - viktigt hitta riskindivider, *Läkartidningen* 8 (2010).