MiSeq™ Dx Instrument

illumina®

Package Insert for Instruments with MOS v4

FOR IN VITRO DIAGNOSTIC USE. FOR EXPORT ONLY.

Intended Use

The MiSeqDx instrument is intended for targeted sequencing of DNA libraries from human genomic DNA extracted from peripheral whole blood or formalin-fixed, paraffin-embedded (FFPE) tissue, when used with *in vitro* diagnostic (IVD) assays performed on the instrument. The MiSeqDx instrument is not intended for whole genome or *de novo* sequencing. The MiSeqDx instrument is to be used with registered and listed, cleared, or approved IVD reagents and analytical software.

Principles of Procedure

The Illumina MiSeqDx is intended for targeted resequencing of human DNA using Illumina sequencing consumables and libraries prepared from human genomic DNA extracted from peripheral whole blood or FFPE tissue using registered and listed, cleared, or approved IVD reagents. Libraries are prepared by amplifying targets and adding sample indexes and capture sequences. Sample libraries are captured on a flow cell and sequenced on the instrument using sequencing by synthesis (SBS) chemistry. SBS chemistry uses a reversible-terminator method to detect single nucleotide bases as they are incorporated into growing DNA strands. The Real-Time Analysis (RTA) software performs image analysis and base calling, and assigns a quality score to each base for each sequencing cycle. When primary analysis finishes, secondary analysis on the MiSeqDx instrument processes base calls.

Processing typically includes demultiplexing, FASTQ file generation, alignment, variant calling, and generation of variant call format (VCF) files that contain information about variants found at specific positions in a reference genome. The MiSeqDx uses different modules for secondary analysis depending on the workflow.

Dual Boot Configuration

The dual boot configuration includes the hardware, software, and installation procedures to allow the MiSeqDx instrument to run both *in vitro* diagnostic (IVD) and research use only (RUO) sequencing assays. The dual boot configuration allows the user to switch between the diagnostic mode of the instrument to the research mode of the instrument. The radio-frequency identifications (RFIDs) on sequencing consumables prevent RUO sequencing reagents from being utilized in diagnostic sequencing runs.

Limitations of the Procedure

- For *in vitro* diagnostic use.

- Results presented in the labeling were obtained with representative assay panels using peripheral whole blood or cell lines for germline performance and FFPE tissue or FFPE cell lines for somatic performance with the reagents and software modules described. The Germline Variant and Somatic Variant Modules were developed for the purpose of evaluating performance with representative assays. Performance characteristics are provided for information purposes only. The validation testing presented only serves to exemplify the instrument's general capabilities and does not establish the instrument's capabilities or suitability with respect to any specific claims. All diagnostic tests developed for use on this instrument require full validation for all aspects of performance.
- This product is limited to delivering the following:
  - Sequencing output ≥ 5 Gb at read length of 2 x 150 bp
  - Reads Passing Filter ≥ 15 million at read length of 2 x 150 bp
  - Bases higher than Q30 ≥ 80% at read length of 2 x 150 bp
    Equal or greater than 80% of bases have Phred scale quality scores greater than 30, indicating base call accuracy greater than 99.9%.
- The MiSeqDx instrument has only been validated to sequence human DNA libraries extracted from peripheral whole blood or FFPE tissue. Libraries generated from other specimen types should not be used with this instrument for *in vitro* diagnostic use. Performance of this instrument for sequencing microbial or viral nucleic acids from clinical specimens has not been established.
- The MiSeqDx is intended for *in vitro* diagnostic use with registered and listed, cleared, or approved IVD reagents or assays. Reagent limitations and performance characteristics described in this package insert are based on representative assays and software modules. For IVD assays, refer to the assay-specific package insert for intended use, variants detected, and sample type.
- Indel (insertions, deletions, and combinations thereof) content of length greater than 25 bp is not aligned by the assay software. Consequently, indels of length greater than 25 bp are not detectable by the assay software.
- The system has been validated for the detection of single nucleotide variants (SNVs) and up to 25 bp deletions and 24 bp insertions when used with the Germline and Somatic Variant module software. For somatic calling, at a variant frequency of 0.05, 25 bp deletions and 18 bp insertions were detected.
- Amplicon reads with extreme variant content might not be aligned by the assay software, resulting in the region being reported as wild-type. Such extreme content includes:
  - Reads containing more than three indels.
  - Reads of length at least 30 bp with SNV content greater than 4% of the total amplicon target length (excluding probe regions).
  - Reads of length less than 30 bp with SNV content greater than 10% of the total amplicon length (including probe regions).
- Large variants, including multi-nucleotide variants (MNVs) and large indels, might be reported as separate smaller variants in the output VCF file.

- Deletion variants may be filtered or missed when spanning two tiled amplicons if the deletion length is greater than or equal to the overlap between the tiled amplicons.
- The system cannot detect indels if they occur directly adjacent to a primer and there is no overlapping amplicon. For regions with overlapping amplicons, the assay cannot detect deletions when the region of overlap is smaller than the size of deletion to be detected. For example, if the region of overlap between two adjacent amplicons is two (2) bases, the assay cannot detect any deletions including both of those bases. A single base deletion at either of those bases can be detected.
- As with any hybridization-based library preparation workflow, underlying polymorphisms, mutations, insertions, or deletions in oligonucleotide-binding regions can affect the alleles being probed. Consequently, the calls made during sequencing are also affected. For example:
  - A variant in phase with a variant in the primer region may not be amplified resulting in a false negative.
  - Variants in the primer region could prevent the amplification of the reference allele resulting in an incorrect homozygous variant call.
  - Indel variants in the primer region may cause a false positive call at the end of the read adjacent to the primer.
- Indels may be filtered due to strand bias if they occur near the end of one read and are soft-clipped during alignment.
- Small MNVs have not been validated.
- Copy number variants or structural variants, such as fusions or translocations, have not been validated.
- Germline-specific limitations:
  - The MiSeqDx system using the Germline Variant Module is designed to deliver qualitative results for germline variant calling (ie, homozygous, heterozygous, wild type).
  - When used with the Germline Variant Module, the minimal coverage per amplicon needed for accurate variant calling is 150x. The number of samples and the total number of bases targeted affect coverage. GC-content and other genomic content can affect coverage.
  - Copy number variation can affect whether a variant is identified as homozygous or heterozygous.
  - Variants in certain repetitive context are filtered out in the VCF files. The RMxN repeat filter is used to filter variants if all or part of the variant sequence is present repeatedly in the reference genome adjacent to the variant position. For germline variant calling, at least 9 repeats in the reference are required for a variant to be filtered, and only repeats with length up to 5 bp are considered (R5x9).
- Somatic-specific limitations:
  - The MiSeqDx system using the Somatic Variant Module is designed to deliver qualitative results for somatic variant calling (ie, presence of a somatic variant with a variant frequency greater than or equal to 0.026 with a limit of detection of 0.05).
  - When used with the Somatic Variant Module, the minimal coverage per amplicon needed for accurate variant calling is 450x per oligonucleotide pool. The number of samples and the total number of bases targeted affect coverage. GC-content and other genomic content can affect coverage.

- Variants in certain repetitive context are filtered out in the VCF files. The RMxN repeat filter is used to filter variants in all or part of the variant sequence is present repeatedly in the reference genome adjacent to the variant position. For somatic variant calling, at least 6 repeats in the reference are required for the variant to be filtered, and only repeats with length up to 3 bp are considered (R3x6).
- The Somatic Variant Module cannot differentiate between germline and somatic variants. The module is designed to detect variants across a range of variant frequencies, but variant frequency cannot be used to differentiate somatic variants from germline variants.
- Normal tissue in the specimen impacts the detection of variants. The reported limit of detection is based on a variant frequency relative to the total DNA extracted from both tumor and normal tissue.

Product Components

The Illumina MiSeqDx consists of the following:

MiSeqDx Instrument (Catalog # DX-410-1001)

The following software components are required for MiSeqDx instrument operation and data analysis:

Software Application	Function	Description
MiSeq Operating Software (MOS)	Controls instrument operation	The MOS software application manages the operation of the instrument during sequencing and generates images for use by Real-Time Analysis (RTA) software. For more information, refer to *MiSeqDx Instrument Reference Guide for MOS v4 (document # 200010452)*.
Real-Time Analysis (RTA)	Performs primary analysis	The RTA software application converts the images generated by MOS for each tile per cycle of the sequencing run into base call files, which are inputs for Local Run Manager analysis modules. The RTA software application does not contain a user interface.
Local Run Manager	Interface for module selection	The Local Run Manager software is an on instrument integrated solution for user management, executing secondary analysis, and monitoring status. For more information, refer to *Local Run Manager v4 Software Guide for MiSeqDx (document # 200046657)*.

Storage and Handling

Element	Specification
Temperature	Transportation and Storage: -10°C to 40°C (14°F to 104°F) Operating Conditions: 19°C to 25°C (66°F to 77°F)

Element	Specification
Humidity	Transportation and Storage: Non-condensing humidity Operating Conditions: 30–75% relative humidity (non-condensing)

Equipment and Materials Required, Not Provided

Sequencing Consumables

MiSeqDx Reagent Kit v3 (Catalog # 20037124)

MiSeqDx Reagent Kit v3 Micro (Catalog # 20063860)

User-Supplied Consumables

Make sure that the following user-supplied consumables are available before beginning a run.

Consumable	Purpose
Alcohol wipes, 70% Isopropyl or Ethanol, 70%	Cleaning the flow cell glass and stage
Lab tissue, low-lint	Cleaning the flow cell stage
Lens paper, 4 x 6 in.	Cleaning the flow cell
Tween 20	Washing the instrument
Tweezers, square-tip plastic (optional)	Removing flow cell from flow cell shipping container
Water, laboratory-grade	Washing the instrument

Guidelines for Laboratory-Grade Water

Always use laboratory-grade water or deionized water to perform instrument procedures. Never use tap water.

Use only the following grades of water or equivalents:

- Deionized water
- Illumina PW1
- 18 Megaohm (MΩ) water
- Milli-Q water
- Super-Q water
- Molecular biology-grade water

Warnings and Precautions

WARNING

This set of reagents contains potentially hazardous chemicals. Personal injury can occur through inhalation, ingestion, skin contact, and eye contact. Ventilation should be appropriate for handling of hazardous materials in reagents. Wear protective equipment, including eye protection, gloves, and laboratory coat appropriate for risk of exposure. Handle used reagents as chemical waste and discard in accordance with applicable regional, national, and local laws and regulations. For additional environmental, health, and safety information, refer to the SDS at support.illumina.com/sds.html.

- Handle all blood specimens as if they are known to be infectious for human immunodeficiency virus (HIV), human hepatitis B virus (HBV), and other bloodborne pathogens agents (universal precautions).
- Failure to follow the procedures as outlined may result in erroneous results or significant reduction in sample quality.
- Use routine laboratory precautions. Do not pipette by mouth. Do not eat, drink, or smoke in designated work areas. Wear disposable gloves and laboratory coats when handling specimens and kit reagents. Wash hands thoroughly after handling specimens and kit reagents.
- Proper laboratory practices and good laboratory hygiene are required to prevent PCR products from contaminating reagents, instrumentation, and genomic DNA samples. PCR contamination may cause inaccurate and unreliable results.
- To prevent contamination, make sure that pre-amplification and post-amplification areas have dedicated equipment and consumables (eg, pipettes, pipette tips, heat blocks, vortexers, and centrifuges).
- Where appropriate, index-sample pairing must match the printed plate layout exactly. Local Run Manager automatically populates the index primers associated with the sample names, when entered in the module. The user is advised to verify the index primers associate with samples before starting the sequencing run. Mismatches between the sample and plate layout results in loss of positive sample identification and incorrect result reporting.
- Installation of user-supplied anti-virus software is strongly recommended to protect the computer against viruses. Consult the user manual for instructions on installation.
- Do not operate the MiSeqDx with any of the panels removed. Operating the instrument with any of the panels removed creates potential exposure to line voltage and DC voltages.
- Do not touch the flow cell state in the flow cell compartment. The heater in this compartment operates between 22°C and 95°C and may result in burns.
- The instrument weighs approximately 126 lbs. and could cause serious injury if dropped or mishandled.
- Immediately report any serious incidents related to this product to Illumina and the Competent Authority of the Member State in which the user and/or patient is established.

Instructions for Use

The following instructions for use of the MiSeqDx instrument require reagents provided in the MiSeqDx Reagent Kit v3.

Create Run with Local Run Manager

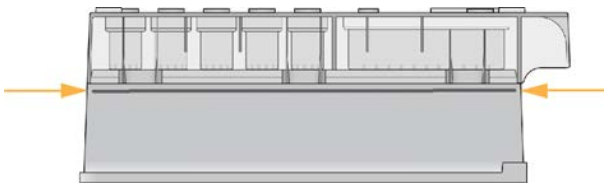
For detailed instructions for creating a run, refer to the *Local Run Manager v4 Software Guide for MiSeqDx (document # 200046657)* and the Local Run Manager module guide for the analysis module you use.

Prepare the Reagent Cartridge

The following instructions describe how to thaw reagents using a room temperature water bath.

1. Remove the reagent cartridge from -15° to -25°C storage.
2. Place the reagent cartridge in a water bath containing enough room temperature deionized water to submerge the base of the reagent cartridge up to the water line printed on the reagent cartridge. Do not allow the water to exceed the maximum water line.

Figure 1 Maximum Water Line



3. Allow the reagent cartridge to thaw in the room temperature water bath for approximately 60–90 minutes or until completely thawed.
4. Remove the cartridge from the water bath and gently tap it on the bench to dislodge water from the base of the cartridge. Dry the base of the cartridge. Make sure that no water has splashed on the top of the reagent cartridge.

Inspect the Reagent Cartridge

1. Invert the reagent cartridge ten times to mix the thawed reagents, and then inspect that all positions are thawed.
2. Inspect reagents in positions 1, 2, and 4 to make sure that they are fully mixed and free of precipitates.

NOTE It is critical that the reagents in the cartridge are thoroughly thawed and mixed to ensure proper sequencing.

3. Gently tap the cartridge on the bench to reduce air bubbles in the reagents.

NOTE The MiSeqDx sipper tubes go to the bottom of each reservoir to aspirate the reagents, so it is important that the reservoirs are free of air bubbles.

4. Place the reagent cartridge on ice or set aside at 2°C to 8°C (up to six hours) until ready to set up the run. For best results, proceed directly to loading the sample and setting up the run.

Prepare Samples for Sequencing

For directions on how to prepare sample libraries for sequencing, including library dilution and pooling, refer to the Instructions for Use section of the library prep package insert.

Dilution of sample libraries is dependent on the complexity of oligonucleotide pools. Optimization of cluster density on the MiSeqDx is required, and optimal cluster density varies depending on the particular library prep assay.

Load Sample Libraries onto Cartridge

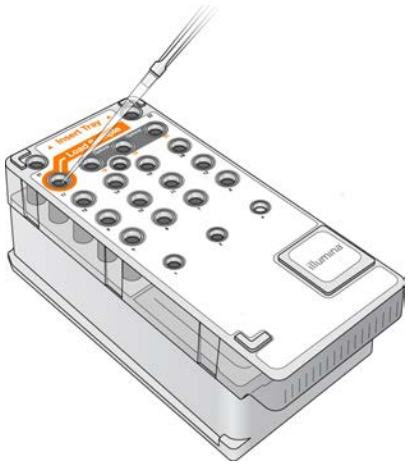
When the reagent cartridge is fully thawed and ready for use, you are ready to load samples into the cartridge.

1. Use a separate, clean, and empty 1 ml pipette tip to pierce the foil seal over the reservoir on the reagent cartridge labeled **Load Samples**.

NOTE Do not pierce any other reagent positions. Other reagent positions are pierced automatically during the run.

2. Pipette 600 µl prepared diluted amplicon library (DAL) sample libraries into the **Load Samples** reservoir. Avoid touching the foil seal.
3. Check for air bubbles in the reservoir after loading the sample. If air bubbles are present, gently tap the cartridge on the bench to release the bubbles.

Figure 2 Load Libraries



4. Proceed directly to the run setup steps using the MiSeq Operating Software (MOS) interface.

Run Setup

Refer to the *MiSeqDx Instrument Reference Guide for MOS v4 (document # 200010452)* for complete run setup instructions.

1. Log in to the MiSeqDx with your Local Run Manager software password.
2. From the Home screen of the MOS software, select **Sequence**.
3. Select a run from the list, and then select **Next**.
   A series of run setup screens open in the following order: Load Flow Cell, Load Reagents, Review, and Pre-Run check.
4. When the Load Flow Cell screen appears, clean and then load the flow cell.
5. Close the flow cell latch and flow cell compartment door.
   Both the latch and compartment door must be closed before beginning the run. When the flow cell is loaded, the software reads and records the RFID. A confirmation that the RFID was successfully read appears in the lower-right corner of the screen.
6. Follow the software prompts to load the MiSeqDx SBS Solution (PR2) bottle, make sure that the waste bottle is empty, and load the reagent cartridge.
   When the MiSeqDx SBS Solution (PR2) bottle and reagent cartridge are loaded, the software reads and records the RFID. A confirmation that the RFID was successfully read appears in the lower-right corner of the screen.
7. The Sequencing screen opens when the run begins. This screen provides a visual representation of the run in progress, including intensities and quality scores (Q-scores).

Results

Real-Time Analysis (RTA) is an integrated software that performs image analysis and base calling, and assigns a quality score to each base for each sequencing cycle. When primary analysis finishes, the module on the MiSeqDx instrument selected in *Create Run with Local Run Manager* on page 7 begins secondary analysis. Refer to assay-specific documentation for other workflows.

Quality Control Procedures

The MiSeqDx software evaluates each run, sample, and base call against quality control metrics. When required, the positive and negative controls included in library preparation should also be evaluated for expected results.

Performance Characteristics

All studies were performed on the MiSeqDx.

Germline studies used either the MiSeqDx Cystic Fibrosis 139-Variant Assay or the TruSeq Custom Amplicon Kit Dx reagents for library preparation. The two kits use identical library preparation reagents and have only one workflow difference: the number of polymerase chain reaction (PCR) cycles (25 and 28, respectively). The additional PCR cycles allow for a lower DNA input with the TruSeq Custom Amplicon Kit Dx (50 ng) relative to the MiSeqDx Cystic Fibrosis 139-Variant Assay (250 ng), as demonstrated in the DNA input study using the TruSeq Custom Amplicon Kit Dx. Libraries prepared with the MiSeqDx Cystic Fibrosis 139-Variant Assay were sequenced with the accompanying sequencing reagents in the kit. Libraries prepared with the TruSeq Custom Amplicon Kit Dx were sequenced with MiSeqDx Reagent Kit v3. The latter sequencing reagents have increased output relative to those in the MiSeqDx Cystic Fibrosis 139-Variant Assay.

The testing encompasses the sample throughput ranges supported by the MiSeqDx Reagent Kit v3 Micro. The MiSeqDx can support 1–96 samples/run, depending upon the assay. The MiSeqDx Reagent Kit v3 Micro is designed to support lower sample throughputs within this range for selected assays.

Somatic studies used the TruSeq Custom Amplicon Kit Dx with the MiSeqDx Reagent Kit v3.

The germline or somatic workflows, described for the TruSeq Custom Amplicon Kit Dx to prepare libraries for sequencing, were followed with analysis using the Germline Variant Module or Somatic Variant Module, respectively, with two exceptions. Studies using the one gene (germline performance; MiSeqDx Cystic Fibrosis 139-Variant Assay) or two gene (somatic performance) as representative mutation panels used assay-specific workflows and analysis modules.

NOTE Amplicon genomic content is summarized relative to the genomic strand that is sequenced. For amplicons designed against the minus strand, the reference genome content is the reverse complement (for example, PolyA regions on minus strand amplicons correspond to PolyT regions on the reference genome).

Definitions of Calculations Used in Performance Characteristics

- Positive Percent Agreement (PPA) is calculated as the proportion of loci classified as variants by a reference method that are correctly reported by the assay.
  - (# variant loci correctly reported by the assay) / (total # of variant loci)
    Variant loci reported by the assay that are concordant with the reference method are true positives (TPs). Variant loci reported as reference calls or as different variant calls by the assay are false negatives (FNs)
- Negative Percent Agreement (NPA) is calculated as the proportion of loci classified as wild-type by a reference method that are correctly reported by the assay.

- (# wild-type loci correctly reported by the assay)/ (total # of wild-type loci)
  Wild-type loci reported by the assay that are concordant with the reference method are true negatives (TNs). Wild-type loci reported as variants by the assay are false positives (FPs).

- Overall percent agreement (OPA) is calculated as the proportion of loci correctly reported by the assay relative to a reference method.
  - ((# variant loci correctly reported by the assay) + (# wild-type loci correctly reported by the assay)) / ((total # of variant loci) + (total # of wild-type loci))
- For variant calling applications, the calculations of PPA, NPA, and OPA do not include no calls (variant or reference loci not meeting one or more quality filters). Two studies specifically include no calls in their "% correct calls" metric, and this inclusion of no calls is noted for the applicable tables.
- Call rate is calculated as total number of loci passing filters divided by the total number of positions sequenced for chromosomes 1-22. Chromosomes X and Y are excluded. This metric does not consider the agreement of the calls with the reference method.

For performance characteristics related to pre-analytical factors (eg, extraction methods or DNA input), refer to the package insert for the applicable library preparation method.

Sample Indexing

Sample index primers, added during library preparation, assign a unique sequence to each sample DNA, allowing multiple samples to be pooled together into a single sequencing run. Sample indexing was tested for germline and somatic workflows.

A total of 96 samples indexes were tested with a representative assay designed to query a variety of genes covering 12,588 bases per strand across all 23 human chromosomes to verify the ability of the assay to consistently make a genotyping call for a given sample across different indexing primer combinations. The Y chromosome does not contain confident regions and was not evaluated. Eight unique sample were tested with 12 different indexing primer combinations per sample. Sample results from the Germline Analysis Module were compared to Platinum Genomes version 2016-01. PPA (SNVs and indels) exceeded 97% (true positive calls were at least 70 for SNVs, 38 for insertions, 36 for deletions) and NPA was 100% (at least 23,440 reference positions per index combination) for each of the 96 index combinations. Independently a single index was tested to verify that the MiSeqDx Reagent Kit v3 sequencing chemistry can support less than eight samples (the predecessor chemistry in the MiSeqDx Universal Kit 1.0 was limited to a minimum of eight samples). The single index had PPA values of 98.9% (180/182) for SNVs, 100% (38/38) for insertions, and 100% (46/46) for deletions. NPA was 100% (23,856/23,856).

Twelve replicates (24 libraries) of a sample were tested to measure index accuracy with somatic variants at frequencies from 0.05-0.10 using the Somatic Variant Module (two index combinations per replicate are used make somatic calls). PPA was 100% for SNVs (64/64), insertions (11/11), and deletions (19/19). NPA was 100% (at least 11590 reference positions per index combination) for all index combinations.

Sample Carryover

The MiSeqDx instrument workflow involves library preparation and sequencing of multiple samples plus controls processed all at once. The Sample Carryover study was conducted to evaluate if false positive results, due to carryover from well to well contamination during sample library preparation as well as run to run contamination between consecutive sequencing runs, impact test results. Somatic variants were used, as they can be detected at lower allele frequency events than germline variants.

The samples consisted of four genomic DNA samples from cell lines, each containing different panel mutations in a two gene representative assay. The samples were such that a mutation at a position in one will have a reference (wild type) sequence in the others.

Well to well carryover is defined as a failure mode potentially created by manual processing steps (pipetting, sample mix-up, etc.). In order to evaluate carryover from one sample well to another, two test runs were performed:

- A checkerboard layout of a high input genomic DNA (gDNA) sample containing a mutation in Gene 1 alternating with a sample of low input gDNA containing a mutant in Gene 2.
- A checkerboard layout of a high input gDNA sample containing an mutation in Gene 2 alternating with a sample of low input gDNA containing a mutation in Gene 1.

In each run, a total of 12 replicates were evaluated for false positives (eg, a Gene 1 mutation reported in a well designated as a Gene 2 mutant sample or vice versa).

Run to run carryover is defined as a failure mode potentially created by residue from a previous sequencing run. In order to determine if there is carryover between sequencing runs, two plates each containing 11 replicates of a single unique sample of high input gDNA plus a blank sample were prepared and sequenced consecutively on one MiSeqDx instrument and evaluated for false positives. The first run contained 11 replicates of a Gene 2 mutant sample plus one blank. The second run contained 11 replicates of a Gene 1 mutant sample plus 1 blank. The Gene 2 mutant sample library was sequenced first followed by a subsequent sequencing run with the Gene 1 mutant sample library, followed by another repeat sequencing run of the Gene 2 mutant sample library. If any Gene 2 mutations are observed in a Gene 1 mutant-only run, and vice versa, this would indicate carryover.

Zero false positives (0/24, 0%) due to *well to well* carryover were reported. All expected mutations were detected. Zero false positives (0/24, 0%) due to *run to run* carryover were reported. All expected mutations were detected. Zero false positives (0/48, 0%) due to *total* carryover (well to well and run to run carryover combined) were reported.

Germline Performance Characteristics

Studies described here used the Germline Variant Module to analyze sequencing data, except those studies that using the one gene panel where an assay-specific module was used.

Accuracy

The following study was conducted to assess the accuracy of the MiSeqDx instrument with the MiSeqDx Reagent Kit v3 and high quality DNA. The study used a representative assay designed to query a variety of genes covering 12,588 bases across 23 different chromosomes using 150 amplicons. The Y chromosome does not contain confident regions and was not evaluated. The 12 unique samples used in this study are from a single family - two parents and 10 children - frequently sequenced by multiple laboratories and sequencing methodologies. There are five samples from females and seven from males. Each of the samples was tested in duplicate. Accuracy was determined for SNVs, insertions, and deletions by comparing the study data to a well-characterized reference database. The reference database sequence (Platinum Genomes version 2016-01) was derived from the combination of multiple sequencing methodologies, publicly available data, and hereditary information. Confident genomic regions were defined based on this reference method unless otherwise specified. In total the samples were run eight times, the tables presented to demonstrate accuracy are based on data from the first run.

Table 1 contains the study data presented with positive and negative percent agreement on a per sample basis, where the variant results are compared to the well-characterized composite reference method for PPA calculations. The three variant types (SNVs, insertions, and deletions) are combined. Because the reference method only provides results for the single nucleotide variants and insertions/deletions, non-variant base results are compared to human genome reference sequence build hg19 for NPA calculations.

Table 1 Agreement of the MiSeqDx Instrument Base Call Results per Sample

Sample	Mean Call Rate	Total # Variants	Total # TP Variants	Total # FN Variants	Total # No Calls	Total # TN Calls	PPA	NPA	OPA
NA12877	> 99.9	152	152	0	4	24024	100	100	100
NA12878	> 99.9	270	266	0	4	23856	100	100	100
NA12879	> 99.9	192	190	1	1	24054	99.5	100	> 99.9

Sample	Mean Call Rate	Total # Variants	Total # TP Variants	Total # FN Variants	Total # No Calls	Total # TN Calls	PPA	NPA	OPA
NA12880	> 99.9	222	220	0	6	24052	100	100	100
NA12881	> 99.9	250	247	1	2	23862	99.6	100	> 99.9
NA12882	> 99.9	200	196	2	2	23962	99.0	100	> 99.9
NA12883	> 99.9	226	224	0	6	23870	100	100	100
NA12884	> 99.9	228	226	1	1	23942	99.6	100	> 99.9
NA12885	> 99.9	244	240	2	2	23942	99.2	100	> 99.9
NA12886	> 99.9	230	228	1	1	23888	99.6	100	> 99.9
NA12888	> 99.9	216	216	0	4	24002	100	100	100
NA12893	> 99.9	236	234	0	2	23810	100	100	100

The representative assay consisted of 150 amplicons designed to cover a variety of genomic content. The GC content of the amplicons ranged from 26-87%. Amplicons also had a range of single nucleotide (eg, PolyA, PolyT), dinucleotide, and trinucleotide repeats. Data were compiled on a per amplicon basis (Table 2) to determine the effect of genomic content on % correct calls. % correct calls consists of variant and reference calls and is less than 100% if there are either incorrect or no calls. No calls occur when one or more filters are not met for variant calling (eg, insufficient coverage).

Of the eight FN variants from Table 2, seven occurred with a 1 bp insertion on amplicon 111 which also contains PolyA homopolymer and GC content of 0.29. The remaining 1 FN (incorrect call) was due to an expected heterozygous SNV, on amplicon 125 with a GC content of 0.68, called as a homozygous variant. The SNV variant frequency was 0.71 which is above the 0.70 threshold for classification as homozygous variant. The amplicon with the lowest % correct calls (98.2%) was amplicon 17 with 40 no calls and containing AT repeats and GC content of 27%.

Table 2 Amplicon-level Accuracy for the MiSeqDx Instrument

Amplicon	Chromosome	Amplicon Start	Amplicon End	Analyzed Fragment Size	Bases in Confident Regions	Amplicon Genomic Content	GC Content	Total # Correct Calls	Total # Incorrect Calls	Total # No Calls	% Correct Calls
1	1	36450499	36450591	93	93	Indel	0.22	2232	0	0	100

Amplicon	Chromosome	Amplicon Start	Amplicon End	Analyzed Fragment Size	Bases in Confident Regions	Amplicon Genomic Content	GC Content	Total # Correct Calls	Total # Incorrect Calls	Total # No Calls	% Correct Calls
2	1	109465122	109465200	79	79	PolyA (5), PolyC (5), indel	0.38	1896	0	0	100
3	1	218353867	218353957	91	91	Indel	0.4	2184	0	0	100
4	1	223906657	223906748	92	92	Indel	0.49	2208	0	0	100
5	1	228526602	228526682	81	81	PolyG (5)	0.69	1944	0	0	100
6	1	236372039	236372108	70	70	PolyT (10), indel	0.39	1680	0	0	100
7	1	247812041	247812128	88	88	PolyA (5), CT(3), TAA(3), indel	0.27	2112	0	0	100
8	2	55862774	55862863	90	90	Indel	0.28	2160	0	0	100
9	2	87003930	87004009	80	80	Indel	0.38	1920	0	0	100
10	2	177016721	177016805	85	81	N/A	0.65	1944	0	0	100
11	2	186625727	186625801	75	75	PolyA (8)	0.35	1800	0	0	100
12	2	190323504	190323591	88	88	PolyT (5)	0.42	2112	0	0	100
13	2	200796740	200796826	87	87	PolyT (5), indel	0.31	2088	0	0	100
14	2	212245049	212245139	91	91	PolyT (5), PolyA (6), indel	0.3	2184	0	0	100
15	2	228147052	228147144	93	93	N/A	0.43	2232	0	0	100
16	2	235016350	235016422	73	73	PolyT (5), indel	0.42	1752	0	0	100
17	3	4466229	4466321	93	93	AT(3), indel	0.27	2192	0	40	98.2
18	3	46620561	46620643	83	83	N/A	0.43	1992	0	0	100
19	3	49851331	49851400	70	70	CT(3), indel	0.49	1680	0	0	100
20	3	189713161	189713248	88	88	PolyA (5), PolyT (5), PolyA (9), TG(3)	0.41	2112	0	0	100
21	3	190106030	190106104	75	74	Indel	0.57	1774	0	2	99.9
22	4	2233667	2233744	78	78	PolyA (6)	0.26	1872	0	0	100
23	4	7780541	7780637	97	97	PolyG (6), PolyT (5), PolyA (5)	0.42	2328	0	0	100
24	4	15688604	15688681	78	78	N/A	0.29	1872	0	0	100
25	4	56236521	56236586	66	62	PolyA (5), indel	0.36	1488	0	0	100
26	4	102839244	102839314	71	69	PolyA (5)	0.46	1656	0	0	100
27	4	164446743	164446804	62	62	PolyA (7), indel	0.27	1488	0	0	100

Amplicon	Chromosome	Amplicon Start	Amplicon End	Analyzed Fragment Size	Bases in Confident Regions	Amplicon Genomic Content	GC Content	Total # Correct Calls	Total # Incorrect Calls	Total # No Calls	% Correct Calls
28	5	1882081	1882158	78	75	N/A	0.78	1800	0	0	100
29	5	14769061	14769144	84	84	GT(3), CCA(3)	0.62	2016	0	0	100
30	5	41069808	41069871	64	64	N/A	0.39	1536	0	0	100
31	5	74077114	74077196	83	83	PolyA (6), indel	0.3	1992	0	0	100
32	5	147475343	147475409	67	67	PolyT (5)	0.37	1608	0	0	100
33	5	149323731	149323821	91	91	CT(4), AG(3)	0.55	2184	0	0	100
34	5	155662213	155662287	75	75	Indel	0.43	1800	0	0	100
35	6	6318713	6318814	10	10	PolyG (6)	0.68	2448	0	0	100
36	6	24949983	24950074	92	92	Indel	0.63	2208	0	0	100
37	6	31084900	31084999	100	94	GCT(5), indel	0.61	2244	0	12	99.5
38	6	32147987	32148084	98	98	PolyT (5), TCT(3), CTT (3)	0.55	2352	0	0	100
39	6	32986864	32986958	95	95	Indel	0.53	2280	0	0	100
40	6	33408498	33408583	86	86	PolyC (6)	0.7	2064	0	0	100
41	6	41647401	41647495	95	94	PolyG (5), indel	0.61	2256	0	0	100
42	6	112435865	112435955	91	91	PolyA (5)	0.44	2184	0	0	100
43	7	22202076	22202148	73	73	N/A	0.44	1752	0	0	100
44	7	66276100	66276187	88	88	Indel	0.35	2112	0	0	100
45	7	77365735	77365821	87	87	PolyA (7), AG(4)	0.26	2088	0	0	100
46	7	110939946	110940030	85	85	Indel	0.38	2040	0	0	100
47	7	128533468	128533557	90	90	PolyG (5), indel	0.62	2160	0	0	100
48	7	149503875	149503965	91	91	PolyG (6), PolyC (6), indel	0.71	2184	0	0	100
49	7	154404519	154404599	81	66	N/A	0.31	1584	0	0	100
50	7	156476507	156476599	93	93	Indel	0.35	2232	0	0	100
51	8	1817312	1817394	83	83	N/A	0.42	1992	0	0	100
52	8	24811020	24811109	90	89	PolyG (7), CTC(4), indel	0.61	2113	0	23	98.9
53	8	76518625	76518691	67	67	Indel	0.3	1608	0	0	100
54	9	103054909	103055006	98	98	PolyG (6)	0.67	2352	0	0	100
55	9	105586150	105586214	65	65	Indel	0.32	1560	0	0	100

Amplicon	Chromosome	Amplicon Start	Amplicon End	Analyzed Fragment Size	Bases in Confident Regions	Amplicon Genomic Content	GC Content	Total # Correct Calls	Total # Incorrect Calls	Total # No Calls	% Correct Calls
56	9	107620823	107620918	96	96	N/A	0.49	2304	0	0	100
57	9	123769149	123769231	83	83	AT(3)	0.37	1992	0	0	100
58	9	138995345	138995441	97	97	PolyC (6), indel	0.68	2328	0	0	100
59	10	5987120	5987198	79	78	PolyG (5), indel	0.47	1872	0	0	100
60	10	11784629	11784726	98	91	GC(3)	0.87	2184	0	0	100
61	10	27317777	27317855	79	79	PolyT (5)	0.3	1896	0	0	100
62	10	33018351	33018440	90	90	PolyA (5), PolyT (5)	0.2	2160	0	0	100
63	10	45084159	45084253	95	95	Indel	0.35	2280	0	0	100
64	10	55892599	55892687	89	88	AC(11), indel	0.42	2102	0	10	99.5
65	10	101611250	101611329	80	80	N/A	0.49	1920	0	0	100
66	10	118351373	118351453	81	81	N/A	0.51	1944	0	0	100
67	11	8159816	8159912	97	96	N/A	0.45	2304	0	0	100
68	11	30177648	30177717	70	70	Indel	0.46	1680	0	0	100
69	11	47470345	47470444	100	100	N/A	0.65	2400	0	0	100
70	11	59837679	59837740	62	62	Indel	0.37	1488	0	0	100
71	11	64418856	64418957	102	102	N/A	0.59	2448	0	0	100
72	11	93529612	93529684	73	73	PolyA (5)	0.4	1752	0	0	100
73	11	101347052	101347136	85	85	N/A	0.42	2040	0	0	100
74	11	102477336	102477426	91	91	PolyG (6)	0.55	2184	0	0	100
75	11	118406285	118406369	85	85	Indel	0.53	2040	0	0	100
76	11	120357801	120357885	85	85	PolyA (5), CA(3), indel	0.34	2040	0	0	100
77	11	125769313	125769397	85	85	GA(3)	0.52	2040	0	0	100
78	12	2834770	2834853	84	84	PolyC (5), indel	0.52	2016	0	0	100
79	12	26811004	26811096	93	93	PolyA (7), AC(4)	0.33	2232	0	0	100
80	12	30881766	30881846	81	81	N/A	0.49	1944	0	0	100
81	12	88474105	88474175	71	71	PolyA (6)	0.35	1704	0	0	100
82	12	120966872	120966966	95	95	PolyG (5)	0.68	2280	0	0	100
83	13	24167504	24167576	73	73	N/A	0.52	1752	0	0	100

Amplicon	Chromosome	Amplicon Start	Amplicon End	Analyzed Fragment Size	Bases in Confident Regions	Amplicon Genomic Content	GC Content	Total # Correct Calls	Total # Incorrect Calls	Total # No Calls	% Correct Calls
84	13	25816961	25817049	89	88	PolyA (5), PolyT (7), PolyA (7), indel	0.22	2112	0	0	100
85	13	44880112	44880200	89	89	Indel	0.49	2136	0	0	100
86	13	77665218	77665294	77	77	Indel	0.39	1848	0	0	100
87	14	31619327	31619393	67	67	GA(3),TA(3)	0.39	1608	0	0	100
88	14	39517884	39517966	83	83	N/A	0.25	1992	0	0	100
89	14	46958962	46959034	73	72	PolyT (5), indel	0.19	1727	0	1	99.9
90	14	58050030	58050110	81	81	Indel	0.38	1944	0	0	100
91	14	82390559	82390649	91	91	Indel	0.35	2184	0	0	100
92	14	92549544	92549609	66	66	PolyA (5)	0.41	1584	0	0	100
93	14	102808496	102808589	94	94	Indel	0.62	2256	0	0	100
94	15	43170751	43170848	98	96	PolyC (5)	0.45	2304	0	0	100
95	15	63446149	63446216	68	68	Indel	0.25	1632	0	0	100
96	15	77879807	77879901	95	93	PolyG (5), indel	0.68	2232	0	0	100
97	15	81625334	81625428	95	95	PolyT (6)	0.43	2280	0	0	100
98	15	85438263	85438334	72	71	Indel	0.65	1704	0	0	100
99	15	89817413	89817503	91	91	N/A	0.36	2184	0	0	100
100	15	89864274	89864343	70	70	Indel	0.56	1680	0	0	100
101	16	1894910	1894972	63	63	N/A	0.27	1512	0	0	100
102	16	28997904	28997998	95	95	PolyC (5)	0.67	2280	0	0	100
103	16	53682908	53682994	87	87	TA(3)	0.41	2088	0	0	100
104	16	57954406	57954509	104	104	PolyC (5)	0.67	2496	0	0	100
105	16	85706375	85706465	91	91	Poly T (5), indel	0.37	2184	0	0	100
106	17	3563920	3564008	89	89	GC(3)	0.64	2136	0	0	100
107	17	3594191	3594277	87	87	PolyC (5), indel	0.67	2088	0	0	100
108	17	3970090	3970180	91	91	Indel	0.46	2184	0	0	100
109	17	16084945	16085037	93	93	Indel	0.26	2232	0	0	100
110	17	33998759	33998849	91	89	PolyT (5)	0.54	2136	0	0	100
111	17	39589691	39589774	84	82	PolyA (13), indel (x2)	0.29	1944	7	17	98.8

Amplicon	Chromosome	Amplicon Start	Amplicon End	Analyzed Fragment Size	Bases in Confident Regions	Amplicon Genomic Content	GC Content	Total # Correct Calls	Total # Incorrect Calls	Total # No Calls	% Correct Calls
112	17	41244394	41244484	91	91	PolyA (5)	0.34	2184	0	0	100
113	17	45438866	45438957	92	92	PolyA (7), AT(3), AT(4), AT(4), indel	0.26	2208	0	0	100
114	17	61502432	61502510	79	79	Indel	0.41	1887	0	9	99.5
115	17	64023582	64023667	86	86	PolyT (7)	0.22	2064	0	0	100
116	17	72308237	72308320	84	84	GAG(3)	0.62	2016	0	0	100
117	18	2616456	2616522	67	67	GA(3)	0.31	1608	0	0	100
118	18	6980478	6980568	91	91	N/A	0.37	2184	0	0	100
119	18	9888026	9888094	69	69	PolyA (6), TG(3)	0.43	1656	0	0	100
120	18	38836999	38837073	75	75	PolyA (5), indel	0.37	1800	0	0	100
121	18	47405382	47405462	81	81	CTC(3), indel	0.47	1944	0	0	100
122	18	54815665	54815749	85	85	CT(3), indel	0.45	2040	0	0	100
123	18	59773996	59774060	65	65	N/A	0.48	1560	0	0	100
124	19	625143	625241	99	99	N/A	0.59	2376	0	0	100
125	19	18121418	18121491	74	74	N/A	0.68	1775	1	0	99.9
126	19	18186574	18186643	70	70	N/A	0.64	1680	0	0	100
127	20	746056	746149	94	94	N/A	0.61	2256	0	0	100
128	20	10633195	10633276	82	82	AC(3)	0.59	1968	0	0	100
129	20	17705633	17705708	76	76	CT(3)	0.58	1824	0	0	100
130	20	21766821	21766890	70	70	GT(3),TG(4), indel	0.46	1680	0	0	100
131	20	25278421	25278521	101	101	Indel	0.63	2424	0	0	100
132	20	50897302	50897368	67	67	Indel	0.36	1608	0	0	100
133	20	62331904	62331994	91	88	PolyG (6)	0.73	2112	0	0	100
134	20	62690860	62690946	87	87	Indel	0.57	2088	0	0	100
135	21	30300823	30300888	66	66	Indel	0.35	1584	0	0	100
136	21	33694176	33694273	98	98	PolyT (6), CA(3)	0.54	2352	0	0	100
137	21	36710706	36710792	87	87	GT(3), indel	0.39	2088	0	0	100
138	21	46644924	46644992	69	69	PolyA (6), AG(3), indel	0.32	1656	0	0	100
139	21	46705575	46705664	90	90	PolyT (5), PolyA (6)	0.5	2160	0	0	100

Amplicon	Chromosome	Amplicon Start	Amplicon End	Analyzed Fragment Size	Bases in Confident Regions	Amplicon Genomic Content	GC Content	Total # Correct Calls	Total # Incorrect Calls	Total # No Calls	% Correct Calls
140	22	25750774	25750873	100	100	Indel	0.63	2400	0	0	100
141	22	32439233	32439329	97	97	N/A	0.68	2328	0	0	100
142	22	37409844	37409940	97	97	Indel	0.46	2328	0	0	100
143	22	37637596	37637694	99	99	N/A	0.6	2376	0	0	100
144	22	47081347	47081438	92	92	Indel	0.66	2208	0	0	100
145	X	15870424	15870492	69	69	PolyT (5)	0.26	1656	0	0	100
146	X	135288543	135288611	69	69	PolyC (5)	0.62	1656	0	0	100
147	X	135290777	135290847	71	71	N/A	0.52	1704	0	0	100
148	Y	2655397	2655461	65	0	N/A	0.55	0	0	0	N/A
149	Y	2655519	2655609	91	0	N/A	0.48	0	0	0	N/A
150	Y	2655609	2655679	71	0	PolyA (5)	0.37	0	0	0	N/A

Variants that were no calls are summarized in Table 3. The particular filters that resulted in the no calls are listed in the table. The insertion on amplicon 111 was filtered for nine of 16 occurrences with the remaining seven occurrences called as reference and are therefore FNs.

Table 3 Summary of Variants No Calls

Amplicon #	Chr:Pos	Variant	Corresponding Amplicon Content	Filter	Missed Variants	Expected Variants	FN Calls
64	10:55892600	TAC > T	AC(11), 42% GC	R5x9[1]	10	10	0
111	17:39589692	C > CA	PolyA (13), 29% GC	R5x9	9	16	7

[1] R5x9: Repeat filter. A variant is filtered if all or part of the variant is present repeatedly in the reference genome adjacent to the variant position. At least nine repeats in the reference are required and only repeats with length up to 5 bp are considered.

The sequencing results for sample NA12878 were compared to a highly confident genotype for NA12878, established by the National Institutes of Standards and Technology (NIST) (v.2.19). Out of the 150 amplicons, 92 amplicons were fully contained within the highly

confident genomic regions, 41 amplicons had partial overlap, and 17 amplicons had no overlap in the NIST sequence. This resulted in 10,000 coordinates per replicate for comparison. Non-variant base calls were compared to human genome reference sequence build 19. The accuracy results are shown in Table 4.

Table 4 Agreement of the MiSeqDx Instrument Base Call Results for NA12878 Sample with NIST Database

Sample	# Amplicons	Mean Call Rate	Total # TP Variants	Total # FN Variants	Total # TN Calls	Total # FP Calls	PPA	NPA	OPA
NA12878	133	99.98	208	0	19380	0	100	100	100

The samples were further analyzed for calling small insertions and deletions (indels) (Table 5). In some cases, the indel was common among two or more samples as reflected in the Total # Sample Replicates with Indel column. Results for both replicates of the 12 valid samples are included in Table 5. There were a total of 71 indels ranging in size from 1–24 bp for insertions and 1–25 bp for deletions. 69 indels were each detected with a positive percent agreement of 100%. One deletion (amplicon 64; 2 bp deletion (chr10 55892600 TAC>T) had no correct calls because each of these variants was a no-call due to the R5x9 filter. Therefore, PPA, which excludes no calls, could not be calculated. Another indel, 1 bp insertion (chr17 39589692 C>CA on amplicon 111), also had no correct calls because nine variants were a no-call due to the R5x9 filter and seven were FN calls.

Table 5 Summary of Indel Detection with the MiSeqDx Instrument

Amplicon	Chromosome	Position	Analyzed Fragment Size	Amplicon Indel Type and Length	Indel	Total # Sample Replicates with Indel	# of No Calls	Total # incorrect Indel Calls	Total # Correct Indel Calls	PPA
1	1	36450544	93	25 bp deletion	GAAAATTTAATGAAACACATTGTCCT>G	2	0	0	2	100
2	1	109465165	79	3 bp deletion	ACTT>A	12	0	0	12	100
3	1	218353908	91	23 bp insertion	T>TTTTAATAGCAAAAAGAGGCTAGA	24	0	0	24	100
4	1	223906701	92	17 bp deletion	GACAGACTGTGAGGAAGA>G	10	0	0	10	100
6	1	236372081	70	5 bp insertion	C>CTTAAG	10	0	0	10	100
7	1	247812083	88	3 bp insertion	C>CATG	10	0	0	10	100
8	2	55862804	90	7 bp insertion	T>TTTGGTAA	14	0	0	14	100
9	2	87003972	80	6 bp deletion	TTATCTC>T	6	0	0	6	100
13	2	200796749	87	5 bp insertion	T>TTAAAA	24	0	0	24	100
14	2	212245090	91	12 bp insertion	C>CTGAAAATAGGAT	14	0	0	14	100
16	2	235016388	73	2 bp insertion	A>ATG	12	0	0	12	100

Amplicon	Chromosome	Position	Analyzed Fragment Size	Amplicon Indel Type and Length	Indel	Total # Sample Replicates with Indel	# of No Calls	Total # incorrect Indel Calls	Total # Correct Indel Calls	PPA
17	3	4466274	93	23 bp deletion	TAACTTAAAATTACAAAATAACCC>T	2	0	0	2	100
19	3	49851375	70	9 bp insertion	C>CCTGGCTCCT	4	0	0	4	100
21	3	190106071	75	1 bp deletion	AG>A	20	0	0	20	100
25	4	56236567	66	8 bp deletion	TAACCGAAA>T	12	0	0	12	100
27	4	164446785	62	11 bp insertion	T>TTATGGTATTGA	12	0	0	12	100
31	5	74077155	83	4 bp deletion	TAGTA>T	10	0	0	10	100
34	5	155662255	75	8 bp insertion	G>GCCTACTGA	20	0	0	20	100
36	6	24950035	92	21 bp deletion	CCCTGGGTGCTATAGCCCACCA>C	10	0	0	10	100
37	6	31084942	100	3 bp deletion	GCTT>G	14	0	0	14	100
39	6	32986905	95	25 bp deletion	CTTTCACTTTCCCGTCTCATGCAAAG>C	12	0	0	12	100
41	6	41647442	95	23 bp deletion	GGCATGAGGCTTGGTGACATGGCA>G	8	0	0	8	100
44	7	66276142	88	1 bp insertion	C>CT	16	0	0	16	100
46	7	110939983	85	4 bp deletion	CAAGT>C	12	0	0	12	100
47	7	128533514	90	1 bp insertion	T>TC	24	0	0	24	100
48	7	149503916	91	4 bp deletion	GGATA>G	8	0	0	8	100
50	7	156476548	93	11 bp deletion	GAATCTGCACTT>G	12	0	0	12	100
52	8	24811064	90	1 bp deletion	AG>A	24	0	0	24	100
53	8	76518677	67	4 bp insertion	T>TACTG	14	0	0	14	100
55	9	105586193	65	4 bp insertion	C>CAATT	2	0	0	2	100
58	9	138995370	97	21 bp deletion	TCTGGGGGGCAGCCCCTGAGGG>T	14	0	0	14	100
59	10	5987158	79	3 bp deletion	TAAC>T	10	0	0	10	100
63	10	45084202	95	16 bp deletion	AGCGTCTATAACCAAAT>A	12	0	0	12	100
64	10	55892600	89	2 bp deletion	TAC>T	10	10	0	0	N/A
68	11	30177690	70	2 bp insertion	C>CTG	10	0	0	10	100
70	11	59837721	62	8 bp insertion	T>TTATGAAAA	12	0	0	12	100
75	11	118406328	85	8 bp deletion	CAGTGTGGA>C	10	0	0	10	100
76	11	120357842	85	2 bp deletion	CTT>C	10	0	0	10	100
78	12	2834814	84	21 bp insertion	T>TTCTCAGTACGGTGAACCCCAG	24	0	0	24	100

Amplicon	Chromosome	Position	Analyzed Fragment Size	Amplicon Indel Type and Length	Indel	Total # Sample Replicates with Indel	# of No Calls	Total # incorrect Indel Calls	Total # Correct Indel Calls	PPA
84	13	25817002	89	19 bp insertion	C>CAAAATATAAAAAGCTCCCT	24	0	0	24	100
85	13	44880152	89	4 bp insertion	C>CCTGT	12	0	0	12	100
86	13	77665265	77	20 bp deletion	ATCTATTTTCTAATAGACGGC>A	14	0	0	14	100
89	14	46958967	73	22 bp deletion	TTTAAAATTTGAATGTGATAAAA>T	24	0	0	24	100
90	14	58050081	81	4 bp insertion	C>CTGAT	20	0	0	20	100
91	14	82390602	91	16 bp deletion	CTTGCTCTATAAACCGT>C	10	0	0	10	100
93	14	102808554	94	5 bp deletion	CGTGGA>C	10	0	0	10	100
95	15	63446199	68	6 bp deletion	CAAAATT>C	12	0	0	12	100
96	15	77879862	95	25 bp deletion	GCCCCTGAGCCAGCCTCCCGCTCTTA>G	14	0	0	14	100
98	15	85438311	72	3 bp insertion	C>CTTG	8	0	0	8	100
100	15	89864316	70	4 bp insertion	G>GCTAC	8	0	0	8	100
105	16	85706416	91	7 bp deletion	ATTATTTC>A	16	0	0	16	100
107	17	3594276	87	1 bp deletion	TG>T	2	0	0	2	100
108	17	3970133	91	18 bp insertion	A>ATCCTATTCTACTCTGAAT	10	0	0	10	100
109	17	16084985	93	4 bp insertion	A>AACAC	10	0	0	10	100
111	17	39589692	84	1 bp insertion	C>CA	16	9	7	0	0
112	17	39589739	84	24 bp insertion	T>TTCTGAAGGTCAAGTCTATCCCTGA	24	0	0	24	100
113	17	45438886	92	4 bp deletion	CAGTG>C	12	0	0	12	100
114	17	61502459	79	12 bp deletion	TTTGTATCTGCTG>T	20	0	0	20	100
120	18	38837054	75	22 bp insertion	T>TGTATCTTAGCAAAAGTTTCTCA	24	0	0	24	100
121	18	47405425	81	3 bp insertion	T>TGAG	20	0	0	20	100
122	18	54815706	85	2 bp deletion	ACT>A	20	0	0	20	100
130	20	21766863	70	15 bp deletion	TACTTGAGAACTGAGG>T	4	0	0	4	100
131	20	25278464	101	5 bp insertion	A>AGTGGG	20	0	0	20	100
132	20	50897361	67	11 bp insertion	G>GGAATGTCAGCC	24	0	0	24	100
134	20	62690925	87	16 bp deletion	TCCTGGCTGGCCTGTGG>T	10	0	0	10	100
135	21	30300873	66	11 bp insertion	G>GATAAAACTTTA	10	0	0	10	100
137	21	36710749	87	21 bp deletion	ACTCAAGATAACTCATGTTATC>A	16	0	0	16	100

Amplicon	Chromosome	Position	Analyzed Fragment Size	Amplicon Indel Type and Length	Indel	Total # Sample Replicates with Indel	# of No Calls	Total # incorrect Indel Calls	Total # Correct Indel Calls	PPA
138	21	46644985	69	5 bp deletion	GTTGTT>G	8	0	0	8	100
140	22	25750814	100	6 bp insertion	C>CAGGGCA	20	0	0	20	100
142	22	37409885	97	5 bp insertion	C>CTGTTT	2	0	0	2	100
144	22	47081407	92	10 bp deletion	GGGCACAGGCA>G	12	0	0	12	100

Reproducibility

Two studies were conducted to assess the reproducibility of the MiSeqDx instrument with cell lines (study 1 and 2) or leukocyte-depleted blood spiked with cell lines (study 2). Study 1 used multiple instruments. Study 2 had multiple sites.

Study 1

The reproducibility of the MiSeqDx instrument was determined using two instruments, two operators, and two reagent lots for a total of eight runs. The representative assay, samples, and reference method are the same as described for the accuracy study.

The results are presented on per amplicon basis for each instrument (Table 6) to demonstrate the reproducibility of calling across instruments. % correct calls included both incorrect and no calls (one or more filters are not met for variant calling). The instruments generated similar numbers of no calls and incorrect calls dependent on the particular amplicon.

Table 6 Study Instrument to Instrument Reproducibility Results for the MiSeqDx Instrument (Amplicon-level)

Amplicon	Chromosome	Amplicon Start	Amplicon End	Analyzed Fragment Size	Bases in Confident Regions	Amplicon Genomic Content	GC Content	MiSeqDx 1			MiSeqDx 2		
								Total Correct Calls	Total Incorrect Calls	Total # No Calls	Total Correct Calls	Total Incorrect Calls	Total # No Calls
1	1	36450499	36450591	93	93	Indel	0.22	8928	0	0	8928	0	0
2	1	109465122	109465200	79	79	PolyA (5), PolyC (5), indel	0.38	7584	0	0	7584	0	0
3	1	218353867	218353957	91	91	Indel	0.4	8736	0	0	8736	0	0
4	1	223906657	223906748	92	92	Indel	0.49	8832	0	0	8832	0	0
5	1	228526602	228526682	81	81	PolyG (5)	0.69	7776	0	0	7776	0	0

Amplicon	Chromosome	Amplicon Start	Amplicon End	Analyzed Fragment Size	Bases in Confident Regions	Amplicon Genomic Content	GC Content	MiSeqDx 1			MiSeqDx 2		
								Total Correct Calls	Total Incorrect Calls	Total # No Calls	Total Correct Calls	Total Incorrect Calls	Total # No Calls
6	1	236372039	236372108	70	70	PolyT (10), indel	0.39	6720	0	0	6720	0	0
7	1	247812041	247812128	88	88	PolyA (5), CT(3), TAA(3), indel	0.27	8448	0	0	8448	0	0
8	2	55862774	55862863	90	90	Indel	0.28	8640	0	0	8640	0	0
9	2	87003930	87004009	80	80	Indel	0.38	7680	0	0	7680	0	0
10	2	177016721	177016805	85	81	N/A	0.65	7775	1	0	7775	1	0
11	2	186625727	186625801	75	75	PolyA (8)	0.35	7200	0	0	7200	0	0
12	2	190323504	190323591	88	88	PolyT (5)	0.42	8448	0	0	8448	0	0
13	2	200796740	200796826	87	87	PolyT (5), indel	0.31	8352	0	0	8352	0	0
14	2	212245049	212245139	91	91	PolyT (5), PolyA (6), indel	0.3	8736	0	0	8736	0	0
15	2	228147052	228147144	93	93	N/A	0.43	8928	0	0	8928	0	0
16	2	235016350	235016422	73	73	PolyT (5), indel	0.42	7008	0	0	7008	0	0
17	3	4466229	4466321	93	93	AT(3), indel	0.27	8761	0	167	8760	0	168
18	3	46620561	46620643	83	83	N/A	0.43	7968	0	0	7968	0	0
19	3	49851331	49851400	70	70	CT(3), indel	0.49	6720	0	0	6720	0	0
20	3	189713161	189713248	88	88	PolyA (5), PolyT (5), PolyA (9), TG(3)	0.41	8448	0	0	8448	0	0
21	3	190106030	190106104	75	74	Indel	0.57	7096	0	8	7096	0	8
22	4	2233667	2233744	78	78	PolyA (6)	0.26	7488	0	0	7488	0	0

Amplicon	Chromosome	Amplicon Start	Amplicon End	Analyzed Fragment Size	Bases in Confident Regions	Amplicon Genomic Content	GC Content	MiSeqDx 1			MiSeqDx 2		
								Total Correct Calls	Total Incorrect Calls	Total # No Calls	Total Correct Calls	Total Incorrect Calls	Total # No Calls
23	4	7780541	7780637	97	97	PolyG (6), PolyT (5), PolyA (5)	0.42	9312	0	0	9312	0	0
24	4	15688604	15688681	78	78	N/A	0.29	7488	0	0	7488	0	0
25	4	56236521	56236586	66	62	PolyA (5), indel	0.36	5952	0	0	5952	0	0
26	4	102839244	102839314	71	69	PolyA (5)	0.46	6624	0	0	6624	0	0
27	4	164446743	164446804	62	62	PolyA (7), indel	0.27	5952	0	0	5952	0	0
28	5	1882081	1882158	78	75	N/A	0.78	7200	0	0	7200	0	0
29	5	14769061	14769144	84	84	GT(3), CCA(3)	0.62	8064	0	0	8064	0	0
30	5	41069808	41069871	64	64	N/A	0.39	6144	0	0	6144	0	0
31	5	74077114	74077196	83	83	PolyA (6), indel	0.3	7968	0	0	7968	0	0
32	5	147475343	147475409	67	67	PolyT (5)	0.37	6432	0	0	6432	0	0
33	5	149323731	149323821	91	91	CT(4), AG (3)	0.55	8736	0	0	8736	0	0
34	5	155662213	155662287	75	75	Indel	0.43	7200	0	0	7200	0	0
35	6	6318713	6318814	102	102	PolyG (6)	0.68	9792	0	0	9792	0	0
36	6	24949983	24950074	92	92	Indel	0.63	8832	0	0	8832	0	0
37	6	31084900	31084999	100	94	GCT(5), indel	0.61	8979	0	45	8979	0	45
38	6	32147987	32148084	98	98	PolyT (5), TCT(3), CTT(3)	0.55	9408	0	0	9408	0	0
39	6	32986864	32986958	95	95	Indel	0.53	9120	0	0	9120	0	0
40	6	33408498	33408583	86	86	PolyC (6)	0.7	8256	0	0	8256	0	0
41	6	41647401	41647495	95	94	PolyG (5), indel	0.61	9024	0	0	9024	0	0
42	6	112435865	112435955	91	91	PolyA (5)	0.44	8736	0	0	8736	0	0

Amplicon	Chromosome	Amplicon Start	Amplicon End	Analyzed Fragment Size	Bases in Confident Regions	Amplicon Genomic Content	GC Content	MiSeqDx 1			MiSeqDx 2		
								Total Correct Calls	Total Incorrect Calls	Total # No Calls	Total Correct Calls	Total Incorrect Calls	Total # No Calls
43	7	22202076	22202148	73	73	N/A	0.44	7008	0	0	7008	0	0
44	7	66276100	66276187	88	88	Indel	0.35	8448	0	0	8448	0	0
45	7	77365735	77365821	87	87	PolyA (7), AG(4)	0.26	8352	0	0	8352	0	0
46	7	110939946	110940030	85	85	Indel	0.38	8160	0	0	8160	0	0
47	7	128533468	128533557	90	90	PolyG (5), indel	0.62	8550	0	90	8550	0	90
48	7	149503875	149503965	91	91	PolyG (6), PolyC (6), indel	0.71	8736	0	0	8736	0	0
49	7	154404519	154404599	81	66	N/A	0.31	6336	0	0	6336	0	0
50	7	156476507	156476599	93	93	Indel	0.35	8928	0	0	8928	0	0
51	8	1817312	1817394	83	83	N/A	0.42	7968	0	0	7968	0	0
52	8	24811020	24811109	90	89	PolyG (7), CTC(4), indel	0.61	8452	0	92	8449	0	95
53	8	76518625	76518691	67	67	Indel	0.3	6432	0	0	6432	0	0
54	9	103054909	103055006	98	98	PolyG (6)	0.67	9408	0	0	9408	0	0
55	9	105586150	105586214	65	65	Indel	0.32	6240	0	0	6240	0	0
56	9	107620823	107620918	96	96	N/A	0.49	9216	0	0	9216	0	0
57	9	123769149	123769231	83	83	AT(3)	0.37	7968	0	0	7968	0	0
58	9	138995345	138995441	97	97	PolyC (6), indel	0.68	9312	0	0	9312	0	0
59	10	5987120	5987198	79	78	PolyG (5), indel	0.47	7488	0	0	7488	0	0
60	10	11784629	11784726	98	91	GC(3)	0.87	8644	1	91	8644	1	91
61	10	27317777	27317855	79	79	PolyT (5)	0.3	7584	0	0	7584	0	0
62	10	33018351	33018440	90	90	PolyA (5), PolyT (5)	0.2	8640	0	0	8640	0	0
63	10	45084159	45084253	95	95	Indel	0.35	9120	0	0	9120	0	0

Amplicon	Chromosome	Amplicon Start	Amplicon End	Analyzed Fragment Size	Bases in Confident Regions	Amplicon Genomic Content	GC Content	MiSeqDx 1			MiSeqDx 2		
								Total Correct Calls	Total Incorrect Calls	Total # No Calls	Total Correct Calls	Total Incorrect Calls	Total # No Calls
64	10	55892599	55892687	89	88	AC(11), indel	0.42	8408	0	40	8407	0	41
65	10	101611250	101611329	80	80	N/A	0.49	7680	0	0	7680	0	0
66	10	118351373	118351453	81	81	N/A	0.51	7776	0	0	7776	0	0
67	11	8159816	8159912	97	96	N/A	0.45	9216	0	0	9216	0	0
68	11	30177648	30177717	70	70	Indel	0.46	6720	0	0	6720	0	0
69	11	47470345	47470444	100	100	N/A	0.65	9600	0	0	9600	0	0
70	11	59837679	59837740	62	62	Indel	0.37	5952	0	0	5952	0	0
71	11	64418856	64418957	102	102	N/A	0.59	9792	0	0	9792	0	0
72	11	93529612	93529684	73	73	PolyA (5)	0.4	7008	0	0	7008	0	0
73	11	101347052	101347136	85	85	N/A	0.42	8160	0	0	8160	0	0
74	11	102477336	102477426	91	91	PolyG (6)	0.55	8736	0	0	8736	0	0
75	11	118406285	118406369	85	85	Indel	0.53	8160	0	0	8160	0	0
76	11	120357801	120357885	85	85	PolyA (5), CA(3), indel	0.34	8160	0	0	8160	0	0
77	11	125769313	125769397	85	85	GA(3)	0.52	8160	0	0	8160	0	0
78	12	2834770	2834853	84	84	PolyC (5), indel	0.52	8064	0	0	8064	0	0
79	12	26811004	26811096	93	93	PolyA (7), AC(4)	0.33	8928	0	0	8928	0	0
80	12	30881766	30881846	81	81	N/A	0.49	7776	0	0	7776	0	0
81	12	88474105	88474175	71	71	PolyA (6)	0.35	6816	0	0	6816	0	0
82	12	120966872	120966966	95	95	PolyG (5)	0.68	9117	3	0	9119	1	0
83	13	24167504	24167576	73	73	N/A	0.52	7008	0	0	7008	0	0
84	13	25816961	25817049	89	88	PolyA (5), PolyT (7), PolyA (7), indel	0.22	8448	0	0	8448	0	0
85	13	44880112	44880200	89	89	Indel	0.49	8544	0	0	8544	0	0

Amplicon	Chromosome	Amplicon Start	Amplicon End	Analyzed Fragment Size	Bases in Confident Regions	Amplicon Genomic Content	GC Content	MiSeqDx 1			MiSeqDx 2		
								Total Correct Calls	Total Incorrect Calls	Total # No Calls	Total Correct Calls	Total Incorrect Calls	Total # No Calls
86	13	77665218	77665294	77	77	Indel	0.39	7392	0	0	7392	0	0
87	14	31619327	31619393	67	67	GA(3),TA (3)	0.39	6432	0	0	6432	0	0
88	14	39517884	39517966	83	83	N/A	0.25	7968	0	0	7968	0	0
89	14	46958962	46959034	73	72	PolyT (5), indel	0.19	6830	0	82	6835	0	77
90	14	58050030	58050110	81	81	Indel	0.38	7776	0	0	7776	0	0
91	14	82390559	82390649	91	91	Indel	0.35	8736	0	0	8736	0	0
92	14	92549544	92549609	66	66	PolyA (5)	0.41	6336	0	0	6336	0	0
93	14	102808496	102808589	94	94	Indel	0.62	9024	0	0	9024	0	0
94	15	43170751	43170848	98	96	PolyC (5)	0.45	9216	0	0	9216	0	0
95	15	63446149	63446216	68	68	Indel	0.25	6528	0	0	6528	0	0
96	15	77879807	77879901	95	93	PolyG (5), indel	0.68	8928	0	0	8926	2	0
97	15	81625334	81625428	95	95	PolyT (6)	0.43	9120	0	0	9120	0	0
98	15	85438263	85438334	72	71	Indel	0.65	6816	0	0	6816	0	0
99	15	89817413	89817503	91	91	N/A	0.36	8736	0	0	8736	0	0
100	15	89864274	89864343	70	70	Indel	0.56	6720	0	0	6720	0	0
101	16	1894910	1894972	63	63	N/A	0.27	6048	0	0	6048	0	0
102	16	28997904	28997998	95	95	PolyC (5)	0.67	9120	0	0	9120	0	0
103	16	53682908	53682994	87	87	TA(3)	0.41	8352	0	0	8352	0	0
104	16	57954406	57954509	104	104	PolyC (5)	0.67	9984	0	0	9984	0	0
105	16	85706375	85706465	91	91	PolyT (5), indel	0.37	8736	0	0	8736	0	0
106	17	3563920	3564008	89	89	GC(3)	0.64	8544	0	0	8544	0	0
107	17	3594191	3594277	87	87	PolyC (5), indel	0.67	8347	0	5	8347	0	5
108	17	3970090	3970180	91	91	Indel	0.46	8736	0	0	8736	0	0
109	17	16084945	16085037	93	93	Indel	0.26	8928	0	0	8928	0	0

Amplicon	Chromosome	Amplicon Start	Amplicon End	Analyzed Fragment Size	Bases in Confident Regions	Amplicon Genomic Content	GC Content	MiSeqDx 1			MiSeqDx 2		
								Total Correct Calls	Total Incorrect Calls	Total # No Calls	Total Correct Calls	Total Incorrect Calls	Total # No Calls
110	17	33998759	33998849	91	89	PolyT (5)	0.54	8544	0	0	8544	0	0
111	17	39589691	39589774	84	82	PolyA (13), indel (x2)	0.29	7776	7	89	7777	12	83
112	17	41244394	41244484	91	91	PolyA (5)	0.34	8736	0	0	8736	0	0
113	17	45438866	45438957	92	92	PolyA (7), AT(3), AT (4), AT(4), indel	0.26	8832	0	0	8832	0	0
114	17	61502432	61502510	79	79	Indel	0.41	7546	0	38	7547	0	37
115	17	64023582	64023667	86	86	PolyT (7)	0.22	8256	0	0	8256	0	0
116	17	72308237	72308320	84	84	GAG(3)	0.62	8064	0	0	8064	0	0
117	18	2616456	2616522	67	67	GA(3)	0.31	6432	0	0	6432	0	0
118	18	6980478	6980568	91	91	N/A	0.37	8736	0	0	8736	0	0
119	18	9888026	9888094	69	69	PolyA (6), TG(3)	0.43	6624	0	0	6624	0	0
120	18	38836999	38837073	75	75	PolyA (5), indel	0.37	7200	0	0	7200	0	0
121	18	47405382	47405462	81	81	CTC(3), indel	0.47	7776	0	0	7776	0	0
122	18	54815665	54815749	85	85	CT(3), indel	0.45	8160	0	0	8160	0	0
123	18	59773996	59774060	65	65	N/A	0.48	6240	0	0	6240	0	0
124	19	625143	625241	99	99	N/A	0.59	9504	0	0	9504	0	0
125	19	18121418	18121491	74	74	N/A	0.68	7102	2	0	7104	0	0
126	19	18186574	18186643	70	70	N/A	0.64	6718	2	0	6718	2	0
127	20	746056	746149	94	94	N/A	0.61	9024	0	0	9024	0	0
128	20	10633195	10633276	82	82	AC(3)	0.59	7872	0	0	7872	0	0
129	20	17705633	17705708	76	76	CT(3)	0.58	7296	0	0	7296	0	0
130	20	21766821	21766890	70	70	GT(3),TG (4), indel	0.46	6720	0	0	6720	0	0
131	20	25278421	25278521	101	101	Indel	0.63	9696	0	0	9696	0	0

Amplicon	Chromosome	Amplicon Start	Amplicon End	Analyzed Fragment Size	Bases in Confident Regions	Amplicon Genomic Content	GC Content	MiSeqDx 1			MiSeqDx 2		
								Total Correct Calls	Total Incorrect Calls	Total # No Calls	Total Correct Calls	Total Incorrect Calls	Total # No Calls
132	20	50897302	50897368	67	67	Indel	0.36	6432	0	0	6432	0	0
133	20	62331904	62331994	91	88	PolyG (6)	0.73	8360	0	88	8360	0	88
134	20	62690860	62690946	87	87	Indel	0.57	8352	0	0	8352	0	0
135	21	30300823	30300888	66	66	Indel	0.35	6336	0	0	6336	0	0
136	21	33694176	33694273	98	98	PolyT (6), CA(3)	0.54	9408	0	0	9408	0	0
137	21	36710706	36710792	87	87	GT(3), indel	0.39	8352	0	0	8352	0	0
138	21	46644924	46644992	69	69	PolyA (6), AG(3), indel	0.32	6603	0	21	6601	0	23
139	21	46705575	46705664	90	90	PolyT (5), PolyA (6)	0.5	8640	0	0	8640	0	0
140	22	25750774	25750873	100	100	Indel	0.63	9600	0	0	9600	0	0
141	22	32439233	32439329	97	97	N/A	0.68	9312	0	0	9312	0	0
142	22	37409844	37409940	97	97	Indel	0.46	9312	0	0	9312	0	0
143	22	37637596	37637694	99	99	N/A	0.6	9504	0	0	9504	0	0
144	22	47081347	47081438	92	92	Indel	0.66	8832	0	0	8832	0	0
145	X	15870424	15870492	69	69	PolyT (5)	0.26	6624	0	0	6624	0	0
146	X	135288543	135288611	69	69	PolyC (5)	0.62	6624	0	0	6624	0	0
147	X	135290777	135290847	71	71	N/A	0.52	6816	0	0	6816	0	0
148	Y	2655397	2655461	65	0	N/A	0.55	0	0	0	0	0	0
149	Y	2655519	2655609	91	0	N/A	0.48	0	0	0	0	0	0
150	Y	2655609	2655679	71	0	PolyA (5)	0.37	0	0	0	0	0	0

Reproducibility study results were analyzed on a per operator basis using variant frequency (Table 7). This analysis demonstrated that variant frequencies were consistent across the operators. Mean variant frequencies +/- 1 standard deviation are presented.

Table 7 Operator to Operator Results for the MiSeqDx Instrument

Variant Frequency Range	# Unique Variants	Total # Variants Analyzed Operator 1	Total # Variants Analyzed Operator 2	Mean (SD) Reported Variant Frequency Operator 1	Mean (SD) Reported Variant Frequency Operator 2
Homozygous (0.70-1.00)	2424	2424	2422	0.94 +/- 0.07	0.96 +/- 0.05
Heterozygous (0.20-0.70)	8240	8132	8128	0.48 +/- 0.04	0.49 +/- 0.04

Reproducibility study results for each sample are shown compounded from all eight runs (Table 8). Detection is evaluated for each variant type – SNVs, insertions, and deletions – separately. Reference positions are excluded. This analysis demonstrated that the results for the variants were reproducible across the samples.

Table 8 Agreement of MiSeqDx Instrument Base Call Results Per Sample

	SNVs				Insertions				Deletions			
Sample	Total #	Total # TP	Total # FP	Total # FN	Total #	Total # TP	Total # FP	Total # FN	Total #	Total # TP	Total # FP	Total # FN
NA12877	592	592	0	0	336	336	0	0	288	288	0	0
NA12878	1456	1456	0	0	320	304	0	0	384	368	0	0
NA12879	912	912	0	0	336	320	0	2	288	288	0	0
NA12880	1072	1071	0	1	384	384	0	0	320	304	0	0
NA12881	1248	1247	0	1	384	368	0	0	368	368	0	0
NA12882	944	943	0	1	352	336	0	4	304	288	0	0
NA12883	1088	1087	0	1	368	368	0	0	352	335	0	1
NA12884	1088	1088	0	0	400	384	0	5	336	336	0	0
NA12885	1200	1189	0	7	400	382	0	4	352	336	0	0
NA12886	1104	1102	0	2	368	352	0	3	368	368	0	0
NA12888	1056	1054	0	2	368	368	0	0	304	304	0	0
NA12893	1168	1168	0	0	352	336	0	1	368	368	0	0

The data provided by the eight runs in this reproducibility study support the claim that the MiSeqDxinstrument can consistently sequence:

- GC content ≥ 19% (all called bases in 192 out of 192 sequenced amplicons with 19% GC content called correctly with no-call rate of 1.1%)
- GC content ≤ 78% (all called bases in 192 out of 192 sequenced amplicons with 78% GC content called correctly with zero no-calls)
- PolyA lengths ≤ 8 (PolyA repeat of 8 nucleotides was called correctly in 192 out of 192 sequenced amplicons containing PolyA = 8)
- PolyT lengths ≤ 10 (PolyT repeat of 10 nucleotides was called correctly in 192 out of 192 sequenced amplicons containing PolyT = 10)
- PolyG lengths ≤ 7 (PolyG repeat of 7 nucleotides was called correctly in 192 out of 192 sequenced amplicons containing PolyG = 7)
- PolyC lengths ≤ 6 (PolyC repeat of 6 nucleotides was called correctly in 576 out of 576 sequenced amplicons containing PolyC = 6)
- Dinucleotide repeat lengths ≤ 11x (all called bases in 192 out of 192 sequenced amplicons with 11x dinucleotide repeat were called correctly with no-call rate of 0.5%)
- Trinucleotide repeat lengths ≤ 5x (all called bases in 192 out of 192 sequenced amplicons with 5x trinucleotide repeat were called correctly with no-call rate of 0.5%)
- 24 or fewer base insertions and 25 or fewer base deletions
  - 24-base insertions called correctly in 192 out of 192 samples
  - 25-base deletions called correctly in 223 samples and miscalled in 1 sample out of 224 samples

Study 2

A site-to-site reproducibility study performed with a representative assay, the Illumina MiSeqDx Cystic Fibrosis 139 Variant Assay, included a subset of *CFTR* clinically significant genetic variations analyzed with the MiSeq Reporter software using the MiSeqDx Platform targeted DNA sequencing workflow. The blinded study used 3 trial sites and 2 operators at each site. Two well-characterized panels of 46 samples each were tested by each of the operators at each site for a total of 810 calls per site. The panels contained a mix of genomic DNA from cell lines with known variants in the *CFTR* gene, as well as leukocyte-depleted blood spiked with cell lines with known variants in the *CFTR* gene. The blood samples were provided to allow incorporation of the extraction steps used to prepare gDNA that serves as the primary input for the assay workflow. The sample pass rate, defined as the number of samples passing QC metrics on the first attempt, was 99.88%. All test results are based on initial testing.

Table 9 Summary of Reproducibility Study Results Performed with a Representative MiSeqDx Cystic Fibrosis 139-Variant Assay

Panel	Sample #	Sample Genotype	Variants	Total Calls per Site	Positive Agreeing Calls (Variants)			Negative Agreeing Calls (Wild Type)			# Miscalls	# No Calls	Positive Agreement (%)	Negative Agreement (%)	Overall Agreement (%)
					Site 1	Site 2	Site 3	Site 1	Site 2	Site 3					
A	1	S549N (HET)		810	6	6	6	804	804	804	0	0	100	100	100
A	2	1812-1 G>A (HET)		810	6	6	6	804	804	804	0	0	100	100	100
A	3	Q493X/F508del (HET)		810	12	12	12	798	798	798	0	0	100	100	100
A	4[1]	F508del/2184delA (HET)		810	12	12	12	797	798	798	0	1[1]	100	100	100
A	5[2]	Y122X/R1158X (HET)		810	12	10	12	798	665	798	0	135[2]	94.44	94.44	94.44
A	6	F508del/2183AA>G (HET)		810	12	12	12	798	798	798	0	0	100	100	100
A	7	R75X (HET)		810	6	6	6	804	804	804	0	0	100	100	100
A	8	I507del/F508del (HET)		810	12	12	12	798	798	798	0	0	100	100	100
A	9[3]	F508del/W1282X (HET)		810	12	11	12	798	797	798	2[3]	0	97.22	99.96	99.92
A	10[3]	F508del/3272-26A>G (HET)		810	12	11	12	798	797	798	2[3]	0	97.22	99.96	99.92
A	11	F508del/3849+10kbC>T (HET)		810	12	12	12	798	798	798	0	0	100	100	100
A	12	621+1G>T/3120+1G>A (HET)		810	12	12	12	798	798	798	0	0	100	100	100
A	13	E60X/F508del (HET)		810	12	12	12	798	798	798	0	0	100	100	100
A	14	M1101K (HET)		810	6	6	6	804	804	804	0	0	100	100	100
A	15	M1101K (HOM)		810	6	6	6	804	804	804	0	0	100	100	100
A	16	F508del (HOM)	I506V, I507V, F508C not present	828	6	6	6	822	822	822	0	0	100	100	100
A	17	F508del/3659delC (HET)		810	12	12	12	798	798	798	0	0	100	100	100
A	18	R117H/F508del (HET)	(TG)10(T)9/(TG)12 (T)5	816	18	18	18	798	798	798	0	0	100	100	100

Panel	Sample #	Sample Genotype	Variants	Total Calls per Site	Positive Agreeing Calls (Variants)			Negative Agreeing Calls (Wild Type)			# Miscalls	# No Calls	Positive Agreement (%)	Negative Agreement (%)	Overall Agreement (%)
					Site 1	Site 2	Site 3	Site 1	Site 2	Site 3					
A	19	621+1G>T/711+1G>T (HET)		810	12	12	12	798	798	798	0	0	100	100	100
A	20	G85E/621+1G>T (HET)		810	12	12	12	798	798	798	0	0	100	100	100
A	21	A455E/F508del (HET)		810	12	12	12	798	798	798	0	0	100	100	100
A	22	F508del/R560T (HET)		810	12	12	12	798	798	798	0	0	100	100	100
A	23	F508del/Y1092X (C>A) (HET)		810	12	12	12	798	798	798	0	0	100	100	100
A	24	N1303K (HET)		810	6	6	6	804	804	804	0	0	100	100	100
A	25	G542X (HOM)		810	6	6	6	804	804	804	0	0	100	100	100
A	26	G542X (HET)		810	6	6	6	804	804	804	0	0	100	100	100
A	27	G551D/R553X (HET)		810	12	12	12	798	798	798	0	0	100	100	100
A	28	3849+10kbC>T (HOM)		810	6	6	6	804	804	804	0	0	100	100	100
A	29	WT		810	0	0	0	810	810	810	0	0	N/A	100	100
A	30	F508del (HET)		810	6	6	6	804	804	804	0	0	100	100	100
A	31	1717-1G>A (HET)		810	6	6	6	804	804	804	0	0	100	100	100
A	32	R1162X (HET)		810	6	6	6	804	804	804	0	0	100	100	100
A	33	R347P/G551D (HET)		810	12	12	12	798	798	798	0	0	100	100	100
A	34	R334W (HET)		810	6	6	6	804	804	804	0	0	100	100	100
A	35	WT		810	0	0	0	810	810	810	0	0	N/A	100	100
A	36	G85E (HET)		810	6	6	6	804	804	804	0	0	100	100	100
A	37	I336K (HET)		810	6	6	6	804	804	804	0	0	100	100	100
A	38	WT		810	0	0	0	810	810	810	0	0	N/A	100	100

Panel	Sample #	Sample Genotype	Variants	Total Calls per Site	Positive Agreeing Calls (Variants)			Negative Agreeing Calls (Wild Type)			# Miscalls	# No Calls	Positive Agreement (%)	Negative Agreement (%)	Overall Agreement (%)
					Site 1	Site 2	Site 3	Site 1	Site 2	Site 3					
A	39	F508del/3849+10kbC>T (HET)		810	12	12	12	798	798	798	0	0	100	100	100
A	40	621+1G>T/3120+1G>A (HET)		810	12	12	12	798	798	798	0	0	100	100	100
A	41	F508del/3659delC (HET)		810	12	12	12	798	798	798	0	0	100	100	100
A	42	R117H/F508del (HET)	(TG)10(T)9/(TG)12(T)5	816	18	18	18	798	798	798	0	0	100	100	100
A	43	G85E/621+1G>T (HET)		810	12	12	12	798	798	798	0	0	100	100	100
A	44	A455E/F508del (HET)		810	12	12	12	798	798	798	0	0	100	100	100
A	45	N1303K (HET)		810	6	6	6	804	804	804	0	0	100	100	100
A	46	G551D/R553X (HET)		810	12	12	12	798	798	798	0	0	100	100	100
B	47	2789+5G>A (HOM)		810	6	6	6	804	804	804	0	0	100	100	100
B	48	CFTR dele2, 3/F508del (HET)		810	12	12	12	798	798	798	0	0	100	100	100
B	49	F508del/1898+1G>A (HET)		810	12	12	12	798	798	798	0	0	100	100	100
B	50	WT		810	0	0	0	810	810	810	0	0	N/A	100	100
B	51	F508del/2143delT (HET)		810	12	12	12	798	798	798	0	0	100	100	100
B	52	3876delA (HET)		810	6	6	6	804	804	804	0	0	100	100	100
B	53	3905insT (HET)		810	6	6	6	804	804	804	0	0	100	100	100
B	54	394delTT (HET)		810	6	6	6	804	804	804	0	0	100	100	100
B	55	F508del (HET)		810	6	6	6	804	804	804	0	0	100	100	100
B	56	WT		810	0	0	0	810	810	810	0	0	N/A	100	100
B	57	WT		810	0	0	0	810	810	810	0	0	N/A	100	100

Panel	Sample #	Sample Genotype	Variants	Total Calls per Site	Positive Agreeing Calls (Variants)			Negative Agreeing Calls (Wild Type)			# Miscalls	# No Calls	Positive Agreement (%)	Negative Agreement (%)	Overall Agreement (%)
					Site 1	Site 2	Site 3	Site 1	Site 2	Site 3					
B	58	F508del (HET)		810	6	6	6	804	804	804	0	0	100	100	100
B	59	WT		810	0	0	0	810	810	810	0	0	N/A	100	100
B	60	L206W (HET)		810	6	6	6	804	804	804	0	0	100	100	100
B	61	WT		810	0	0	0	810	810	810	0	0	N/A	100	100
B	62	G330X (HET)		810	6	6	6	804	804	804	0	0	100	100	100
B	63	WT		810	0	0	0	810	810	810	0	0	N/A	100	100
B	64	R347H (HET)		810	6	6	6	804	804	804	0	0	100	100	100
B	65	1078delT (HET)		810	6	6	6	804	804	804	0	0	100	100	100
B	66	G178R/F508del (HET)		810	12	12	12	798	798	798	0	0	100	100	100
B	67	S549R (c.1647T>G) (HET)		810	6	6	6	804	804	804	0	0	100	100	100
B	68	S549N (HET)		810	6	6	6	804	804	804	0	0	100	100	100
B	69	W846X (HET)		810	6	6	6	804	804	804	0	0	100	100	100
B	70	WT		810	0	0	0	810	810	810	0	0	N/A	100	100
B	71	E92X/F508del (HET)		810	12	12	12	798	798	798	0	0	100	100	100
B	72[4]	621+1G>T/1154insTC (HET)		810	12	12	12	798	798	797	0	1[4]	100	99.96	99.96
B	73	G542X (HET)		810	6	6	6	804	804	804	0	0	100	100	100
B	74	F508del (HET)		810	6	6	6	804	804	804	0	0	100	100	100
B	75[2]	F508del (HET)		810	6	5	6	804	670	804	0	135[2]	94.44	94.44	94.44
B	76	F508del (HET)		810	6	6	6	804	804	804	0	0	100	100	100
B	77	621+1G>T/A455E (HET)		810	12	12	12	798	798	798	0	0	100	100	100

Panel	Sample #	Sample Genotype	Variants	Total Calls per Site	Positive Agreeing Calls (Variants)			Negative Agreeing Calls (Wild Type)			# Miscalls	# No Calls	Positive Agreement (%)	Negative Agreement (%)	Overall Agreement (%)
					Site 1	Site 2	Site 3	Site 1	Site 2	Site 3					
B	78	1812-1 G>A (HET)		810	6	6	6	804	804	804	0	0	100	100	100
B	79	WT		810	0	0	0	810	810	810	0	0	N/A	100	100
B	80	F508del/R553X (HET)		810	12	12	12	798	798	798	0	0	100	100	100
B	81	F508del/G551D (HET)		810	12	12	12	798	798	798	0	0	100	100	100
B	82	R347P/F508del (HET)		810	12	12	12	798	798	798	0	0	100	100	100
B	83	R117H/F508del (HET)	(TG)10(T)9/(TG)12 (T)5	816	18	18	18	798	798	798	0	0	100	100	100
B	84	I507del (HET)		810	6	6	6	804	804	804	0	0	100	100	100
B	85	2789+5G>A (HOM)		810	6	6	6	804	804	804	0	0	100	100	100
B	86[4]	CFTR dele2, 3/F508del (HET)		810	12	12	12	798	797	798	0	1[4]	100	99.96	99.96
B	87	F508del/1898+1G>A (HET)		810	12	12	12	798	798	798	0	0	100	100	100
B	88	WT		810	0	0	0	810	810	810	0	0	N/A	100	100
B	89	F508del/2143delT (HET)		810	12	12	12	798	798	798	0	0	100	100	100
B	90	3905insT (HET)		810	6	6	6	804	804	804	0	0	100	100	100
B	91	394delTT (HET)		810	6	6	6	804	804	804	0	0	100	100	100
B	92	F508del (HET)		810	6	6	6	804	804	804	0	0	100	100	100
		Total		74556	2209			221182			4	273	99.77	99.88	99.88

[1] The wild type location corresponding to the N1303K variant for one replicate resulted in a No Call due to insufficient coverage.
[2] One replicate of samples 5 and 75 had a 0% call rate. Further investigation indicates that samples may not have been added to the sample plate prior to library preparation because the sample volumes remaining in the tubes were consistent with no volume having been removed.
[3] Evidence indicates that samples 9 and 10 were likely switched by the operator prior to library preparation.
[4] The wild type location corresponding to the M1V variant for one replicate of each of two samples resulted in a No Call due to insufficient coverage.

Somatic Performance Characteristics

Studies described here used the Somatic Variant Module to analyze sequencing data, except those studies using the two gene panel where an assay-specific module was used.

Accuracy

Three studies were conducted to assess the accuracy of the MiSeqDx instrument with DNA extracted from FFPE samples.

Study 1

The study used a representative assay designed to query a variety of genes covering 12,588 bases across 23 different chromosomes using 150 amplicons. The Y chromosome does not contain confident regions and was not evaluated. The five unique samples used in this study are from a single family – two parents and three children – frequently sequenced by multiple laboratories and sequencing methodologies. There are three samples from females and two from males. All the samples were formalin fixed and paraffin embedded before DNA was extracted for the study. Sample GM12877 was diluted, at the DNA level, with sample GM12878 to create GM12877-D to make a set of variants with frequencies near 5% and 10%. Each of the samples was tested in duplicate except GM12877-D, which was tested with five replicates. Accuracy was determined for SNVs, insertions, and deletions by comparing the study data to a well-characterized reference database. The reference database sequence (Platinum Genomes version 2016-01) was derived from the combination of multiple sequencing methodologies, publicly available data, and hereditary information. Confident genomic regions were defined based on this reference method unless otherwise specified. In total the samples were run eight times. The tables presented to demonstrate accuracy are based on data from the first run.

Table 10 contains the study data presented with positive and negative percent agreement on a per sample basis, where the variant results are compared to the well-characterized composite reference method for PPA calculations. The three variant types (SNVs, insertions, and deletions) are combined. Because the reference method only provides results for the single nucleotide variants and insertions/deletions, non-variant base results are compared to human genome reference sequence build hg19 for NPA calculations.

Table 10 Agreement of the MiSeqDx Instrument Base Call Results with Reference Data for 6 Well-characterized Samples

Sample	Mean Call Rate	Total # Variants	Total # TP Variants	Total # FN Variants	Total # TN Calls	PPA	NPA	OPA
GM12877	98.7	152	147	0	23719	100	100	100
GM12878	98.4	270	260	0	23482	100	100	100
GM12879	98.7	192	186	0	23744	100	100	100

Sample	Mean Call Rate	Total # Variants	Total # TP Variants	Total # FN Variants	Total # TN Calls	PPA	NPA	OPA
GM12885	99.1	244	236	0	23713	100	100	100
GM12886	98.7	230	226	0	23652	100	100	100
GM12877-D[1]	98.4	675	650	0	57608	100	100	100
GM12877-D[2]		155	155	0		100	100	100

[1] Variants with frequency greater than 20%.
[2] Variant with frequency less than 20%.

The 150 amplicons were designed to cover a variety of genomic content. The GC content of the amplicons ranged from 26-87%. Amplicons also had a range of single nucleotide (e.g. PolyA, PolyT), dinucleotide, and trinucleotide repeats. 6 samples unique samples were used in the assay. Data were compiled on a per amplicon basis (Table 11) to determine the effect of genomic content on % correct calls. % correct calls consists of variant and reference calls and is less than 100% if there are either incorrect or no calls. No calls occur when one or more filters are not met for variant calling (eg, insufficient coverage). There were no incorrect calls. The number of no calls varied considerably across the amplicons. GC content and several interactions with GC content were the most significant predictors of no calls. 2040/2580 (79%) of no calls were due to not meeting the coverage specification. Amplicons with GC content greater than 78% resulted in the most no calls. A representative amplicon with 78% GC content had a total of 675 no calls. A representative amplicon with 87% had a total of 1365 no calls. Coverage can be increased by reducing the number of samples loaded on the flow cell, which can allow detection on amplicons with high GC content.

Table 11 Amplicon-level Accuracy Data

Amplicon	Chromosome	Amplicon Start	Amplicon End	Analyzed Fragment Size	Bases in Confident Regions	Amplicon Genomic Content	GC Content	Total # Correct Calls	Total # Incorrect Calls	Total # No Calls	% Correct Calls
1	1	36450499	36450591	93	93	Indel	0.22	1395	0	0	100
2	1	109465122	109465200	79	79	PolyA (5), PolyC (5), indel	0.38	1185	0	0	100
3	1	218353867	218353957	91	91	Indel	0.4	1364	0	1	99.9
4	1	223906657	223906748	92	92	Indel	0.49	1380	0	0	100
5	1	228526602	228526682	81	81	PolyG (5)	0.69	1215	0	0	100
6	1	236372039	236372108	70	70	PolyT (10), indel	0.39	1050	0	0	100
7	1	247812041	247812128	88	88	PolyA (5), CT(3), TAA(3), indel	0.27	1320	0	0	100
8	2	55862774	55862863	90	90	Indel	0.28	1350	0	0	100

Amplicon	Chromosome	Amplicon Start	Amplicon End	Analyzed Fragment Size	Bases in Confident Regions	Amplicon Genomic Content	GC Content	Total # Correct Calls	Total # Incorrect Calls	Total # No Calls	% Correct Calls
9	2	87003930	87004009	80	80	Indel	0.38	1200	0	0	100
10	2	177016721	177016805	85	81	N/A	0.65	1215	0	0	100
11	2	186625727	186625801	75	75	PolyA (8)	0.35	1117	0	10	99.1
12	2	190323504	190323591	88	88	PolyT (5)	0.42	1320	0	0	100
13	2	200796740	200796826	87	87	PolyT (5), indel	0.31	1302	0	8	99.4
14	2	212245049	212245139	91	91	PolyT (5), PolyA (6), indel	0.3	1365	0	0	100
15	2	228147052	228147144	93	93	N/A	0.43	1395	0	0	100
16	2	235016350	235016422	73	73	PolyT (5), indel	0.42	1095	0	0	100
17	3	4466229	4466321	93	93	AT(3), indel	0.27	1349	0	46	96.7
18	3	46620561	46620643	83	83	N/A	0.43	1245	0	0	100
19	3	49851331	49851400	70	70	CT(3), indel	0.49	1050	0	0	100
20	3	189713161	189713248	88	88	PolyA (5), PolyT (5), PolyA (9), TG(3)	0.41	1305	0	30	97.8
21	3	190106030	190106104	75	74	Indel	0.57	1108	0	2	99.8
22	4	2233667	2233744	78	78	PolyA (6)	0.26	1170	0	0	100
23	4	7780541	7780637	97	97	PolyG (6), PolyT (5), PolyA (5)	0.42	1455	0	0	100
24	4	15688604	15688681	78	78	N/A	0.29	1169	0	1	99.9
25	4	56236521	56236586	66	62	PolyA (5), indel	0.36	930	0	0	100
26	4	102839244	102839314	71	69	PolyA (5)	0.46	1035	0	0	100
27	4	164446743	164446804	62	62	PolyA (7), indel	0.27	920	0	10	98.9
28	5	1882081	1882158	78	75	N/A	0.78	450	0	675	40.0
29	5	14769061	14769144	84	84	GT(3), CCA(3)	0.62	1260	0	0	100
30	5	41069808	41069871	64	64	N/A	0.39	960	0	0	100
31	5	74077114	74077196	83	83	PolyA (6), indel	0.3	1245	0	0	100
32	5	147475343	147475409	67	67	PolyT (5)	0.37	1005	0	0	100
33	5	149323731	149323821	91	91	CT(4), AG(3)	0.55	1365	0	0	100
34	5	155662213	155662287	75	75	Indel	0.43	1125	0	0	100
35	6	6318713	6318814	102	102	PolyG (6)	0.68	1530	0	0	100

Amplicon	Chromosome	Amplicon Start	Amplicon End	Analyzed Fragment Size	Bases in Confident Regions	Amplicon Genomic Content	GC Content	Total # Correct Calls	Total # Incorrect Calls	Total # No Calls	% Correct Calls
36	6	24949983	24950074	92	92	Indel	0.63	1380	0	0	100
37	6	31084900	31084999	100	94	GCT(5), indel	0.61	1383	0	27	98.1
38	6	32147987	32148084	98	98	PolyT (5), TCT(3), CTT(3)	0.55	1455	0	15	99.0
39	6	32986864	32986958	95	95	Indel	0.53	1425	0	0	100
40	6	33408498	33408583	86	86	PolyC (6)	0.7	1290	0	0	100
41	6	41647401	41647495	95	94	PolyG (5), indel	0.61	1410	0	0	100
42	6	112435865	112435955	91	91	PolyA (5)	0.44	1365	0	0	100
43	7	22202076	22202148	73	73	N/A	0.44	1095	0	0	100
44	7	66276100	66276187	88	88	Indel	0.35	1320	0	0	100
45	7	77365735	77365821	87	87	PolyA (7), AG(4)	0.26	1299	0	6	99.5
46	7	110939946	110940030	85	85	Indel	0.38	1275	0	0	100
47	7	128533468	128533557	90	90	PolyG (5), indel	0.62	1350	0	0	100
48	7	149503875	149503965	91	91	PolyG (6), PolyC (6), indel	0.71	1365	0	0	100
49	7	154404519	154404599	81	66	N/A	0.31	990	0	0	100
50	7	156476507	156476599	93	93	Indel	0.35	1395	0	0	100
51	8	1817312	1817394	83	83	N/A	0.42	1245	0	0	100
52	8	24811020	24811109	90	89	PolyG (7), CTC(4), indel	0.61	1305	0	30	97.8
53	8	76518625	76518691	67	67	Indel	0.3	1005	0	0	100
54	9	103054909	103055006	98	98	PolyG (6)	0.67	1470	0	0	100
55	9	105586150	105586214	65	65	Indel	0.32	973	0	2	99.8
56	9	107620823	107620918	96	96	N/A	0.49	1440	0	0	100
57	9	123769149	123769231	83	83	AT(3)	0.37	1242	0	3	99.8
58	9	138995345	138995441	97	97	PolyC (6), indel	0.68	1455	0	0	100
59	10	5987120	5987198	79	78	PolyG (5), indel	0.47	1170	0	0	100
60	10	11784629	11784726	98	91	GC(3)	0.87	0	0	1365	0
61	10	27317777	27317855	79	79	PolyT (5)	0.3	1185	0	0	100
62	10	33018351	33018440	90	90	PolyA (5), PolyT (5)	0.2	1350	0	0	100
63	10	45084159	45084253	95	95	Indel	0.35	1425	0	0	100

Amplicon	Chromosome	Amplicon Start	Amplicon End	Analyzed Fragment Size	Bases in Confident Regions	Amplicon Genomic Content	GC Content	Total # Correct Calls	Total # Incorrect Calls	Total # No Calls	% Correct Calls
64	10	55892599	55892687	89	88	AC(11), indel	0.42	1290	0	69	94.9
65	10	101611250	101611329	80	80	N/A	0.49	1200	0	0	100
66	10	118351373	118351453	81	81	N/A	0.51	1215	0	0	100
67	11	8159816	8159912	97	96	N/A	0.45	1440	0	0	100
68	11	30177648	30177717	70	70	Indel	0.46	1050	0	0	100
69	11	47470345	47470444	100	100	N/A	0.65	1500	0	0	100
70	11	59837679	59837740	62	62	Indel	0.37	930	0	0	100
71	11	64418856	64418957	102	102	N/A	0.59	1530	0	0	100
72	11	93529612	93529684	73	73	PolyA (5)	0.4	1095	0	0	100
73	11	101347052	101347136	85	85	N/A	0.42	1275	0	0	100
74	11	102477336	102477426	91	91	PolyG (6)	0.55	1365	0	0	100
75	11	118406285	118406369	85	85	Indel	0.53	1275	0	0	100
76	11	120357801	120357885	85	85	PolyA (5), CA(3), indel	0.34	1275	0	0	100
77	11	125769313	125769397	85	85	GA(3)	0.52	1275	0	0	100
78	12	2834770	2834853	84	84	PolyC (5), indel	0.52	1260	0	14	98.9
79	12	26811004	26811096	93	93	PolyA (7), AC(4)	0.33	1395	0	0	100
80	12	30881766	30881846	81	81	N/A	0.49	1215	0	0	100
81	12	88474105	88474175	71	71	PolyA (6)	0.35	1065	0	0	100
82	12	120966872	120966966	95	95	PolyG (5)	0.68	1425	0	0	100
83	12	24167504	24167576	73	73	N/A	0.52	1095	0	0	100
84	13	25816961	25817049	89	88	PolyA (5), PolyT (7), PolyA (7), indel	0.22	1305	0	15	98.9
85	13	44880112	44880200	89	89	Indel	0.49	1335	0	0	100
86	13	77665218	77665294	77	77	Indel	0.39	1155	0	0	100
87	14	31619327	31619393	67	67	GA(3),TA(3)	0.39	1005	0	0	100
88	14	39517884	39517966	83	83	N/A	0.25	1245	0	0	100
89	14	46958962	46959034	73	72	PolyT (5), indel	0.19	1038	0	42	96.1
90	14	58050030	58050110	81	81	Indel	0.38	1215	0	0	100
91	14	82390559	82390649	91	91	Indel	0.35	1365	0	0	100

Amplicon	Chromosome	Amplicon Start	Amplicon End	Analyzed Fragment Size	Bases in Confident Regions	Amplicon Genomic Content	GC Content	Total # Correct Calls	Total # Incorrect Calls	Total # No Calls	% Correct Calls
92	14	92549544	92549609	66	66	PolyA (5)	0.41	975	0	60	94.2
93	14	102808496	102808589	94	94	Indel	0.62	1410	0	0	100
94	15	43170751	43170848	98	96	PolyC (5)	0.45	1440	0	0	100
95	15	63446149	63446216	68	68	Indel	0.25	1020	0	0	100
96	15	77879807	77879901	95	93	PolyG (5), indel	0.68	1395	0	0	100
97	15	81625334	81625428	95	95	PolyT (6)	0.43	1425	0	0	100
98	15	85438263	85438334	72	71	Indel	0.65	1065	0	0	100
99	15	89817413	89817503	91	91	N/A	0.36	1365	0	0	100
100	15	89864274	89864343	70	70	Indel	0.56	1050	0	0	100
101	16	1894910	1894972	63	63	N/A	0.27	945	0	0	100
102	16	28997904	28997998	95	95	PolyC (5)	0.67	1425	0	0	100
103	16	3682908	53682994	87	87	TA(3)	0.41	1305	0	0	100
104	16	57954406	57954509	104	104	PolyC (5)	0.67	1560	0	0	100
105	16	85706375	85706465	91	91	Poly T (5), indel	0.37	1362	0	3	99.8
106	17	3563920	3564008	89	89	GC(3)	0.64	1335	0	0	100
107	17	3594191	3594277	87	87	PolyC (5), indel	0.67	1303	0	2	99.8
108	17	3970090	3970180	91	91	Indel	0.46	1365	0	0	100
109	17	16084945	16085037	93	93	Indel	0.26	1395	0	0	100
110	17	33998759	33998849	91	89	PolyT (5)	0.54	1335	0	0	100
111	17	39589691	39589774	84	82	PolyA (13), indel (x2)	0.29	1215	0	78	94.0
112	17	41244394	41244484	91	91	PolyA (5)	0.34	1365	0	0	100
113	17	45438866	45438957	92	92	PolyA (7), AT(3), AT(4), AT (4), indel	0.26	1365	0	15	98.9
114	17	61502432	61502510	79	79	Indel	0.41	1175	0	10	99.2
115	17	64023582	64023667	86	86	PolyT (7)	0.22	1289	0	1	99.9
116	17	72308237	72308320	84	84	GAG(3)	0.62	1260	0	0	100
117	18	2616456	2616522	67	67	GA(3)	0.31	1005	0	0	100
118	18	6980478	6980568	91	91	N/A	0.37	1365	0	0	100
119	18	9888026	9888094	69	69	PolyA (6), TG(3)	0.43	1035	0	0	100

Amplicon	Chromosome	Amplicon Start	Amplicon End	Analyzed Fragment Size	Bases in Confident Regions	Amplicon Genomic Content	GC Content	Total # Correct Calls	Total # Incorrect Calls	Total # No Calls	% Correct Calls
120	18	38836999	38837073	75	75	PolyA (5), indel	0.37	1121	0	19	98.3
121	18	47405382	47405462	81	81	CTC(3), indel	0.47	1215	0	0	100
122	18	54815665	54815749	85	85	CT(3), indel	00.45	1275	0	0	100
123	18	59773996	59774060	65	65	N/A	0.48	975	0	0	100
124	19	625143	625241	99	99	N/A	0.59	1478	0	7	99.5
125	19	18121418	18121491	74	74	N/A	0.68	1110	0	0	100
126	19	18186574	18186643	70	70	N/A	0.64	1050	0	0	100
127	20	746056	746149	94	94	N/A	0.61	1410	0	0	100
128	20	10633195	10633276	82	82	AC(3)	0.59	1230	0	0	100
129	20	17705633	17705708	76	76	CT(3)	0.58	1140	0	0	100
130	20	21766821	21766890	70	70	GT(3),TG(4), indel	0.46	1050	0	0	100
131	20	25278421	25278521	101	101	Indel	0.63	1515	0	0	100
132	20	50897302	50897368	67	67	Indel	0.36	1005	0	6	99.4
133	20	62331904	62331994	91	88	Poly G (6)	0.73	1320	0	0	100
134	20	62690860	62690946	87	87	Indel	0.57	1305	0	0	100
135	21	30300823	30300888	66	66	Indel	0.35	990	0	0	100
136	21	33694176	33694273	98	98	PolyT (6), CA(3)	0.54	1470	0	0	100
137	21	36710706	36710792	87	87	GT(3), indel	0.39	1305	0	0	100
138	21	46644924	46644992	69	69	PolyA (6), AG(3), indel	0.32	1029	0	7	99.3
139	21	46705575	46705664	90	90	PolyT (5), PolyA (6)	0.5	1350	0	0	100
140	22	25750774	25750873	100	100	Indel	0.63	1500	0	1	99.9
141	22	32439233	32439329	97	97	N/A	0.68	1455	0	0	100
142	22	37409844	37409940	97	97	Indel	0.46	1455	0	0	100
143	22	37637596	37637694	99	99	N/A	0.6	1485	0	0	100
144	22	47081347	47081438	92	92	Indel	0.66	1380	0	0	100
145	X	15870424	15870492	69	69	PolyT (5)	0.26	1035	0	0	100
146	X	135288543	135288611	69	69	PolyC (5)	0.62	1035	0	0	100
147	X	135290777	135290847	71	71	N/A	0.52	1065	0	0	100

Amplicon	Chromosome	Amplicon Start	Amplicon End	Analyzed Fragment Size	Bases in Confident Regions	Amplicon Genomic Content	GC Content	Total # Correct Calls	Total # Incorrect Calls	Total # No Calls	% Correct Calls
148	Y	2655397	2655461	65	0	N/A	0.55	0	0	0	N/A
149	Y	2655519	2655609	91	0	N/A	0.48	0	0	0	N/A
150	Y	2655609	2655679	71	0	PolyA (5)	0.37	0	0	0	N/A

Variants that were no calls are summarized in Table 12. The particular filters that resulted in the no calls are listed in the table.

Table 12 Summary of Variants No Calls

Amplicon #	Chr:Pos	Variant	Corresponding Amplicon Content	Filter	Missed Variants	Expected Variants
28	5:1882129	T > G	78% GC	LowDP[1]	8	13
52	8:24811064	AG > A	PolyG (7), CTC(4), 61% GC	R3x6[2]	15	15
60	10:11784633	C > T	PolyGC (3), 87% GC	LowDP	13	13
64	10:55892600	TAC > T	AC(11), 42% GC	R3x6	9	9
111	17:39589692	C > CA	PolyA (13), 29% GC	R3x6	13	13

[1] LowDP: Low coverage. A variant is filtered if the depth in at least one of the pools at this particular position is below 900.
[2] R3x6: Repeat filter. A variant is filtered if all or part of the variant is present repeatedly in the reference genome adjacent to the variant position. At least six repeats in the reference are required and only repeats with length up to 3 bp are considered.

The sequencing results for sample were compared to a highly confident genotype for NA12878, established by the National Institutes of Standards and Technology (NIST) (v.2.19). Out of the 150 amplicons, 92 amplicons were fully contained within the highly confident genomic regions, 41 amplicons had partial overlap, and 17 amplicons had no overlap in the NIST sequence. This resulted in 10,000 coordinates per replicate for comparison. Non-variant base calls were compared to human genome reference sequence build hg19. The accuracy results are shown in Table 13.

Table 13 Agreement of the MiSeqDx Instrument Base Call Results with NIST Reference for GM12878 Sample

Sample	# Amplicons	Mean Call Rate	Total # TP Variant Calls	Total # FN Variant Calls	Total # TN Calls	Total # FP Calls	PPA	NPA	OPA
GM12878	150	98.43	206	0	19231	0	100	100	100

The five undiluted samples were further analyzed for calling small insertions and deletions (indels) (Table 14). In some cases, the indel was common among two or more samples as reflected in the Total # Sample Replicates with Indel column. Results for both replicates of the five samples are included in Table 14. There were a total of 71 indels ranging in size from 1-24 bp for insertions and 1-25 bp for deletions. 68 indels were each detected with a positive percent agreement of 1. Three insertions and deletions had no correct calls

because each of these variants was a no-call due to the R3x6 filter. Therefore, PPA, which excludes no calls, could not be calculated. The three variants were 1 bp deletion (chr8 24811064 AG>A); 2 bp deletion (chr10 55892600 TAC>T); and 1 bp insertion (chr17 39589692 C>CA).

Table 14 Summary of Indel Detection with the MiSeqDx Instrument

Amplicon	Chromosome	Position	Analyzed Fragment Size	Amplicon Indel Type and Length	Indel	Total # Sample Replicates with Indel	# of no calls	Total # Incorrect Indel Calls	Total # Correct Indel Calls	PPA
1	1	36450544	93	25 bp deletion	GAAAATTTAATGAAACACATTGTCCT>G	7	0	0	7	100
2	1	109465165	79	3 bp deletion	ACTT>A	9	0	0	9	100
3	1	218353908	91	23 bp insertion	T>TTTTAATAGCAAAAAGAGGCTAGA	15	0	0	15	100
4	1	223906701	92	17 bp deletion	GACAGACTGTGAGGAAGA>G	11	0	0	11	100
6	1	236372081	70	5 bp insertion	C>CTTAAG	9	0	0	9	100
7	1	247812083	88	3 bp insertion	C>CATG	9	0	0	9	100
8	2	55862804	90	7 bp insertion	T>TTTGGTAA	13	0	0	13	100
9	2	87003972	80	6 bp deletion	TTATCTC>T	11	0	0	11	100
13	2	200796749	87	5 bp insertion	T>TTAAAA	15	0	0	15	100
14	2	212245090	91	12 bp insertion	C>CTGAAAATAGGAT	11	0	0	11	100
16	2	235016388	73	2 bp insertion	A>ATG	9	0	0	9	100
17	3	4466274	93	23 bp deletion	TAACTTAAAATTACAAAATAACCC>T	13	0	0	13	100
19	3	49851375	70	9 bp insertion	C>CCTGGCTCCT	7	0	0	7	100
21	3	190106071	75	1 bp deletion	AG>A	13	0	0	13	100
25	4	56236567	66	8 bp deletion	TAACCGAAA>T	9	0	0	9	100
27	4	164446785	62	11 bp insertion	T>TTATGGTATTGA	9	0	0	9	100
31	5	74077155	83	4 bp deletion	TAGTA>T	7	0	0	7	100
34	5	155662255	75	8 bp insertion	G>GCCTACTGA	13	0	0	13	100
36	6	24950035	92	21 bp deletion	CCCTGGGTGCTATAGCCCACCA>C	11	0	0	11	100
37	6	31084942	100	3 bp deletion	GCTT>G	15	0	0	15	100
39	6	32986905	95	25 bp deletion	CTTTCACTTTCCCGTCTCATGCAAAG>C	7	0	0	7	100
41	6	41647442	95	23 bp deletion	GGCATGAGGCTTGGTGACATGGCA>G	11	0	0	11	100
44	7	66276142	88	1 bp insertion	C>CT	13	0	0	13	100
46	7	110939983	85	4 bp deletion	CAAGT>C	13	0	0	13	100

Amplicon	Chromosome	Position	Analyzed Fragment Size	Amplicon Indel Type and Length	Indel	Total # Sample Replicates with Indel	# of no calls	Total # Incorrect Indel Calls	Total # Correct Indel Calls	PPA
47	7	128533514	90	1 bp insertion	T>TC	15	0	0	15	100
48	7	149503916	91	4 bp deletion	GGATA>G	7	0	0	7	100
50	7	156476548	93	11 bp deletion	GAATCTGCACTT>G	13	0	0	13	100
52	8	24811064	90	1 bp deletion	AG>A	15	15	0	0	N/A
53	8	76518677	67	4 bp insertion	T>TACTG	9	0	0	9	100
55	9	105586193	65	4 bp insertion	C>CAATT	13	0	0	13	100
58	9	138995370	97	21 bp deletion	TCTGGGGGGCAGCCCCTGAGGG>T	9	0	0	9	100
59	10	5987158	79	3 bp deletion	TAAC>T	11	0	0	11	100
63	10	45084202	95	16 bp deletion	AGCGTCTATAACCAAAT>A	11	0	0	11	100
64	10	55892600	89	2 bp deletion	TAC>T	9	9	0	0	100
68	11	30177690	70	2 bp insertion	C>CTG	7	0	0	7	100
70	11	59837721	62	8 bp insertion	T>TTATGAAAA	11	0	0	11	100
75	11	118406328	85	8 bp deletion	CAGTGTGGA>C	9	0	0	9	100
76	11	120357842	85	2 bp deletion	CTT>C	11	0	0	11	100
78	12	2834814	84	21 bp insertion	T>TTCTCAGTACGGTGAACCCCAG	15	0	0	15	100
84	13	25817002	89	19 bp insertion	C>CAAAATATAAAAAGCTCCCT	15	0	0	15	100
85	13	44880152	89	4 bp insertion	C>CCTGT	11	0	0	11	100
86	13	77665265	77	20 bp deletion	ATCTATTTTCTAATAGACGGC>A	9	0	0	9	100
89	14	46958967	73	22 bp deletion	TTTAAAATTTGAATGTGATAAAA>T	15	0	0	15	100
90	14	58050081	81	4 bp insertion	C>CTGAT	13	0	0	13	100
91	14	82390602	91	16 bp deletion	CTTGCTCTATAAACCGT>C	11	0	0	11	100
93	14	102808554	94	5 bp deletion	CGTGGA>C	9	0	0	9	100
95	15	63446199	68	6 bp deletion	CAAAATT>C	11	0	0	11	100
96	15	77879862	95	25 bp deletion	GCCCCTGAGCCAGCCTCCCGCTCTTA>G	9	0	0	9	100
98	15	85438311	72	3 bp insertion	C>CTTG	9	0	0	9	100
100	15	89864316	70	4 bp insertion	G>GCTAC	9	0	0	9	100
105	16	85706416	91	7 bp deletion	ATTATTTC>A	11	0	0	11	100
107	17	3594276	87	1 bp deletion	TG>T	13	0	0	13	100

Amplicon	Chromosome	Position	Analyzed Fragment Size	Amplicon Indel Type and Length	Indel	Total # Sample Replicates with Indel	# of no calls	Total # Incorrect Indel Calls	Total # Correct Indel Calls	PPA
108	17	3970133	91	18 bp insertion	A>ATCCTATTCTACTCTGAAT	11	0	0	11	100
109	17	16084985	93	4 bp insertion	A>AACAC	7	0	0	7	100
111	17	39589692	84	1 bp insertion	C>CA	13	13	0	0	100
112	17	39589739	84	24 bp insertion	T>TTCTGAAGGTCAAGTCTATCCCTGA	15	0	0	15	100
113	17	45438886	92	4 bp deletion	CAGTG>C	7	0	0	7	100
114	17	61502459	79	12 bp deletion	TTTGTATCTGCTG>T	13	0	0	13	100
120	18	38837054	75	22 bp insertion	T>TGTATCTTAGCAAAAGTTTCTCA	15	0	0	15	100
121	18	47405425	81	3 bp insertion	T>TGAG	11	0	0	11	100
122	18	54815706	85	2 bp deletion	ACT>A	13	0	0	13	100
130	20	21766863	70	15 bp deletion	TACTTGAGAACTGAGG>T	9	0	0	9	100
131	20	25278464	101	5 bp insertion	A>AGTGGG	13	0	0	13	100
132	20	50897361	67	11 bp insertion	G>GGAATGTCAGCC	15	0	0	15	100
134	20	62690925	87	16 bp deletion	TCCTGGCTGGCCTGTGG>T	9	0	0	9	100
135	21	30300873	66	11 bp insertion	G>GATAAAACTTTA	9	0	0	9	100
137	21	36710749	87	21 bp deletion	ACTCAAGATAACTCATGTTATC>A	9	0	0	9	100
138	21	46644985	69	5 bp deletion	GTTGTT>G	13	0	0	13	100
140	22	25750814	100	6 bp insertion	C>CAGGGCA	13	0	0	13	100
142	22	37409885	97	5 bp insertion	C>CTGTTT	13	0	0	13	100
144	22	47081407	92	10 bp deletion	GGGCACAGGCA>G	7	0	0	7	100

Study 2

This study used banked FFPE colorectal cancer tissue samples and a representative two gene assay, which was compared to the reference method, bidirectional Sanger sequencing (Sanger). Of the 1183 total subjects, 441 subjects had valid Sanger and representative assay results. When evaluated at the subject level (Table 15), 230 of the 441 subjects were positive by Sanger (mutation detected by Sanger). Of these, 227 were positive by the representative assay. The remaining 211 of 441 subjects were negative by Sanger (no mutation detected by Sanger). Of these, 206 were negative by the representative assay. This resulted in a Positive Percent Agreement (PPA) of 98.7% and a Negative Percent Agreement (NPA) of 97.6% (Table 15).

Table 15 Positive and Negative Percent Agreement of Subject-level Results

Representative Assay	Sanger		Total
	Positive	Negative	
Positive	227[1]	5	232
Negative	3[2]	206	209
Total	230	211	441

Performance Summary		
Agreement Statistic	Point Estimate	Exact 95% CI
PPA	227/230 = 98.7%	[96.2%,99.7%]
NPA	206/211 = 97.6%	[94.6%,99.2%]

[1]There were 224 exact matches for within-subject, all-mutation level results. For two subjects, MiSeqDx detected the Sanger detected mutation and one additional mutation. For one subject, MiSeqDx and Sanger detected different mutations.
[2]One subject had two mutations detected by Sanger. Two subjects had one mutation detected by Sanger.

Study 3

This study assessed DNA libraries prepared with FFPE specimens across multiple tissue types. A total of 109 FFPE specimens from eight different tissues (colon, ovary, pancreas, adrenal, bladder, liver, thyroid, and breast) with at least 11 FFPE specimens representing each tissue type). The adrenal tissue included metastasis from esophageal, lung, colon tumors. The other tissue had primary tumors. This study used a representative assay design to query 26 genes covering 21,577 bases across 17 different chromosomes. A total of six different genes (*KRAS*, *NRAS*, *TP53*, *PIK3CA*, *EGFR*, and *BRAF*) were Sanger sequenced with each tumor having 1-3 genes Sanger sequenced based on expected prevalence of somatic mutations for that tumor. Sanger sequencing results identified 39 SNV somatic mutations in 33 of 109 FFPE specimens. The MiSeqDx identified 36 SNV somatic mutations in 32 out of 109 FFPE specimens with one false negative and two variant position no calls. PPA was 97.3%. The MiSeqDx identified 78,975 reference bases across the 109 FFPE specimens with 29 false positives relative to Sanger sequencing and 3416 no calls. NPA was 99.9%. A two base deletion was concordant between the two methods. Table 16 summarizes the results by tissue type.

Table 16 Positive and Negative Percent Agreement by Tissue Type

Tissue Type	# Samples	Total # Variants	Total # TP Variants	Total # FN Variants	Total # TN Calls	Total # FP Calls	Total # No Calls	PPA	NPA
Adrenal	16	6	4	1	11823	2	607	80	>99.9
Bladder	12	4	4	0	7070	3	273	100	>99.9
Breast	16	3	3	0	13439	7	479	100	99.9
Colon	11	6	5	0	8720	2	133	100	>99.9
Liver	13	3	3	0	7984	1	59	100	>99.9
Ovary	13	7	7	0	10581	1	724	100	>99.99
Pancreas	17	7	7	0	11929	12	489	100	99.9
Thyroid	11	3	3	0	7429	1	652	100	>99.9
Total	109	39	36	1	78975	29	3416	97.3	>99.9

Reproducibility

Two studies were conducted to assess the reproducibility of the MiSeqDx instrument with DNA extracted from FFPE samples. Study 1 used multiple instruments. Study 2 had multiple sites.

Study 1

The reproducibility of the MiSeqDx instrument was determined using two instruments and two trained operators for a total of eight runs. The representative assay, amplicon genomic context, samples, and reference method are the same as described for accuracy study 1 above. The results are presented on per amplicon basis for each instrument (Table 17) to demonstrate reproducibility of calling across instruments. % correct calls included both incorrect and no calls (one or more filters are not met for variant calling). The instruments generated similar numbers of no calls dependent on the particular amplicon. A single incorrect call within a confident region as defined by the Platinum Genomes reference standard was observed for MiSeqDx 1. The incorrect call was a false positive call of an insertion variant in amplicon 64 querying chromosome 10 at positions 55892599 to 55892687. The amplicon had a dinucleotide repeat of 11.

Table 17 Study Instrument to Instrument Reproducibility Results for the MiSeqDx Instrument (Amplicon-level)

Amplicon	Chromosome	Analyzed Fragment Size	Bases in Confident Regions	Amplicon Genomic Content	GC Content	MiSeqDx 1			MiSeqDx 2		
						Total Correct Calls	Total Incorrect Calls	Total # No Calls	Total Correct Calls	Total Incorrect Calls	Total # No Calls
1	1	93	93	Indel	0.22	5580	0	0	5580	0	0
2	1	79	79	PolyA (5), PolyC (5), indel	0.38	4740	0	0	4740	0	0
3	1	91	91	Indel	0.4	5448	0	12	5453	0	8
4	1	92	92	Indel	0.49	5518	0	2	5518	0	2
5	1	81	81	PolyG (5)	0.69	4858	0	2	4860	0	0
6	1	70	70	PolyT (10) indel	0.39	4200	0	0	4200	0	0
7	1	88	88	PolyA (5), CT(3), TAA(3), indel	0.27	5279	0	1	5279	0	1
8	2	90	90	Indel	0.28	5400	0	0	5400	0	0
9	2	80	80	Indel	0.38	4800	0	0	4800	0	0
10	2	85	81	N/A	0.65	4859	0	1	4859	0	1
11	2	75	75	PolyA (8)	0.35	4468	0	40	4468	0	40
12	2	88	88	PolyT (5)	0.42	5280	0	0	5280	0	0
13	2	87	87	PolyT (5), indel	0.31	5211	0	43	5214	0	40
14	2	91	91	PolyT (5), PolyA (6), indel	0.3	5453	0	7	5449	0	11
15	2	93	93	N/A	0.43	5579	0	1	5579	0	1
16	2	73	73	PolyT (5), indel	0.42	4378	0	2	4379	0	1

Amplicon	Chromosome	Analyzed Fragment Size	Bases in Confident Regions	Amplicon Genomic Content	GC Content	MiSeqDx 1			MiSeqDx 2		
						Total Correct Calls	Total Incorrect Calls	Total # No Calls	Total Correct Calls	Total Incorrect Calls	Total # No Calls
17	3	93	93	AT(3), indel	0.27	5396	0	184	5396	0	184
18	3	83	83	N/A	0.43	4980	0	0	4980	0	0
19	3	70	70	CT(3), indel	0.49	4193	0	7	4194	0	6
20	3	88	88	PolyA (5), PolyT (5), PolyA (9), TG(3)	0.41	5220	0	120	5220	0	120
21	3	75	74	Indel	0.57	4432	0	8	4432	0	8
22	4	78	78	PolyA (6)	0.26	4676	0	4	4676	0	4
23	4	97	97	PolyG (6), PolyT (5), PolyA (5)	0.42	5820	0	0	5820	0	0
24	4	78	78	N/A	0.29	4679	0	1	4677	0	3
25	4	66	62	PolyA (5), indel	0.36	3720	0	0	3720	0	0
26	4	71	69	PolyA (5)	0.46	4140	0	0	4140	0	0
27	4	62	62	PolyA (7), indel	0.27	3676	0	45	3671	0	51
28	5	78	75	N/A	0.78	3368	0	1132	3485	0	1015
29	5	84	84	GT(3), CCA(3)	0.62	5040	0	0	5040	0	0
30	5	64	64	N/A	0.39	3840	0	0	3840	0	0
31	5	83	83	PolyA (6), indel	0.3	4979	0	1	4980	0	0
32	5	67	67	PolyT (5)	0.37	4020	0	0	4020	0	0
33	5	91	91	CT(4), AG (3)	0.55	5460	0	0	5460	0	0

Amplicon	Chromosome	Analyzed Fragment Size	Bases in Confident Regions	Amplicon Genomic Content	GC Content	MiSeqDx 1			MiSeqDx 2		
						Total Correct Calls	Total Incorrect Calls	Total # No Calls	Total Correct Calls	Total Incorrect Calls	Total # No Calls
34	5	75	75	Indel	0.43	4498	0	6	4500	0	1
35	6	102	102	PolyG (6)	0.68	6120	0	0	6120	0	0
36	6	92	92	Indel	0.63	5520	0	0	5520	0	0
37	6	100	94	GCT(5), indel	0.61	5532	0	108	5532	0	108
38	6	98	98	Poly T (5), TCT(3), CTT(3)	0.55	5820	0	60	5820	0	60
39	6	95	95	Indel	0.53	5697	0	3	5698	0	2
40	6	86	86	PolyC (6)	0.7	5159	0	1	5160	0	0
41	6	95	94	PolyG (5), indel	0.61	5638	0	2	5638	0	2
42	6	91	91	PolyA (5)	0.44	5460	0	0	5460	0	0
43	7	73	73	N/A	0.44	4380	0	0	4380	0	0
44	7	88	88	Indel	0.35	5279	0	1	5276	0	4
45	7	87	87	PolyA (7), AG(4)	0.26	5184	0	36	5181	0	39
46	7	85	85	Indel	0.38	5100	0	0	5100	0	0
47	7	90	90	PolyG (5), indel	0.62	5398	0	2	5399	0	1
48	7	91	91	PolyG (6), PolyC (6), indel	0.71	5460	0	0	5459	0	1
49	7	81	66	N/A	0.31	3960	0	0	3960	0	0
50	7	93	93	Indel	0.35	5580	0	0	5579	0	1
51	8	83	83	N/A	0.42	4980	0	0	4980	0	0

Amplicon	Chromosome	Analyzed Fragment Size	Bases in Confident Regions	Amplicon Genomic Content	GC Content	MiSeqDx 1			MiSeqDx 2		
						Total Correct Calls	Total Incorrect Calls	Total # No Calls	Total Correct Calls	Total Incorrect Calls	Total # No Calls
52	8	90	89	PolyG (7), CTC(4), indel	0.61	5219	0	121	5220	0	120
53	8	67	67	Indel	0.3	4020	0	0	4020	0	0
54	9	98	98	PolyG (6)	0.67	5879	0	1	5880	0	0
55	9	65	65	Indel	0.32	3894	0	6	3895	0	5
56	9	96	96	N/A	0.49	5760	0	0	5760	0	0
57	9	83	83	AT(3)	0.37	4973	0	7	4978	0	2
58	9	97	97	PolyC (6), indel	0.68	5817	0	3	5818	0	2
59	10	79	78	PolyG (5), indel	0.47	4679	0	1	4680	0	0
60	10	98	91	GC(3)	0.87	450	0	5010	632	0	4828
61	10	79	79	PolyT (5)	0.3	4740	0	0	4740	0	0
62	10	90	90	PolyA (5), PolyT (5)	0.2	5400	0	0	5400	0	0
63	10	95	95	Indel	0.35	5699	0	1	5699	0	1
64	10	89	88	AC(11), indel	0.42	5157	0	276	5153	2	273
65	10	80	80	N/A	0.49	4800	0	0	4800	0	0
66	10	81	81	N/A	0.51	4860	0	0	4860	0	0
67	11	97	96	N/A	0.45	5760	0	0	5760	0	0
68	11	70	70	Indel	0.46	4199	0	2	4200	0	1
69	11	100	100	N/A	0.65	5999	0	1	5998	0	2
70	11	62	62	Indel	0.37	3720	0	0	3720	0	0
71	11	102	102	N/A	0.59	6120	0	0	6118	0	2
72	11	73	73	PolyA (5)	0.4	4380	0	0	4380	0	0

Amplicon	Chromosome	Analyzed Fragment Size	Bases in Confident Regions	Amplicon Genomic Content	GC Content	MiSeqDx 1			MiSeqDx 2		
						Total Correct Calls	Total Incorrect Calls	Total # No Calls	Total Correct Calls	Total Incorrect Calls	Total # No Calls
73	11	85	85	N/A	0.42	5100	0	0	5100	0	0
74	11	91	91	PolyG (6)	0.55	5437	0	23	5441	0	19
75	11	85	85	Indel	0.53	5100	0	0	5100	0	0
76	11	85	85	Poly A (5), CA(3), indel	0.34	5100	0	0	5100	0	0
77	11	85	85	GA(3)	0.52	5100	0	0	5100	0	0
78	12	84	84	PolyC (5), indel	0.52	5040	0	60	5038	0	63
79	12	93	93	PolyA (7), AC(4)	0.33	5577	0	3	5573	0	7
80	12	81	81	N/A	0.49	4860	0	0	4860	0	0
81	12	71	71	PolyA (6)	0.35	4260	0	0	4260	0	0
82	2	95	95	PolyG (5)	0.68	5605	0	95	5605	0	95
83	13	73	73	N/A	0.52	4380	0	0	4379	0	1
84	13	89	88	PolyA (5), PolyT (7), PolyA (7), indel	0.22	5220	0	60	5220	0	60
85	13	89	89	Indel	0.49	5340	0	0	5340	0	0
86	13	77	77	Indel	0.39	4620	0	0	4620	0	0
87	14	67	67	GA(3),TA (3)	0.39	4020	0	0	4020	0	0
88	14	83	83	N/A	0.25	4980	0	0	4980	0	0
89	14	73	72	PolyT (5), indel	0.19	4173	0	147	4173	0	147
90	14	81	81	Indel	0.38	4860	0	2	4860	0	0
91	14	91	91	Indel	0.35	5459	0	1	5460	0	0

Amplicon	Chromosome	Analyzed Fragment Size	Bases in Confident Regions	Amplicon Genomic Content	GC Content	MiSeqDx 1			MiSeqDx 2		
						Total Correct Calls	Total Incorrect Calls	Total # No Calls	Total Correct Calls	Total Incorrect Calls	Total # No Calls
92	14	66	66	PolyA (5)	0.41	3900	0	240	3900	0	240
93	14	94	94	Indel	0.62	5637	0	3	5637	0	3
94	15	98	96	PolyC (5)	0.45	5760	0	0	5760	0	0
95	15	68	68	Indel	0.25	4079	0	1	4078	0	2
96	15	95	93	PolyG (5), indel	0.68	5475	0	105	5487	0	93
97	15	95	95	PolyT (6)	0.43	5699	0	1	5700	0	0
98	15	72	71	Indel	0.65	4260	0	0	4260	0	0
99	15	91	91	N/A	0.36	5460	0	0	5460	0	0
100	15	70	70	Indel	0.56	4200	0	0	4200	0	0
101	16	63	63	N/A	0.27	3780	0	0	780	0	0
102	16	95	95	PolyC (5)	0.67	5700	0	0	5700	0	0
103	16	87	87	TA(3)	0.41	5220	0	0	5220	0	0
104	16	104	104	PolyC (5)	0.67	6238	0	3	6238	0	3
105	16	91	91	PolyT (5), indel	0.37	5443	0	17	5444	0	16
106	17	89	89	GC(3)	0.64	5251	0	89	5339	0	1
107	17	87	87	PolyC (5), indel	0.67	5212	0	8	5212	0	8
108	17	91	91	Indel	0.46	5459	0	1	5459	0	1
109	17	93	93	Indel	0.26	5580	0	0	5580	0	0
110	17	91	89	PolyT (5)	0.54	5340	0	0	5340	0	0
111	17	84	82	Poly A (13), indel (x2)	0.29	4860	0	308	4860	0	07
112	17	91	91	PolyA (5)	0.34	5459	0	1	5459	0	1

Amplicon	Chromosome	Analyzed Fragment Size	Bases in Confident Regions	Amplicon Genomic Content	GC Content	MiSeqDx 1			MiSeqDx 2		
						Total Correct Calls	Total Incorrect Calls	Total # No Calls	Total Correct Calls	Total Incorrect Calls	Total # No Calls
113	17	92	92	PolyA (7), AT(3), AT (4), AT(4), indel	0.26	5460	0	60	5460	0	60
114	17	79	79	Indel	0.41	4699	0	41	4700	0	40
115	17	86	86	PolyT (7)	0.22	5153	0	7	5156	0	4
116	17	84	84	GAG(3)	0.62	5039	0	1	5039	0	1
117	18	67	67	GA(3)	0.31	4020	0	0	4020	0	0
118	18	91	91	N/A	0.37	5460	0	0	5460	0	0
119	18	69	69	PolyA (6), TG(3)	0.43	4132	0	8	4131	0	9
120	18	75	75	PolyA (5), indel	0.37	4475	0	85	4480	0	79
121	18	81	81	CTC(3), indel	0.47	4860	0	0	4860	0	0
122	18	85	85	CT(3), indel	0.45	5098	0	2	5098	0	2
123	18	65	65	N/A	0.48	3900	0	0	3900	0	0
124	19	99	99	N/A	0.59	5926	0	14	5924	0	16
125	19	74	74	N/A	0.68	4440	0	0	4438	0	2
126	19	70	70	N/A	0.64	4199	0	1	4200	0	0
127	20	94	94	N/A	0.61	5640	0	1	5638	0	3
128	20	82	82	AC(3)	0.59	4920	0	0	4920	0	0
129	20	76	76	CT(3)	0.58	4559	0	1	4558	0	2
130	20	70	70	GT(3),TG (4), indel	0.46	4200	0	0	4200	0	0
131	20	101	101	Indel	0.63	6060	0	0	6060	0	0
132	20	67	67	Indel	0.36	4020	0	31	4020	0	25

Amplicon	Chromosome	Analyzed Fragment Size	Bases in Confident Regions	Amplicon Genomic Content	GC Content	MiSeqDx 1			MiSeqDx 2		
						Total Correct Calls	Total Incorrect Calls	Total # No Calls	Total Correct Calls	Total Incorrect Calls	Total # No Calls
133	20	91	88	PolyG (6)	0.73	5277	0	3	5274	0	6
134	20	87	87	Indel	0.57	5218	0	2	5218	0	2
135	21	66	66	Indel	0.35	3959	0	1	3957	0	3
136	21	98	98	PolyT (6), CA(3)	0.54	5880	0	0	5880	0	0
137	21	87	87	GT(3), indel	0.39	5220	0	0	5220	0	0
138	21	69	69	PolyA (6), AG(3), indel	0.32	4119	0	31	4113	0	37
139	21	90	90	PolyT (5), PolyA (6)	0.5	5399	0	1	5399	0	1
140	22	100	100	Indel	0.63	5998	0	7	5997	0	5
141	22	97	97	N/A	0.68	5819	0	1	5819	0	1
142	22	97	97	Indel	0.46	5818	0	2	5816	0	4
143	22	99	99	N/A	0.6	5940	0	0	5940	0	0
144	22	92	92	Indel	0.66	5519	0	1	5519	0	1
145	X	69	69	PolyT (5)	0.26	4139	0	1	4140	0	0
146	X	69	69	PolyC (5)	0.62	4136	0	4	4137	0	3
147	X	71	71	N/A	0.52	4260	0	0	4260	0	0
148	Y	65	0	N/A	0.55	0	0	0	0	0	0
149	Y	91	0	N/A	0.48	0	0	0	0	0	0
150	Y	71	0	PolyA (5)	0.37	0	0	0	0	0	0

Reproducibility study results were analyzed on a per operator basis using variant frequency (Table 18). This analysis demonstrated that variant frequencies were consistent across the operators. Mean variant frequencies +/- 1 standard deviation are presented.

Table 18 Operator to Operator Results for the MiSeqDx Instrument

Variant Frequency Range	# Unique Variants	Total # Variants Analyzed Operator 1	Total # Variants Analyzed Operator 2	Mean (SD) Reported Variant Frequency Operator 1	Mean (SD) Reported Variant Frequency Operator 2
High-frequency (~100%)	1112	1072	1072	0.96 +/- 0.05	0.96 +/- 0.05
Medium-frequency (~50%)	3240	3151	3161	0.49 +/- 0.04	0.49 +/- 0.04
Low-frequency (3-7%)	620	618	612	0.05 +/- 0.01	0.05 +/- 0.01

Reproducibility study results for each sample are shown compounded from all eight runs (Table 19). Detection is evaluated separately for each variant type— SNVs, insertions and deletions separately. Referenced positions are excluded. This analysis demonstrated that the results for the variants were reproducible across the samples.

Table 19 Agreement of MiSeqDx Instrument Base Call Results Per Sample

	SNVs				Insertions				Deletions			
Sample	Total #	Total # TP	Total # FP	Total # FN	Total #	Total # TP	Total # FP	Total # FN	Total #	Total # TP	Total # FP	Total # FN
GM12877	592	574	2	0	336	336	0	0	228	272	0	0
GM12878	1456	1432	0	0	320	304	0	0	384	352	0	0
GM12879	912	896	0	0	336	320	0	0	288	272	0	0
GM12885	1200	1192	0	0	400	384	0	0	352	320	0	0
GM12886	1104	1104	0	0	368	352	0	0	368	352	0	0
GM12877-D1[1]	3640	3582	0	0	800	760	0	0	960	880	0	0
GM12877-D2[2]	400	398	0	0	520	516	0	0	560	556	0	0

[1]Variants with frequency greater than 20%.
[2]Variant with frequency less than 20%.

The data provided by the 8 runs in this reproducibility study support the claim that the MiSeqDx instrument can consistently sequence:

- GC content ≥ 19% (all called bases in 120 out of 120 sequenced amplicons with 19% GC content called correctly with no-call rate of 3.4%)
- GC content ≤ 73% (all called bases in 120 out of 120 sequenced amplicons with 73% GC content called correctly with no-call rate of 0.1%)
- PolyA lengths ≤ 8 (PolyA repeat of 8 nucleotides was called correctly in 120 out of 120 sequenced amplicons containing PolyA = 8)
- PolyT lengths ≤ 10 (PolyT repeat of 10 nucleotides was called correctly in 120 out of 120 sequenced amplicons containing PolyT = 10)
- PolyG lengths ≤ 6 (PolyG repeat of 6 nucleotides was called correctly in 720 out of 720 sequenced amplicons containing PolyG = 6)
- PolyC lengths ≤ 6 (PolyC repeat of 6 nucleotides was called correctly in 359 out of 360 sequenced amplicons containing PolyC = 6, with 1 no-call)

- Dinucleotide repeat lengths ≤ 4x (all called bases in 600 out of 600 sequenced amplicons with 4x dinucleotide repeat were called correctly with no-call rate of 0.4%)
- Trinucleotide repeat lengths ≤ 5x (all called bases in 120 out of 120 sequenced amplicons with 5x trinucleotide repeat were called correctly with no-call rate of 1.9%)
- 24 or fewer base insertions and 25 or fewer base deletions
  - 24-base insertions called correctly in 120 out of 120 samples
  - 25-base deletions called correctly in 182 samples and reported as no-call in 2 samples out of 184 samples

Study 2

An external study was performed to assess the reproducibility of the representative two gene assay, described in accuracy study 2, across three external testing sites (two operators per site), one reagent lot, and three non-consecutive testing days. Testing was conducted with six well-characterized sample panels of genomic DNA samples from FFPE clinical specimens or cell lines. Each panel consisted of 10 members, for a total of 60 members across panels.

The 60 panel members consisted of duplicates of four unique wild type (for panel mutations) specimens, 12 unique mutant specimens (with a single mutation) prepared at both high and low mutation frequency levels, and two unique mutant specimens (with a single mutation) prepared at a low mutation frequency level only. Each unique specimen/mutation frequency level sample (tested in duplicate in each run) had 36 possible results (2 replicates × 2 operators × 3 days × 3 sites) if all results were valid.

Percent Expected Call (PEC) across all positive and negative variants was evaluated by comparing the representative assay result to the expected mutation result (expected mutation detected or not detected) in each sample. PEC is calculated as 100% times the number of expected calls divided by the number of calls attempted. The two-sided 95% confidence interval is calculated using the Wilson score method.

Combining sites, sample pass rates were ≥ 94.7% for the first run of the sample or in samples tested in runs that were valid on first pass. The mutation-level PEC across all mutant samples was 99.6% (905/909) (95% CI; 98.9, 99.8). The number of attempted calls across all 56 panel mutations (regardless of whether a detected mutation was expected or not) for all valid samples was 58,856 (56 × 1051). Out of these 58,856 mutation level observations, there were only six incidents where the observed and expected results were discordant. The mutation-level PEC across all positive and negative variants from all mutant and wild type panel members combined was 99.99% (58,850/58,856).

Analytical Sensitivity (Limit of Blank (LoB) and Limit of Detection (LoD)

This study verified the assay cutoff and determined the Limit of Detection (LoD) for the MiSeqDx with a representative panel. Briefly, the well characterized, Platinum Genome cell lines GM12878 and GM12877 were formalin fixed and embedded in paraffin and then DNA was extracted. GM12878 was diluted with GM12877 such that the variant frequencies of seventy variants (52 SNVs, nine insertions, and nine deletions) were near 0.05. The two DNA samples were tested by two operators, using two instruments and two lots of reagents, for a total of 10 MiSeqDx sequencing runs. This resulted in 40 replicates for each variant in GM12878 and 60 replicates for each corresponding wild-type coordinate in GM12877 for each reagent lot. LoB and LoD were calculated using the classical approach stated in CLSI EP17-A2 using the nonparametric option. LoB and LoD were calculated for SNVs, insertions, and deletions separately by pooling the variant frequencies for a given variant type. The Type I error was defined as 0.01, and the type II error was defined as 0.05.

For the LoB, the pooled variant frequencies were sorted from lowest to highest, and the 99th rank position for each reagent lot for each variant type was calculated (Table 20). The MiSeqDx software uses a cutoff (the effective LoB) of 0.026 variant frequency for determining the qualitative detection of variants. The calculated limits of blank verified that this cutoff results in a type I error of no more than 0.01.

Table 20 Limit of Blank

Variant Type	Total Number of Variant Frequencies	LoB Reagent Lot 1 (%)	LoB Reagent Lot 2 (%)
SNV	3120	0.87	0.75
Insertion	540	0.79	0.60
Deletion	540	0.96	0.84

For the LoD, the percentage of individual mutation frequency for each reagent lot for each variant type falling below the cutoff of 0.026 was calculated (Table 21). Because the percentages were less than the type II error of 5% (0.05), the median of the combined variant frequencies was calculated as the LoD (Table 21). The LoD for each variant type was taken as the larger of the two values calculated for the two reagents lots: 5.45% for SNVs, 4.88% for insertions, and 5.44% for deletions.

Table 21 Limit of Detection

Reagent Lot	Variant Type	Total Number of Variant Frequencies	# of VF Measurements < 2.6%	% of VF Measurements < 2.6%	Limit of Detection (%)
1	SNV	2080	4	0.20	5.45
	Insertion	360	0	0.00	4.86
	Deletion	360	2	0.60	5.44
2	SNV	2080	26	1.30	5.44
	Insertion	360	0	0.00	4.88
	Deletion	360	0	0.00	5.24

The following studies demonstrate the performance characteristics of the MiSeqDx with another representative assay targeting 56 mutations in two clinically relevant cancer genes (Mutation panel). The mutation panel is designed to specifically detect 56 mutations in two clinically relevant cancer genes (Gene 1 and Gene 2). The assay simultaneously determines the presence or absence of each of the 56 mutations in every sequenced sample. The reference method for these studies was bi-directional Sanger sequencing.

Lot to Lot Precision

A lot to lot precision study was conducted to evaluate the performance of the MiSeqDx instrument across manufactured reagent kit lots (consisting of sample qualification, library preparation and sequencing reagents) using the two gene representative assay using a panel of five blended FFPE specimens meeting sample qualification requirements. Each FFPE specimen contained two unique mutations: one at a lower (approximately 8%) mutation frequency level and one at a high (approximately 14%) mutation frequency level. Twelve (12) observations of each of the five specimen mixes were collected over three nonconsecutive days with three reagent kit lots. The total number of observations for the study across all reagent lots was 180 observations across all specimen mixes and 360 observations across all mutation frequency levels. Across all lots and days, 99.7% (359/360) of observations showed the expected mutation result. One low frequency mutation was incorrectly called as a wild type. A variance component analysis was performed for each of the mutations/mutation frequency levels to estimate the variability of the system. The total standard deviation ranged from 0.011 to 0.029. The reagent lot component of the total standard deviation ranged from 0 to 0.015.

Revision History

Document #	Date	Description of Change
Document # 200006218 v02	October 2023	Updated references to Local Run Manager for version 4. Updated the labeling to add Australian sponsor details. Corrected discrepancies where the package insert differs from the Instrument verification study report for MiSeqDx.
Document # 200006218 v01	May 2022	Added MiSeqDx Reagent Kit v3 Micro Package Insert CN to Equipment and Materials Provided, Not Required. Added testing information for MiSeqDx Reagent Kit v3 Micro to Performance Characteristics. Removed US-specific caution note from Warnings and Precautions.
Document # 200006218 v00	November 2021	Initial release to support MOS v4.0 and Local Run Manager v3.0.

Patents and Trademarks

Contact Information

Illumina, Inc.
5200 Illumina Way
San Diego, California 92122 U.S.A.
+1.800.809.ILMN (4566)
+1.858.202.4566 (outside North America)
techsupport@illumina.com
www.illumina.com

EC REP
Illumina Netherlands B.V.
Steenoven 19
5626 DK Eindhoven
The Netherlands

Australian Sponsor
Illumina Australia Pty Ltd
Nursing Association Building
Level 3, 535 Elizabeth Street
Melbourne, VIC 3000
Australia

Product Labeling

For a complete reference of symbols that appear on product packaging and labeling, refer to the symbol key at support.illumina.com on the *Documentation* tab for your kit.