EVALUATION OF AUTOMATIC CLASS III DESIGNATION FOR
The 23andMe Personal Genome Service Carrier Screening Test for Bloom Syndrome

DECISION SUMMARY

This decision summary corrects the decision summary dated February 2015.

A. DEN Number:

DEN140044

B. Purpose for Submission:

De Novo request for evaluation of automatic class III designation for the 23andMe Personal Genome Service (PGS) Carrier Screening Test for Bloom Syndrome

C. Measurands:

Genomic DNA obtained from a human saliva sample

D. Type of Test:

The 23andMe PGS Carrier Screening Test for Bloom Syndrome, using the Illumina Infinium BeadChip (23andMe BeadChip), is designed to be capable of detecting specific single nucleotide polymorphisms (SNPs) as well as other genetic variants. The 23andMe PGS Carrier Screening Test for Bloom Syndrome is a molecular assay indicated for use for the detection of the BLM^{Ash} variant in the BLM gene from saliva collected using the OrageneDx® saliva collection device (OGD-500.001). Results are analyzed using the Illumina iScan System and Genome Studio and Coregen software. The 23andMe PGS Carrier Screening Test for Bloom Syndrome can be used to determine carrier status for Bloom syndrome, but cannot determine if a person has two copies of the BLM^{Ash} variant.

E. Applicant:

23andMe, Inc.

F. Proprietary and Established Names:

23andMe Personal Genome Service Carrier Screening Test for Bloom Syndrome

G. Regulatory Information:

1. Regulation section:

   21 CFR 866.5940

2. Classification:

Class II

3. Product code(s):

PKB

4. Panel:

82- Immunology

H. Indication(s) for use:

1. Indication(s) for use:

The 23andMe PGS Carrier Screening Test for Bloom Syndrome is indicated for the detection of the BLM^{Ash} variant in the BLM gene from saliva collected using an FDA cleared collection device (Oragene DX model OGD-500.001). This test can be used to determine carrier status for Bloom syndrome in adults of reproductive age, but cannot determine if a person has two copies of the BLM^{Ash} variant. The test is most relevant for people of Ashkenazi Jewish descent.

2. Special conditions for use statement(s):

1. For over-the-counter (OTC) use.

2. This test is not intended to diagnose a disease, to tell you anything about the health of your fetus, or your risk or your new born child's risk of developing a particular disease later on in life.

3. This test is not a substitute for visits to a healthcare provider. It is recommended that you consult with a healthcare provider if you have any questions or concerns about your results.

4. 23andMe PGS Carrier Screening Test for Bloom Syndrome does not detect all genetic variants associated with Bloom Syndrome. The absence of a variant tested does not rule out the presence of other genetic variants that may be disease related.

5. The test is intended only for autosomal recessive carrier screening in adults of reproductive age.

6. The test does not diagnose any health conditions. Results should be used along with other clinical information for any medical purposes.

7. The laboratory may not be able to process a patient's sample. The probability that the laboratory cannot process the sample can be up to 7.6%.

8. A user's ethnicity may affect how the genetic test results are interpreted.

9. Subject to meeting the limitations contained in the special controls under regulation 21 CFR 866.5940.

4. Special instrument requirements:

The 23andMe PGS Carrier Screening Test for Bloom Syndrome is to be performed using the Tecan Evo and Illumina iScan instruments.

GenomeStudio is a modular software application that is used to view and analyze genotypic data obtained from the iScan. Coregen software conducts a variety of control checks on the file, resulting in a final analytical genotype profile for each sample. The data is used to generate test reports that are based on information from reported scientific findings on genotypes.

I. Device Description:

The 23andMe Personal Genome Service (PGS) Carrier Screening Test for Bloom Syndrome (hereafter the "PGS") is a non-invasive genetic information service that combines qualitative genotyping data for an individual. The PGS is indicated for use for the detection of the BLM^{Ash} variant in the BLM gene from saliva collected using the Oragene•Dx Saliva Collection Device (Oragene Dx model OGD-500.01). The core components of the PGS consist of the saliva collection kit; custom genotyping chip; laboratory procedures, equipment and analysis; and result reporting software.

The saliva collection kit includes a sample collection tube with a unique barcode printed by the manufacturer, funnel, preservative solution, instructions for use, and pre-paid packaging for returning the sample to the processing laboratory. Saliva may be collected by spitting directly into the Oragene•Dx container or may be transferred into the Oragene•Dx container using a sponge. Saliva samples collected using Oragene•Dx are stabilized and can be transported and/or stored long term at ambient conditions.

The PGS is indicated for the detection of the BLM^{Ash} variant in the BLM gene using DNA extracted from 2 mL saliva samples that are collected in a FDA cleared or approved collection device. Illumina manufactures a custom Infinium BeadChip genotyping chip for the device. The chip is designed to detect specific single nucleotide polymorphisms (SNPs) as well as other genetic variants; all markers refer to specific positions in the National Center for Biotechnology Information (NCBI) reference human genome.

After placing an order, an individual receives via post an Oragene•Dx ® saliva collection kit. Once the saliva sample is received by the laboratory, DNA extraction and quantitation steps occur. Samples meeting a minimum DNA concentration of 15 ng/µL are processed and prepared for amplification and BeadChip addition. BeadChips are read by the Illumina iScan, which is a laser-based, high-resolution optical imaging system. The instrument reads BeadChips by employing red and green lasers to excite the fluorophores of the allele-specific extended products found on the beads. Light emissions from these fluorophores are then recorded in high-resolution images of each BeadChip section. Data from these images are analyzed to determine genotypes using Illumina's GenomeStudio software package. GenomeStudio is a modular software application that allows viewing and analyzing of

genotypic data obtained from the iScan.

J. Substantial Equivalence Information:

1. Predicate device name(s):

   No predicate device exists.

2. Predicate 510(k) number(s):

   Not applicable.

3. Comparison with predicate:

   Not applicable.

K. Standard/Guidance Document Referenced (if applicable):

Not applicable.

L. Test Principle:

The PGS is indicated to be performed using the BeadChip v4 assay (Illumina HumanOmniExpress-24 format chip), which covers ~750,000 SNPs. The BeadChip consists of silicon wafers etched to form wells loaded with silica beads, on which oligonucleotide capture probes are immobilized. DNA from saliva is fragmented and captured on a bead array by hybridization to immobilized SNP-specific primers, followed by extension with hapten-labeled nucleotides. The primers hybridize adjacent to the SNPs and are extended with a single nucleotide corresponding to the SNP allele. The incorporated hapten-modified nucleotides are detected by adding fluorescently labeled antibodies in several steps to amplify the signals. The Tecan Evo and Illumina iScan instruments are used for extraction/processing and BeadChip quantification and scanning, respectively. The genotype content is separated, analyzed, and then integrated into pre-defined report templates specific for each condition associated with each genotype. The iScan software uses the dmap file to associate signal intensity measured by the iScan Reader with bead type. The algorithm uses sequential hybridizations of dye-labeled oligonucleotides, or decoders, complementary to bead sequences to create a combinatorial decoding scheme for arrays. The approach uses longer sequences; each is designed to hybridize to a defined target with high specificity. It is capable of decoding, with high accuracy, many thousands of bead types. Each bead type is defined by a unique DNA sequence that is recognized by a complementary decoder. Genotypes are determined using the GenomeStudio software package.

The PGS is indicated to detect the BLM^{Ash} variant. The BLM^{Ash} variant results from a deletion of six letters and an insertion of seven letters in the sequence of the gene. It results in a shortened protein that does not work properly (Typical Sequence → Deletion and Insertion → TAGATTC Variant Sequence).

Intensity values that fall outside of the expected range of intensities are returned as "no-calls"; however, all homozygous variant (BLM^{Ash} II) genotype samples receive a "no call" result, since the calling software was designed not to detect BLM^{Ash} II homozygous variant genotypes. Such results will be reported as "could not be determined" to the customer in their test report. Genetic results are returned to the customer in a secure online account on the 23andMe website.

M. Performance Characteristics:

Saliva samples were collected using the Oragene·Dx saliva collection device (OGD-500.001). The samples were tested on the Illumina Infinium BeadChip. Results were analyzed using the Illumina iScan System and GenomeStudio and Coregen software.

1. Analytical performance:

   a. *Precision/Reproducibility*

   For evaluation of precision/reproducibility, two studies were conducted: the first study was conducted with human cell line samples and a second study was conducted with human saliva samples.

   Precision Study with Human Cell Line Samples
   The precision study was conducted at the b(4) TS/CCI b(4) TS/CCI Six DNA samples taken from b(4) TS/CCI cell lines (4 BLM homozygous common ("DD") samples, 1 BLM^{Ash} heterozygous ("DI") sample and 1 BLM^{Ash} homozygous rare ("II") sample) were tested over 5 days by 3 operator teams at site 1 and 33 operator teams at site 2. To confirm the BeadChip genotype, each b(4) TS/CCI sample was also sequenced by bi-directional Sanger Sequencing. Samples were tested with 4 lots of reagents (lot 1, lot 2 and lot 3 at site 1 and lot 2, lot 3 and lot 4 at site 2) using 3 Tecan instruments and 3 iScan instruments in different combinations. For each of 4 "DD" samples, there were 36 replicates per day per laboratory site with total number of 360 replicates (36 x 5 x 2). The other two samples, "DI" and "II", had 72 replicates per day per site with total number of 720 replicates (72 x 5 x 2). Information regarding samples that failed quality control (FQC) was also evaluated.

   Results of the Study Stratified by Site are presented below:

Site 1	Total number of replicates	Number of correct calls	Number of miscalls	Number of "FQC" replicates	Percent of "FQC" replicates
Sample 1 "DD"	180	173	0	7	3.89%
Sample 2 "DD"	180	174	0	6	3.33%
Sample 3	180	175	0	5	2.78%

"DD"					
Sample 4 "DD"	180	175	0	5	2.78%
Sample "DI"	360	350	0	10	2.78%
Sample "II"	360	343	0	17	4.72%
Total	1,440	1,390	0	50	3.47%

Site 2	Total number of replicates	Number of correct calls	Number of miscalls	Number of "FQC" replicates	Percent of "FQC" replicates
Sample 1 "DD"	180	171	0	9	5.00%
Sample 2 "DD"	180	174	0	6	3.33%
Sample 3 "DD"	180	176	0	4	2.22%
Sample 4 "DD"	180	179	0	1	0.56%
Sample "DI"	360	350	0	10	2.78%
Sample "II"	360	350	0	10	2.78%
Total	1,440	1,400	0	40	2.78%

Results of the study (percent of "FQC" replicates) stratified by instrument combinations are shown for each site below:

Percent of "FQC" for nine different combinations of 3 Tecan and 3 iScan instruments at b(4) TS/CCI **site**

		Percent of "FQC"			
Tecan	iScan	4 Samples "DD"	Sample "DI"	Sample "II"	Total
1	1	0.00% (0/96)	0.00% (0/48)	0.00% (0/48)	0.00%
1	2	0.00% (0/48)	4.17% (1/24)	4.17% (1/24)	2.08%
1	3	8.33% (8/96)	4.17% (2/48)	2.08% (1/48)	5.73%
2	1	0.00% (0/96)	0.00% (0/48)	0.00% (0/48)	0.00%
2	2	7.29% (7/96)	4.17% (2/48)	14.58% (7/48)	8.33%
2	3	4.17% (2/48)	12.50% (3/24)	12.50% (3/24)	8.33%
3	1	4.17% (2/48)	0.00% (0/24)	0.00% (0/24)	2.08%
3	2	0.00% (0/48)	0.00% (0/48)	0.00% (0/48)	0.00%

3	3	4.17% (4/96)	4.17% (2/48)	10.42% (5/48)	5.73%
Percent of "FQC" for nine different combinations of 3 Tecan and 3 iScan instruments at b(4) TS/CCI		Percent of "FQC"			
Tecan	iScan	4 Samples "DD"	Sample "DI"	Sample "II"	Total
4	4	4.17% (4/96)	0.00% (0/48)	4.17% (2/48)	3.13%
4	5	2.08% (2/96)	2.08% (1/48)	6.25% (3/48)	3.13%
4	6	0.00% (0/48)	0.00% (0/24)	0.00% (0/24)	0.00%
5	4	0.00% (0/112)	5.36% (3/56)	0.00% (0/56)	1.34%
5	5	6.25% (3/48)	0.00% (0/24)	0.00% (0/24)	3.13%
5	6	12.5% (10/80)	7.50% (3/40)	2.50% (1/40)	8.75%
6	4	0.00% (0/48)	0.00% (0/24)	0.00% (0/24)	0.00%
6	5	1.04% (1/96)	4.17% (2/48)	4.17% (2/48)	2.60%
6	6	0.00% (0/96)	2.08% (1/48)	4.17% (2/48)	1.56%

The percent of FQC for 9 combinations of Tecan and iScan instruments at b(4) TS/CCI ranged from 0% to 8.33%. The percent of FQC for 9 combinations of Tecan and iScan instruments at b(4) TS/CCI ranged from 0% to 8.75%

Results of the Study (percent of "FQC" replicates) Stratified by Lot of Reagents are presented below:

	Sample 1 "DD"	Sample 2 "DD"	Sample 3 "DD"	Sample 4 "DD"	Sample "DI"	Sample "II"	Total
Lot 1	8.33% (5/60)	8.33% (5/60)	0.00% (0/60)	0.00% (0/60)	5.00% (6/120)	8.33% (10/120)	**5.42%** (26/480)
Lot 2	5.00% (6/120)	0.83% (1/120)	2.50% (3/1200	0.00% (0/120)	1.67% (4/240)	3.75% (9/240)	**2.40%** (23/960)
Lot 3	4.17% (5/120)	4.17% (5/120)	5.00% (6/120)	4.17% (5/120)	1.67% (4/240)	1.67% (4/240)	**3.02%** (29/960)

Lot 4	0.00% (0/60)	1.67% (1/60)	0.00% (0/60)	1.67% (1/60)	5.00% (6/120)	3.33% (4/120)	**2.50%** (12/480)

The percent of "FQC" replicates for 4 different lots of reagents ranged from 2.4% to 5.4%.

The combined data of the reproducibility study for 6 human cell line samples are presented in the table below:

	Total number of replicates	Number of correct calls	Number of miscalls	Number of "FQC" replicates	Percent of "FQC" replicates
Sample 1 "DD"	360	344	0	16	4.44%
Sample 2 "DD"	360	348	0	12	3.33%
Sample 3 "DD"	360	351	0	9	2.50%
Sample 4 "DD"	360	354	0	6	1.67%
Sample "DI"	720	700	0	20	2.78%
Sample "II"	720	693	0	27	3.75%
Total	2,880	2,790	0	90	3.13%

96.9% (2,790/2,880) replicates produced correct genotyping results and 3.1% (90/2,880) replicates did not pass Quality Control (QC) acceptance criteria. Samples with failed QC on the first run are re-tested per laboratory SOPs; therefore, an anticipated rate of samples with two times failed QC based on precision study data of the human cell line samples is 0.1% (=0.0313 x 0.0313).

Laboratory Reproducibility Study with Saliva Samples

A reproducibility study was performed at the same 2 sites as the reproducibility study with human cell line samples with a total of 105 BLM^{Ash} homozygous common ("DD") saliva samples obtained from individuals using the 23andMe Saliva Collection kit (Oragene-DX, OGD500.001, saliva collection kit). Sample processing was performed at both sites and tested with PGS test for Bloom syndrome. Fifty samples were initially processed at the first site and 55 samples were processed at the second site. A sample swap was performed, where an aliquot of the initially processed samples were shipped to the other lab. Results for 105 saliva samples are presented below for both sites.

Laboratory Reproducibility Study with First Run and Re-run Summary Results

			Saliva samples "DD"		
			Site 1		**Site 2**
PGS Carrier test for Bloom Syndrome	First run "1 Variant Detected"		0		0
	First run "0 Variants Detected"		104		87
	First run FQC Result	Re-run "1 Variant Detected"	0		0
		Re-run "0 Variants Detected"	1		10
		Re-run FQC	0		8
Total			105		105

Since the only samples tested were homozygous common (DD), there were no results for the "1 Variant Detected" category for either site. The percent of saliva samples with "FQC" on the first run was 1.0% (1/105) at site 1 and 17.1% (18/105) at site 2. Samples with the "FQC" on the first run were re-tested and the percent of saliva samples with a final failed QC result (Re-run FQC) was 0% (0/105) at site 1 and 7.6% (8/105) at site 2.

b. *Linearity/assay Reportable Range:*

Not applicable

c. *Traceability, Stability, Expected Values (controls, calibrators, or methods):*

The PGS requires two types of controls, the sample processing control and the reproducibility control. The sample processing control material is generated from b(4) TS/CCI the 23andMe BeadChip according to routine Standard Operating Procedures (SOPs) at the contracted laboratory sites.

The reproducibility control consists of b(4) TS/CCI Each new lot of the reproducibility control is tested by comparison with reference BeadChip genotype results. DNA is extracted from this cell suspension and genotyped using the 23andMe BeadChip, according to routine SOPs at the contracted laboratory.

The sample processing control is routinely run on every sample genotyping plate and the reproducibility control is routinely run approximately once per week. Historical data from all such runs were analyzed for one lot of the sample processing control spanning 3 months and one lot of the reproducibility control spanning 1 year.

Stability protocols and acceptance criteria were reviewed and deemed acceptable. The information provided demonstrates that the sample processing control is stable for up to 3 months and the reproducibility control is stable for up to 12 months.

d. Detection Limit:

Limit of Detection testing was performed using DNA samples b(4) TS/CCI cell lines. The following samples/genotypes/replicates were tested: BLM^{Ash} (homozygous wild type, DD), 4 samples/4 replicates per sample; BLM^{Ash} (heterozygous variant, DI), 1 sample/8 replicates per sample; BLM^{Ash} (homozygous variant, II), 1 sample/8 replicates per sample. Each DNA sample was tested at the following concentrations: 5, 15, and 50 ng/µL. BeadChip genotyping was performed b(4) TS/CCI and b(4) TS/CCI DNA, where each site tested the same DNA sample replicates for each of 3 BLM^{Ash} genotypes at 3 DNA concentrations, with 3 lots of reagents. To confirm the BeadChip genotype, each b(4) TS/CCI sample was also sequenced by bi-directional Sanger sequencing. BeadChip genotypes were compared with sequenced genotypes to determine the rates of correct BeadChip genotype calls at each DNA concentration. If a sample replicate failed BeadChip or sequencing Quality Control (QC) criteria, it was marked as "FQC" ("failed QC") if the sample replicate did not demonstrate a call rate ≥ 0.980.

The lower LoD was defined as the lowest DNA concentration at which at least 95% of samples yielded the correct call at each of two laboratory sites. The LoD study yielded 100% correct call rates for all samples across all reagent lots, at all sample concentrations tested at two independent laboratory sites. Therefore, the study passed the acceptance criteria of 95% correct calls at the lowest concentration tested (5 ng/µL). The performance requirement for the PGS, has been set at a minimum of 15 ng/µL DNA and maximum of 50 ng/µL DNA.

Results from the studies also indicated that all but one replicate passed QC acceptance criteria on the re-run (per laboratory SOP) and yielded correct genotype calls. One replicate failed QC (due to call rate < 0.980) on both runs at the b(4) TS/CCI b(4) TS/CCI laboratory site, with one lot of reagent, at the 5 ng/µL concentration.

e. Interfering Substances

Endogenous Interference Study
A study was conducted to determine whether endogenous substances present during saliva collection affect BLM^{Ash} genotyping results. Saliva samples were collected from 10 individuals (with the homozygous common "DD" genotype) and sent to the contracted laboratory for DNA extraction and BeadChip genotyping. Saliva samples were split b(4) TS/CCI for addition of interfering endogenous substances. The

following substances were spiked separately into an aliquot of each saliva sample before DNA extraction:

• salivary α-amylase, 395 U/mL
• hemoglobin, 20 mg/mL
• immunoglobulin A (IgA) 0.43 mg/mL
• total protein 2.67 mg/mL (composed of 0.185 mg/mL salivary α-amylase, 0.43 mg/mL IgA, and 2.05 mg/mL Human Serum Albumin)

An additional saliva sample aliquot was not spiked and served as control. Three replicates were tested for each aliquot. For each replicate, an aliquot of the saliva sample was processed according to laboratory SOPs to extract DNA. Extracted DNA was diluted to a concentration within the analytical range. The minimum required DNA concentration was 15 ng/µL. All replicates were tested on the same day, with one lot of reagents, by one operator team, using one set of instruments.

The PGS genotypes of the samples containing each endogenous substance were compared with the genotype of the control sample to determine percent concordance. The acceptance criteria were defined as a minimum of 95% concordant genotype calls across all individual samples for each of the 4 interfering substance conditions.

Results of the study are shown in the table below:

Sample	Interfering Substance	Replicate Genotypes			Correct	Incorrect	No Call	FQC	% Concordance with control
		Rep 1	Rep 2	Rep 3					
1	**Control**	DD	DD	DD			0	0	
	Amylase	DD	DD	DD	3	0	0	0	100
	Hemoglobin	DD	DD	DD	3	0	0	0	100
	IgA	DD	DD	DD	3	0	0	0	100
	Total protein	DD	DD	DD	3	0	0	0	100
2	**Control**	DD	DD	DD			0	0	
	Amylase	DD	DD	DD	3	0	0	0	100
	Hemoglobin	DD	DD	DD	3	0	0	0	100
	IgA	DD	FQ	FQC	1	0	0	2	100
	Total protein	DD	DD	DD	3	0	0	0	100
3	**Control**	DD	DD	DD			0	0	
	Amylase	DD	DD	DD	3	0	0	0	100
	Hemoglobin	DD	DD	FQC	2	0	0	1	100

	IgA	DD	DD	FQC	2	0	0	1	100
	Total protein	DD	DD	DD	3	0	0	0	100
4	**Control**	DD	DD	DD			0	0	
	Amylase	DD	FQ	DD	2	0	0	1	100
	Hemoglobin	DD	DD	DD	3	0	0	0	100
	IgA	DD	DD	DD	3	0	0	0	100
	Total protein	DD	DD	DD	3	0	0	0	100
5	**Control**	DD	DD	DD			0	0	
	Amylase	DD	DD	DD	3	0	0	0	100
	Hemoglobin	DD	DD	DD	3	0	0	0	100
	IgA	DD	DD	DD	3	0	0	0	100
	Total protein	DD	DD	DD	3	0	0	0	100
7	**Control**	DD	DD	DD			0	0	
	Amylase	DD	DD	DD	3	0	0	0	100
	Hemoglobin	DD	DD	DD	3	0	0	0	100
	IgA	DD	DD	DD	3	0	0	0	100
	Total protein	DD	DD	DD	3	0	0	0	100
8	**Control**	DD	DD	DD			0	0	
	Amylase	DD	DD	DD	3	0	0	0	100
	Hemoglobin	DD	DD	DD	3	0	0	0	100
	IgA	DD	DD	DD	3	0	0	0	100
	Total protein	DD	DD	DD	3	0	0	0	100
9	**Control**	DD	DD	DD			0	0	
	Amylase	DD	DD	DD	3	0	0	0	100
	Hemoglobin	DD	DD	DD	3	0	0	0	100
	IgA	DD	DD	DD	3	0	0	0	100
	Total protein	DD	DD	DD	3	0	0	0	100
10	**Control**	DD	DD	DD			0	0	
	Amylase	DD	DD	DD	3	0	0	0	100

	Hemoglobin	DD	DD	DD	3	0	0	0	100
	IgA	DD	DD	DD	3	0	0	0	100
	Total protein	DD	DD	DD	3	0	0	0	100
11	**Control**	DD	DD	DD			0	0	
	Amylase	DD	FQC	FQC	1	0	0	2	100
	Hemoglobin	DD	DD	DD	3	0	0	0	100
	IgA	DD	DD	DD	3	0	0	0	100
	Total protein	DD	DD	DD	3	0	0	0	100

FQC = Replicate failed QC according to QC criteria; FQC replicates were re-run on a separate day

Out of a total 150 replicates run, 7/150 (4.7%) failed QC in this study. Of these, 5 replicates did not meet the QC criterion "call rate ≥ 0.980". Another 2 replicates had insufficient saliva volume to perform the first run. All 5 replicates with sufficient saliva volume produced the correct genotype upon re-run. The results demonstrate that there was no negative impact upon PGS test performance with all interferents tested.

Exogenous Interference Study

A study was performed to determine whether exogenous substances present in saliva during saliva collection would affect successful assignment of the correct BLM^{Ash} genotype using the 23andMe BeadChip assay.

Saliva samples were collected from 5 individuals at 3 time points – before consuming an exogenous substance (baseline), immediately (0 minutes) after, and 30 minutes after – and sent to the laboratory for genotyping. For each individual, genotype test results of the 0 and 30 minute samples were compared with genotypes of the baseline sample (considered control) to determine percent concordance. The following exogenous interference conditions were tested separately:

- Eating (food containing beef)
- Eating (food other than beef)
- Drinking
- Chewing gum
- Using mouthwash

Each exogenous substance sample was tested in triplicate for a total of 225 test samples (5 donors x 5 conditions x 3 time points x 3 replicates). For each replicate, an aliquot of the saliva sample was processed according to laboratory SOPs to extract DNA. Extracted DNA was diluted to a concentration within the analytical range. The minimum required DNA concentration was 15 ng/µL. The acceptance criteria were defined as a minimum of 95% concordant genotype calls across all individual samples at each time point. The study results are shown below:

Exogenous substance activity	Individual ID	Time point	Genotype			Concordance results			
			rep 1	rep 2	rep 3	Correct	Incorrect	No Call	FQC
Eating (beef)	1	baseline	DD	DD	DD				
		0 min	DD	DD	DD	3	0	0	0
		30 min	DD	DD	DD	3	0	0	0
	2	baseline	DD	DD	DD				
		0 min	FQC	FQC	DD	1	0	0	2
		30 min	FQC	DD	DD	2	0	0	1
	3	baseline	DD	DD	DD				
		0 min	DD	DD	DD	3	0	0	0
		30 min	DD	DD	DD	3	0	0	0
	4	baseline	DD	FQC	DD				1
		0 min	FQC	FQC	FQC	all FQC	0	0	3
		30 min	DD	FQC	DD	2	0	0	1
	5	baseline	DD	DD	DD				
		0 min	FQC	FQC	FQC	all FQC	0	0	3
		30 min	DD	DD	DD	3	0	0	0
Eating (no beef)	1	baseline	DD	DD	DD				
		0 min	DD	DD	FQC	2	0	0	1
		30 min	DD	DD	DD	3	0	0	0
	2	baseline	DD	DD	DD				
		0 min	DD	DD	DD	3	0	0	0
		30 min	DD	DD	DD	3	0	0	0
	3	baseline	DD	DD	DD				
		0 min	DD	DD	DD	3	0	0	0
		30 min	DD	DD	DD	3	0	0	0
	4	baseline	DD	DD	DD				
		0 min	DD	DD	DD	3	0	0	0
		30 min	DD	DD	DD	3	0	0	0
	5	baseline	DD	FQC	DD				1

		0 min	FQC	FQC	FQC	all FQC	0	0	3
		30 min	FQC	DD	DD	2	0	0	1
Drinking	1	baseline	DD	DD	DD				
		0 min	DD	DD	DD	3	0	0	0
		30 min	DD	DD	DD	3	0	0	0
	2	baseline	DD	DD	DD				
		0 min	DD	DD	DD	3	0	0	0
		30 min	DD	DD	DD	3	0	0	0
	3	baseline	DD	FQC	DD				1
		0 min	DD	DD	DD	3	0	0	0
		30 min	DD	DD	DD	3	0	0	0
	4	baseline	DD	FQC	DD				1
		0 min	DD	FQC	FQC	1	0	0	2
		30 min	DD	FQC	FQC	1	0	0	2
	5	baseline	DD	DD	DD				
		0 min	FQC	FQC	DD	1	0	0	2
		30 min	DD	DD	DD	3	0	0	0

Exogenous substance activity	**Individual ID**	**Time point**	**Genotype**			**Concordance results**			
			rep 1	**rep 2**	**rep 3**	**Correct**	**Incorrect**	**No Call**	**FQC**
Chewing gum	1	baseline	DD	DD	DD				
		0 min	DD	DD	DD	3	0	0	0
		30 min	DD	DD	DD	3	0	0	0
	2	baseline	DD	DD	DD				
		0 min	DD	FQC	FQC	1	0	0	2
		30 min	DD	DD	DD	3	0	0	0
	3	baseline	DD	DD	DD				
		0 min	DD	DD	DD	3	0	0	0
		30 min	DD	DD	DD	3	0	0	0
	4	baseline	DD	DD	DD				
		0 min	DD	DD	DD	3	0	0	0

		30 min	DD	DD	DD	3	0	0	0
	5	baseline	DD	DD	DD				
		0 min	FQC	DD	DD	2	0	0	1
		30 min	FQC	DD	DD	2	0	0	1
Using mouth-wash	1	baseline	DD	DD	DD				
		0 min	DD	DD	DD	3	0	0	0
		30 min	DD	DD	DD	3	0	0	0
	2	baseline	DD	DD	DD				
		0 min	DD	DD	DD	3	0	0	0
		30 min	DD	DD	DD	3	0	0	0
	3	baseline	DD	DD	DD				
		0 min	DD	DD	FQC	2	0	0	1
		30 min	DD	DD	DD	3	0	0	0
	4	baseline	DD	DD	DD				
		0 min	DD	DD	DD	3	0	0	0
		30 min	DD	DD	DD	3	0	0	0
	5	baseline	FQC	DD	DD				1
		0 min	DD	FQC	DD	2	0	0	1
		30 min	DD	DD	DD	3	0	0	0

FQC = Replicate failed QC according to QC criteria on the first run; FQC replicates were re-run on a separate day if saliva sample was available

Out of a total 225 replicates run, 32/225 failed QC in this study. Of these, 18 replicates did not meet the QC criterion "call rate ≥ 0.980". Another 3 replicates yielded DNA concentrations below the QC criterion of 15 ng/µL (all of these replicates were from a sample collected at the 0 minute time point). Another 11 replicates did not have sufficient saliva volume for automated DNA extraction.

All replicates that failed BeadChip QC on the original run and had sufficient saliva volume or DNA concentration ≥15 ng/µL (10/32 replicates) for a re-run passed QC acceptance criteria on the re-run (per laboratory SOPs) and yielded correct genotype calls. The results indicate that saliva samples should be collected at least 30 minutes after eating, drinking, chewing gum, or using mouthwash.

Smoking Interference Study

To evaluate the effects of smoking before saliva collection on the accuracy of the BeadChip assay for the BLM^{Ash} genotype, a group of 5 donors provided 3 samples each (baseline/control sample taken at least 60 minutes prior to smoking, samples

collected immediately after smoking and samples collected 30 minutes after smoking). The smoking exogenous substance samples were tested in triplicate for a total of 45 test samples (5 donors x 1 condition x 3 time points x 3 replicates). To confirm the BeadChip genotype, each b(4) TS/CCI sample was also sequenced by bi-directional Sanger sequencing. The smoking interference study also indicates that saliva samples should be collected at least 30 minutes after smoking. The passing criteria were a minimum of 95% concordant genotype calls across all individuals at each time point. Results are shown in the table below:

Smoking Interference Study Genotype Call Raw Data

Individual	Time Point	Replicates Genotype*		
		replicate 1	replicate 2	replicate 3
1	baseline	DD	DD	DD
	0 min	DD	DD	DD
	30 min	DD	DD	DD
2	baseline	DD	FQC	DD
	0 min	DD	DD	DD
	30 min	DD	DD	FQC
3	baseline	DD	DD	DD
	0 min	DD	DD	DD
	30 min	DD	DD	DD
4	baseline	DD	DD	DD
	0 min	DD	FQC	DD
	30 min	DD	DD	DD
5	baseline	DD	DD	DD
	0 min	FQC	FQC	DD
	30 min	DD	DD	DD

FQC = Sample failed QC according to QC criteria
*DD genotype=wild type homozygous non-carrier

Out of a total 45 replicates run, 5/45 failed QC in this study. Of these, 3 replicates did not meet the QC criterion "call rate ≥ 0.980". Another 2 replicates yielded DNA concentrations below the QC criterion of 15 ng/µL (both replicates were from a sample collected immediately after smoking). Out of the 5 samples that failed QC, 2 samples had DNA concentration <15 ng/µL, thus 3 samples that passed QC acceptance criteria on the re-run (per laboratory SOPs) yielded correct genotype calls. The results demonstrate that saliva samples should be collected at least 30 minutes after smoking.

Microbial Interference Study

A microbial interference study was performed to determine whether microbial DNA affects successful assignment of the correct BLM^{Ash} genotype using the 23andMe BeadChip assay. Six DNA samples taken from b(4) TS/CCI cell lines were tested (4 BLM^{Ash} homozygous common ("DD") samples/3 replicates, 1 BLM^{Ash} heterozygous

("DI") sample/3 replicates and 1 BLM^{Ash} homozygous rare ("II") sample/3 replicates). Genomic DNA from five microbes was spiked into an aliquot of DNA from each b(4) TS/CCI sample, followed by BeadChip genotyping. An additional aliquot was spiked with buffer and served as control. Human and microbial DNA was diluted so that the DNA concentration in the final 50/50 mixture was within the analytical range. The minimum required DNA concentration was 15 ng/µL. To confirm the BeadChip genotype, each b(4) TS/CCI sample was also sequenced by bi-directional Sanger sequencing. BeadChip genotypes were compared with sequenced genotypes to determine the rates of correct BeadChip genotype calls. The passing criteria were defined as a minimum of 95% correct genotype calls across all individual samples for each of the five microbe conditions. Results are shown in the table below:

Microbial Interference Study Genotype Call Raw Data

			Replicate Genotypes		
Genotype	**Sample**	**Microbe**	**replicate 1**	**replicate 2**	**replicate 3**
BLM^{Ash} DD	b(4) TS/CCI	**buffer control**	DD	DD	DD
		S. epidermidis	DD	DD	DD
		S. mutans	DD	DD	DD
		L. casei	DD	DD	DD
		A.	DD	DD	DD
		C. albicans	DD	DD	DD
	b(4) TS/CCI	**buffer control**	DD	DD	DD
		S. epidermidis	DD	DD	DD
		S. mutans	DD	DD	DD
		L. casei	DD	DD	DD
		A.	DD	DD	DD
		C. albicans	DD	DD	DD
	b(4) TS/CCI	**buffer control**	DD	DD	DD
		S. epidermidis	DD	DD	DD
		S. mutans	DD	DD	DD
		L. casei	DD	FQC	FQC
		A.	DD	DD	DD
		C. albicans	DD	DD	DD
	b(4) TS/CCI	**buffer control**	DD	DD	DD
		S. epidermidis	DD	DD	DD
		S. mutans	DD	DD	FQC
		L. casei	DD	DD	DD
		A.	DD	DD	DD
		C. albicans	DD	DD	DD
BLM^{Ash} DI	b(4) TS/CCI	**buffer control**	DI	DI	DI
		S. epidermidis	DI	DI	DI
		S. mutans	DI	DI	DI
		L. casei	DI	DI	DI
		A.	DI	DI	DI

		C. albicans	DI	DI	DI
BLM^{Ash} II *	b(4) TS/CCI	**buffer control**	NC	NC	NC
		S. epidermidis	NC	NC	NC
		S. mutans	NC	NC	FQC
		L. casei	NC	NC	NC
		A.	NC	NC	NC
		C. albicans	NC	NC	NC

*NC = No Call, as per protocol for BLM^{Ash} II genotype samples.
FQC = Replicate failed QC according to QC criteria; FQC replicates were re-run on a separate day

Four replicates that failed QC in this study did not meet the QC criterion "call rate ≥ 0.980". Upon re-running these samples, all replicates produced correct genotype calls.

f. Assay Cut-off:

Not applicable.

g. Specimen Stability at 2-8° C

Saliva samples for testing are collected with the Oragene·Dx OGD-500.001 collection device. See k141410 for sample stability information.

h. Shipping Stability

Saliva samples are shipped for testing in the Oragene·Dx OGD-500.001 collection device. See k141410 for sample shipping stability information.

2. Comparison Studies:

a. Method Comparison with Predicate Device:

Accuracy was evaluated by the agreement of the genetic variant determinations by this test with bi-directional sequencing results.

Saliva samples were randomly selected from the 23andMe Biobank, in which saliva samples were collected using the DNA Genotek OrageneDx 500.001 collection device. Saliva sample selection was blinded to previously determined genotypes and at least 20 carrier samples (DI, as detected by genotyping) were selected. A total of 65 saliva samples were selected for the study (25 DD and 22 DI samples tested at Site 1; 18 DI samples samples tested at Site 2) in addition to 6 human cell line samples tested (4 DD samples, 1 DI sample and 1 II sample tested at both sites). All 71 samples were sequenced using bi-directional sequencing. The comparison study was conducted at 2 sites; results of the test were compared with sequencing results. If a replicate fails QC ("FQC") criteria on the first run, the replicate was re-run once using the same sample. Five saliva samples failed QC at Site 1 and 1 saliva sample failed QC at the Site 2. All re-run samples produced correct genotype results. Study

results and % agreement are provided in the tables below:

Summary of Comparison Data of Saliva Samples for b(4) TS/CCI

			Bi-directional sequencing		
			"DI"		**"DD"**
PGS Carrier test for Bloom Syndrome	First run "1 Variant Detected"		21		0
	First run "0 Variants Detected"		0		21
	First Run FQC Result	Re-run "1 Variant Detected"	1		0
		Re-run "0 Variants Detected"	0		4
		Re-run FQC	0		0
Total			22		25

Summary of Comparison Data of Saliva Samples for b(4) TS/CCI

			Bi-directional sequencing		
			"DI"		**"DD"**
PGS Carrier test for Bloom Syndrome	First run "1 Variant Detected"		17		0
	First run "0 Variants Detected"		0		0
	First Run FQC Result	Re-run "1 Variant Detected"	1		0
		Re-run "0 Variants Detected"	0		0
		Re-run FQC	0		0
Total			18		0

The homozygous common (DD) genotype results at both sites did not produce results for the "1 Variant Detected Category" and vice versa for the heterozygous variant genotype regarding the "0 Variants Detected" category. All samples that failed QC on the first run produced correct genotype results upon re-running them at both sites.

Positive and Negative Percent Agreements for Saliva Samples for Both Sites

	Positive Percent Agreement Saliva Samples with "DI" by bi-directional sequencing		**Negative Percent Agreement** Saliva Samples with "DD" by bi-directional sequencing	
	Percent correct results	%FQC on the first run	Percent correct results	%FQC on the first run
Site 1	100% (22/22)	4.5% (1/22)	100% (25/25)	16.0% (4/25)
Site 2	100% (18/18)	5.6% (1/18)	n/a	n/a
Combined	100% (40/40)	5.0% (2/40)	100% (25/25)	16.0% (4/25)

Results for the cell line samples tested at both sites are as follows:

i) The DI sample had correct genotyping results on the first run at both sites;

ii) All 4 DD samples had correct genotyping on the first run at both sites;

iii) The II sample failed QC on the first run at Site 2 and produced a correct (no call) result upon re-running the sample.

The following table presents PPA and NPA for saliva and human cell line samples combined.

PPA and NPA for Saliva and Human Cell Line Samples Combined for Both Sites

	Percent of correct calls	**95% CI**
Positive Percent Agreement (PPA)	100% (41/41)	91.4% to 100%*
Negative Percent Agreement (NPA)	100% (29/29)	88.3% to 100%*
Overall Agreement	100% (70/70)	96.3% to 100%**

*95% two-sided confidence interval
** 95% one-sided confidence interval

Overall agreement was 100% (70/70) with 95% confidence interval of 96.3% to 100%.

b. Matrix Comparison:
Not applicable. This test is for use with human saliva samples only.

3. Clinical Studies:

Clinical Performance
The BLM^{Ash} variant covered by this test is mainly found in people of Ashkenazi Jewish descent. Approximately 1 in 107 people[1] (0.93%) with this ethnicity carries this variant. The BLM^{Ash} variant is rare and not well studied in other ethnic groups.

Carrier Frequency of the BLM^{Ash} Mutation in the 23andMe Database

Ancestry Group	Frequency	Number of Tested
Ashkenazi Jewish	1.03%	b(4) TS/CCI
European	0.02%	
Latino	<0.04%	
African American	0.00	
Asian	0.00	

More than 99% of all patients of individuals of Ashkenazi Jewish descent known to have Bloom syndrome have two copies of this mutation[1,2]; therefore, this test is expected to detect more than 99% of Bloom syndrome carriers in people of Ashkenazi Jewish descent.

Pre-test and Post-test Carrier Risks for Different Results of PGS test for Bloom Syndrome

	Ashkenazi Jewish	Other Ancestry Groups
BLM^{Ash} frequency among patients with Bloom syndrome	>99%[1]	Unknown
Pre-test carrier risk	1 in 107[1]	Likely < 1 in 107
Carrier risk for result "1 Variant Detected" of PGS test for Bloom Syndrome	99%	50%-99%*
Carrier risk for result "0 Variants Detected" of PGS test for Bloom Syndrome	< 1 in 11,000	Likely < 1 in 107

* The carrier risk depends on subject ethnicity; for some ethnicities, this risk can be lower than 50%.

References:

1. Gross, S.J., Pletcher, B.A., Monaghan, K.G. (2008). ACMG Practice Guidelines: Carrier screening in individuals of Ashkenazi Jewish descent. Genet Med. 10(1):54–56.

2. German, J., Sanz, M.M. Syndrome-Causing Mutations of the BLM Gene in Persons in the Bloom's Syndrome Registry. Hum Mut. (2007) 28(8):743-753.

Other clinical supportive data (when a. and b. are not applicable):
A user comprehension study was performed in order to assess comprehension of the proposed labeling of the PGS Test report in a demographically diverse sample and to evaluate potential factors influencing comprehension. Specifically, the user studies were performed in order to assess user comprehension of representative test reports. Furthermore, the studies were comprised of naïve participants representative of a broad intended use population, and conducted in a controlled, lab-based setting so that participant experiences with the survey could be recorded in detail.

The user comprehension study used quota-based sampling to recruit a naïve sample of participants that were demographically diverse according to age, race/ethnicity, and education level. The study was conducted at 5 locations across the U.S. A target of at least 100 subjects were tested across each of 5 different representative test reports for the PGS (Full Ashkenazi Jewish combined with either variant present, variant absent, not determined; partial Ashkenazi Jewish combined with variant absent and no Ashkenazi Jewish and variant absent) with each location recruiting at least 120 participants to achieve this target. A total of 11 of 678 (1.6%) participants were excluded after enrollment for the following reasons: careless responders (indicating an incorrect answer on a main survey question with an obvious correct answer), error in data recording, previous 23andMe study participants.

Participants were assigned to study arms at each facility in the order in which they started their session. Participants were asked to complete the following online tasks:

a. Background survey – familiarity with genetics and interest in genetic testing
b. Pre-test comprehension survey – comprehension of concepts assessed in the main comprehension survey
c. Comprehension survey instructions and education module (view only) – description of the main survey task and explanation of genetic testing concepts
d. Main comprehension survey – Representative Bloom Syndrome Test report with results for a fictional individual, paired with comprehension survey
e. Post-test usability survey – participant's general feedback about the Bloom Syndrome Test report and the survey experience

Primary comprehension assessment addressed the following comprehension concepts: purpose of the PGS, limitations of the test (variants covered), relevant ethnicities for the test, meaning of test results, and appropriate follow-up actions. Secondary analyses included assessment of participants' baseline knowledge of genetic testing concepts, survey completion rates, and evaluation of qualitative feedback from participants and moderators. The table below demonstrates the results for each study arm and testing concept:

Comprehension concept	Comprehension rates (%) within each study arm					Overall comprehension rates (%)
	Full Match Variant Absent	Full Match Variant Present	Full Match Not Determined	Partial Match Variant Absent	No Match Variant Absent	
Purpose of test	90.4	95.5	94.8	91.7	89.2	92.4
Test results		95.5	94.8			95.2
Limitations of test	88.9	91.8	94.1	93.2	90.0	91.6
Ethnicity relevance	96.3		95.6	91.7	94.6	94.6
Meaning of results	89.6	97.8		94.0	87.7	92.3
Appropriate follow-up	94.1	95.5	94.1	92.5	92.3	93.7

The overall comprehension rates across all study arms for each comprehension concept ranged from 91.6% to 95.2%. Overall comprehension rates were above 90% for all comprehension concepts across all study arms.

4. Expected values/Reference range:

Testing for the BLM^{Ash} mutation is expected to have clinical sensitivity (test coverage) greater than 99% in Ashkenazi Jewish individuals (Gross et al., 2008). Approximately 0.75-1% of individuals with Ashkenazi Jewish ancestry carry the mutation (Peleg et al., 2002, Gross et al., 2008; Scott et al., 2010), which is consistent with the observed frequency in 23andMe customers of Ashkenazi Jewish descent (see table below). The frequency of this mutation in other populations has not been reported and it is assumed to be much rarer in non-Jewish populations.

Carrier Frequency of the BLM^{Ash} Mutation in the 23andMe Database

Ancestry Group	Frequency	Number of Tested Individuals
Ashkenazi Jewish	1.03%	b(4) TS/CCI
European	0.02%	
Latino	<0.04%	
African American	0.00	
Asian	0.00	

More than 99% of all patients of individuals of Ashkenazi Jewish descent known to have Bloom syndrome have two copies of this mutation (Gross et al., 2008, German et al., 2007); therefore, this test is expected to detect more than 99% of Bloom syndrome carriers in people of Ashkenazi Jewish descent.

N. Instrument Name:

Illumina iScan BeadChip scanner with GenomeStudio software

O. System Descriptions:

1. Modes of Operation:

   The Illumina iScan is a table top laser-based, high-resolution optical imaging system that produces genotype information for up to 4 beadchips/beadchip carrier. Carriers are loaded into the instrument through an Autoloader2.

2. Software:

   FDA has reviewed applicant's Hazard Analysis and software development processes for this line of product types:

   Yes __X__ or No _____

   Level of Concern:

   Moderate

   Software Description:

   Illumina iScan System with iScan Control Software and Genome Studio performs the following:

   - iScan Control Software drives the iScan hardware in Beadchip scanning and image data generation.
   - GenomeStudio software allows viewing and analyzing of genotypic data obtained from the iScan. Processing includes primary data analyses, such as raw data normalization, clustering, and genotype calling. To ensure data quality, the software performs internal controls and data quality control checks.
   - Coregen software takes raw data and putative genotype calls from the FTL file received from the laboratory and performs the following functions in order to generate the final analytical genotype information for each sample.

   Device Hazard Analysis:

   A device hazard analysis was performed to identify all hazards associated with the device and its components. For each identified hazard, a proposed method(s) for reducing that hazard was identified. Based upon the system examination, hazards, their associated mitigation method(s) and mode of control have been identified in the supplied risk analysis table.

Architecture Design Chart:

The 23andMe software system workflow and associated software architecture design charts were provided along with illustrations and a description of each illustration.

Software Requirements Specification (SRS):

The SRS describes the major functional aspects and corresponding software components of the 23andMe Software Requirements Specification. This document translates the design inputs and the requirements into the overall 23andMe system architecture and design.

Software Design Specification (SDS):

Requirements defined in the SRS document are implemented according to the design specifications described.

Traceability Analysis:

A traceability matrix which links requirements, specifications, hazards, mitigations and verification & validation testing for the software was acceptable.

Software Development Environment Description:

Software development life cycle plan and software development configuration management plan for the 23andMe medical device software, which provides reports based on data derived from a customer's DNA as assayed by a genotyping chip and curated information from scientific literature, was acceptable.

Verification and Validation Testing:

Based on the Software Requirements Specification ("SRS") and Software Design Specifications ("SDS"), software verification test plans were devised to verify that the software meets the requirements. Each element of the SRS was tested and found to meet the requirements.

Revision Level History:

A software revision history record for the 23andMe software system software was acceptable.

Unresolved Anomalies:

There were three unresolved anomalies regarding kit registration (Rev C), none of which affected the actual functionality of the registration process only and did not affect how the customer entered and verified their sample barcode number. An impact analysis

concluded that all tests resulted in a passing state with no limitations.

EMC Testing:

Not applicable.

3. Specimen Identification:

Consumers must register their saliva collection kit, linking their saliva sample to a secure online account with a valid email address through a unique barcode, in order to use the test. The software allows consumers b(4) TS/CCI b(4) TS/CCI The barcode is also matched to records of kits shipped to consumers to ensure it is a valid kit. A timestamp of the consumer completing the entries is recorded

4. Specimen Sampling and Handling:

Saliva samples should be collected using collection devices, such as the Oragene® Dx Saliva Collection Device (OGD-500.001 or FDA cleared/approved or legally marketed exempt equivalent). The recommended volume of saliva is 2 mL. After saliva is collected, the stabilizing liquid is mixed with the sample. Saliva can be delivered directly by spitting or using provided sponges to transfer saliva into the device. Upon contacting saliva cells, the stabilizing liquid lyses cellular and nuclear membranes to release and stabilize nucleic acids. Samples can be immediately processed, transported or stored for future use. Device and sample integrity are preserved during typical ambient transport and storage conditions for up to 12 months.

After the saliva samples arrive b(4) TS/CCI, they are received and accessioned, sample information is entered into the accessioning system, a routelist is generated and reconciled, and then the routelist is imported into the processing database.

5. Calibration:

Calibration and calibration verification procedures are established to demonstrate continued accuracy of the test systems.

6. Quality Control:

The following quality control materials are used and qualified according to established material specification sheets and related procedures. The assay is monitored at each process in the following manner:

- (-) Negative Control for the extraction step consists of b(4) TS/CCI that is prepared with DNA Genotek's OrageneDx® Kit (K110701). The negative control is placed randomly on an extraction plate and is used to check for evidence of plate contamination during aliquot and extraction; there should be no DNA present in the buffer.

- The Sample Processing Control (SP) is introduced after extraction to confirm plate orientation from b(4) TS/CCI with a previously characterized genotype. The SP is considered to meet acceptance criteria if the reproducibility frequency is >0.950.

- Reproducibility Control (RC) evaluates performance at weekly intervals from extraction through detection. RC consists of the same b(4) TS/CCI as the SP control, but is assayed from b(4) TS/CCI.

P. Other Supportive Instrument Performance Characteristics Data Not Covered In The "Performance Characteristics" Section above:

Refer to k141410 for saliva collection device details and study results.

Q. Identified Risks and Required Mitigations:

Carrier screening is a type of genetic testing performed on people who display no symptoms for a recessive genetic disorder but may be at risk for passing it on to their children. A carrier for a genetic disorder has inherited one normal and one abnormal allele for a gene associated with the disorder. A child must inherit two abnormal alleles in order for symptoms to appear.

By definition, autosomal (non-sex chromosome-related) recessive disorders require that two abnormal copies of a gene, one inherited from each parent, be present in order for the disorder to be manifested. Therefore, to have a child with an autosomal recessive disorder both parents must be carriers of an abnormal gene copy. When both parents are carriers for the abnormal copy, there is a 1 in 4 chance (25%) that the child will inherit two abnormal copies of the gene and manifest the specific disease/condition or trait.

Most autosomal recessive genetic diseases are very rare with frequencies much less than 1% in the general population. Some autosomal recessive genetic diseases are more common in certain ethnic groups; however, even in these ethnic groups disease frequencies tend to be very low.

In defining an inheritance pattern of genetic disease, one must demonstrate the appropriate genetic patterns are present in an informed population that includes affected persons. Reputable journals and professional bodies would be very likely to reject a weak or mistaken

interpretation of the mode of inheritance of a genetic disease, because in order to demonstrate heritability, certain criteria must be met, e.g., a family with affected and unaffected individuals would have to display the expected segregation of alleles, or the inheritance pattern would be considered inconclusive. Therefore, there is very little danger that fraudulent or wishful claims for autosomal recessive inheritance would be accepted by such groups.

Risk of False Results:

When considering the risks of tests for being a carrier of an abnormal gene for an autosomal recessive genetic disease, one should consider the effects of both false positive and false negative results, as well as the appropriate mitigations for these risks. Note that the risks of carrier testing are generally similar regardless of the genetic carrier condition to be detected.

- Because carrier screening is only intended to detect heterozygotes (carriers) and not homozygotes, false positive results would suggest that a person was a carrier of a mutation, but would not generally contain any information that could lead to conclusions of disease for the tested person. Further, no conclusion about an individual's future children could be made given the contribution of the carrier status of the child's second parent would be needed for such a conclusion.

- As stated above, false positive results obtained by one individual would have to be complemented by a partner who was also a carrier to cause a couple to consider action based upon the test result. In this scenario, the false-positive could lead to couples choosing not to get married nor have children, or could lead to unnecessary fetal testing in current or future pregnancies. Fetal testing may consist of amniocentesis or chorionic villus sampling (CVS), which pregnant women might already be scheduled for due to other risk factors such as age. Amniocentesis and CVS do carry a risk of spontaneous abortion, so they are not risk-free themselves.

- The probability of two false-positive carrier results for a couple is expected to be significantly smaller than for one false-positive.

- A false-positive result for an individual could potentially lead to adverse psychological effects, particularly if that individual didn't fully understand the nature of autosomal recessive disorders (i.e. that both the mother and father must both be carriers in order to have a 25% chance that their child would have the disorder).

Risk of False-Negative Results:

- False negative results would suggest that a person was not a carrier, but would not directly affect the health of the tested person.

- The risk of a clinical false-negative result is *already* "high" because not all clinically relevant mutations are known or tested for most diseases. The number of people who are true carriers who would be detected by any test is known as the test's "coverage". The clinical false-negative rate due to "coverage" less than 100% is likely higher than the false-negative rate from analytical failure or random error of a test.

- Current genetic testing recommendations typically recommend that initially only one member of a couple be tested for carrier status; therefore, the risks associated with false-negative results generally occur when only one member of a couple is tested and experiences a false-negative result. The risk of the false-negative would only have consequence if the non-tested partner was a carrier of the condition or disorder. In this case, there is a 25% chance that a future child would inherit the condition or disorder.

Additional Risks:
In addition to the risks associated with false-positive or negative results, we identified an additional risk associated with this device system, which is the risk that results from an incorrect interpretation of the test result. This could be characterized by an individual not understanding the nature of autosomal recessive carrier screening tests (i.e. that both the mother and father must both be carriers in order to have a 25% chance that their child would have the disorder) and making critical decisions based upon this information. The risk for this element is considered to be greater for over-the-counter devices where no healthcare professional is directly involved with test ordering or interpretation.

Special Controls:

The special controls outlined in the Order address the risks identified above:

- **Special control 1** mandates that OTC manufacturers of these tests must provide information to a potential or actual test report recipient about how to obtain access to a board-certified clinical molecular geneticist or equivalent to assist in pre and post-test counseling to a potential purchaser and actual test report recipient.
- **Special control 2** requires the use of a collection device that is FDA cleared/approved or classified as 510(k) exempt, with an indication for use in in vitro diagnostic use in DNA testing. The use of a FDA-compliant collection device provides assurances regarding safety, effective and quality of that component, which helps assure safety and effectiveness of the test system.
- **Special control 3** includes a detailed outline of clinical and analytical and performance information that must be generated and posted on the manufacturer's website. This special control also provides details on how analytical testing must be performed and provides criteria on the appropriate standard for performance for many

of these elements. This mitigates risk through lowering the probability of inaccurate test results and provides transparency on test limitations and performance.

- **Special control 4** outlines what the test should and should not be used for and mandates appropriate warning statements be included in product labeling. These controls help ensure that users have the information available to enable them to understand the limitations of the test results prior to ordering and after receiving results. The controls also provide context for the use and further interpretation of the results.
- **Special control 5** provides specific requirements for accuracy and precision/reproducibility performance characteristics. Providing such specific requirements mitigates risk through lowering the probability of inaccurate test results by ensuring a minimum level of performance.
- **Special control 6** limits the distribution of devices, excluding the collection device to the manufacturer, manufacturer's subsidiaries and laboratories subject to regulation under the Clinical Laboratory Improvement Amendments. This limitation is intended to mitigate risk through lowering the probability of inaccurate test results by ensuring that testing is performed by qualified individuals and in a manner that provides greater assurance of quality of the testing process.

R. Identified Risks and Required Mitigations Table

Identified Risks	**Required Mitigations**
Incorrect understanding of the device and test system	Special Controls (1) and (4)
Incorrect test results	Special Controls (2), (3), (5) and (6)
Incorrect interpretation of test results	Special Controls (1), (3), (4), and (5)

S. Benefit/Risk Analysis:

Summary	
Summary of the Benefit(s)	People ≥ 18 years of age may be able to determine preconception carrier status to assist in reproductive decision making.

Summary of the Risk(s)	Associated device risks include erroneous false negative results due to device or user error or false positive results due to device or user error. A person with a false positive result could have unnecessary additional testing (e.g., invasive prenatal testing with risk of miscarriage) or make an inappropriate reproductive choice (e.g., deciding against pregnancy due to unwarranted concern for having an affected child). A person with a false negative result could forego prenatal testing and/or have a child with an autosomal recessive disorder.
Conclusions Do the probable benefits outweigh the probable risks?	Based upon the nature of autosomal recessive carrier testing, which requires that both parents carry a mutation in order to have a 25% chance of bearing an affected child, and the level of scientific evidence required to establish autosomal recessive inheritance patterns which are straightforward to verify, in conjunction with the special controls drafted for this class of test systems, OIR determines that the probable benefits of Autosomal Recessive Carrier Screening Gene Mutation Detection Systems outweigh their probable risks.

T. Conclusion:

The information provided in this *de novo* submission is sufficient to classify this device into class II under regulation 21 CFR 866.5940. FDA believes that special controls, along with the applicable general controls, provide reasonable assurance of the safety and effectiveness of the device type. The device is classified under the following:

Product Code: PKB

Device Type: Autosomal Recessive Carrier Screening Gene Mutation Detection System.

Class: II (special controls).

Regulation: 21 CFR 866.5940

(a)*Identification*. Autosomal recessive carrier screening gene mutation detection system is a qualitative in vitro molecular diagnostic system used for genotyping of clinically relevant variants in genomic DNA isolated from human specimens intended for prescription use or over-the-counter use. The device is intended for autosomal recessive disease carrier screening in adults of reproductive age. The device is not intended for copy number variation, cytogenetic, or biochemical testing.

(b) Classification. Class II (special controls). Autosomal recessive carrier screening gene mutation detection system must comply with the following special controls:

(1) If the device is offered over-the-counter, the device manufacturer must provide information to a potential purchaser or actual test report recipient about how to obtain access to a board-certified clinical molecular geneticist or equivalent to assist in pre and post-test counseling.

(2) The device must use a collection device that is FDA cleared, approved, or classified as 510(k) exempt, with an indication for in vitro diagnostic use in DNA testing.

(3) The device's labeling must include a prominent hyperlink to the manufacturer's public website where the manufacturer shall make the information identified in this subsection publicly available. The manufacturer's home page, as well as the primary part of the manufacturer's website that discusses the device, must provide a prominently-placed hyperlink to the web page containing this information and must allow unrestricted viewing access. If the device can be purchased from the website or testing using the device can be ordered from the website, the same information must be found on the web page for ordering the device or provided in a prominently-placed and publicly accessible hyperlink on the web page for ordering the device. Any changes to the device that could significantly affect safety or effectiveness would require new data or information in support of such changes, which would also have to be posted on the manufacturer's website. The information must include:

(i) A detailed device description including:

(A) Gene (or list of the genes if more than one) and variants the test detects (using standardized nomenclature, Human Genome Organization (HUGO) nomenclature and coordinates).

(B) Scientifically established clinical validity of each variant detected and reported by the test, which must be well-established in peer-reviewed journal articles, authoritative summaries of the literature such as Genetics Home Reference (http://ghr.nlm.nih.gov/), GeneReviews (http://www.ncbi.nlm.nih.gov/books/NBK1116/) or similar summaries of valid scientific evidence, and/or professional society recommendations, including:

(1) Genotype-phenotype information for the reported mutations.

(2) Relevant American College of Medical Genetics (ACMG) or American Congress of Obstetricians and Gynecologists (ACOG) guideline recommending testing of the specific gene(s) and variants the test detects and recommended populations, if available. If not available, a statement stating that professional guidelines currently do not recommend testing for this specific gene(s) and variants.

(3) Table of expected prevalence of carrier status in major ethnic and racial populations and the general population.

(C) The specimen type (e.g., saliva, whole blood), matrix and volume.

(D) Assay steps and technology used.

(E) Specification of required ancillary reagents, instrumentation and equipment.

(F) Specification of the specimen collection, processing, storage and preparation methods.

(G) Specification of risk mitigation elements and description of all additional procedures, methods, and practices incorporated into the directions for use that mitigate risks associated with testing.

(H) Information pertaining to the probability of test failure (e.g., failed quality control) based on data from clinical samples, description of scenarios in which a test can fail (i.e., low sample volume, low DNA concentration, etc.), how customers will be notified and follow-up actions to be taken.

(I) Specification of the criteria for test result interpretation and reporting.

(ii) Information that demonstrates the performance characteristics of the device, including:

(A) Accuracy (method comparison) of study results for each claimed specimen type.

(1) Accuracy of the device shall be evaluated with fresh clinical specimens collected and processed in a manner consistent with the device's instructions for use. If this is impractical, fresh clinical samples may be substituted or supplemented with archived clinical samples. Archived samples shall have been collected previously in accordance with the device's instructions for use, stored appropriately and randomly selected. In some instances, use of contrived samples or human cell line samples may also be appropriate; the contrived or human cell line samples shall mimic clinical specimens as much as is feasible and provide an unbiased evaluation of the device's accuracy.

(2) Accuracy must be evaluated as compared to bidirectional sequencing or other methods identified as appropriate by FDA. Performance criteria for both the comparator method and device must be pre-defined and appropriate to the test's intended use. Detailed appropriate study protocols must be provided.

(3) Information provided shall include the number and type of specimens, broken down by clinically relevant variants, that were compared to bidirectional sequencing or other methods identified as appropriate by FDA. The accuracy, defined as positive percent agreement (PPA) and negative percent agreement (NPA), must be measured; accuracy point estimates must be > 99% (both per reported variant and overall) and uncertainty of the point estimate must be presented using the 95% confidence interval. Clinical specimens must include both homozygous wild type and heterozygous genotypes. The number of clinical specimens for each variant reported that must be included in the accuracy study must be based on the variant prevalence. Common variants (> 0.1% allele frequency in ethnically relevant population) must have at least 20 unique heterozygous clinical specimens tested. Rare variants (≤ 0.1% allele frequency in ethnically relevant population) shall have at least 3 unique mutant heterozygous specimens tested.

Any no calls (i.e., absence of a result) or invalid calls (e.g., failed quality control) in the study must be included in accuracy study results and reported separately. Variants that have a point estimate for PPA or NPA of < 99% (incorrect test results as compared to bidirectional sequencing or other methods identified as appropriate by FDA) must not be incorporated into test claims and reports. Accuracy measures generated from clinical specimens versus contrived samples or cell lines must be presented separately. Results must be summarized and presented in tabular format, by sample and by genotype. Point estimate of PPA should be calculated as the number of positive results divided by the number of specimens known to harbor variants (mutations) without "no calls" or invalid calls. The point estimate of NPA should be calculated as the number of negative results divided by the number of wild type specimens tested without "no calls" or invalid calls, for each variant that is being reported. Point estimates should be calculated along with 95% two-sided confidence intervals.

(4) Information shall be reported on the clinical positive predictive value (PPV) and negative predictive value (NPV) for carrier status (and where possible, for each variant) in each population. Specifically, to calculate PPV and NPV, estimate test coverage (TC) and the percent of persons with variant(s) included in the device among all carriers: $PPV = (PPA*TC * \pi) / (PPA*TC*\pi + (1- NPA) * (1-\pi))$ and $NPV = (NPA*(1- \pi)) / (NPA*(1-\pi) + (1- PPA*TC) * \pi)$ where PPA and NPA described either in (i) or in (ii) below and π is prevalence of carriers in the population (pre-test risk to be a carrier for the disease).

(i) For the point estimates of PPA and NPA less than 100%, use the calculated estimates in the PPV and NPV calculations.

(ii) Point estimates of 100% may have high uncertainty. If these variants are measured using highly multiplexed technology, calculate the random error rate for the overall device and incorporate that rate in the estimation of the PPA and NPA as calculated above. Then use these calculated estimates in the PPV and NPV calculations. This type of accuracy study is helpful in determining that there is no systematic error in such devices.

(B) Precision (reproducibility): Precision data must be generated using multiple instruments and multiple operators, on multiple non-consecutive days, and using multiple reagent lots. The sample panel must include specimens with claimed sample type (e.g. saliva samples) representing different genotypes (i.e., wild type, heterozygous). Performance criteria must be pre-defined. A detailed study protocol must be created in advance of the study and then followed. The "failed quality control" rate must be indicated. It must be clearly documented whether results were generated

from clinical specimens, contrived samples, or cell lines. The study results shall state, in a tabular format, the variants tested in the study and the number of replicates for each variant, and what testing conditions were studied (i.e., number of runs, days, instruments, reagent lots, operators, specimens/type, etc). The study must include all nucleic acid extraction steps from the claimed specimen type or matrix, unless a separate extraction study for the claimed sample type is performed. If the device is to be used at more than one laboratory, different laboratories must be included in the precision study (and reproducibility must be evaluated). The percentage of "no calls" or invalid calls, if any, in the study must be provided as a part of the precision (reproducibility) study results.

(C) Analytical specificity data: Data must be generated evaluating the effect on test performance of potential endogenous and exogenous interfering substances relevant to the specimen type, evaluation of cross-reactivity of known cross-reactive alleles and pseudogenes, and assessment of cross-contamination.

(D) Analytical sensitivity data: Data must be generated demonstrating the minimum amount of DNA that will enable the test to perform accurately in 95% of runs.

(E) Device stability data: The manufacturer must establish upper and lower limits of input nucleic acid and sample stability that will achieve the claimed accuracy and reproducibility. Data supporting such claims must be described.

(F) Specimen type and matrix comparison data: Specimen type and matrix comparison data must be generated if more than one specimen type or anticoagulant can be tested with the device, including failure rates for the different specimen types.

(iii) If the device is offered over-the-counter, including cases in which the test results are provided direct-to-consumer, the manufacturer must conduct a study that assesses user comprehension of the device's labeling and test process and provide a concise summary of the results of the study. The following items must be included in the user study:

(A) The test manufacturer must perform pre- and post-test user comprehension studies to assess user ability to understand the possible results of a carrier test and their clinical meaning. The comprehension test questions must directly evaluate the material being presented to the user in the test reports.

(B) The test manufacturer must provide a carrier testing education module to potential and actual test report recipients. The module must define terms that are used in the test reports and explain the significance of carrier status.

(C) The user study must meet the following criteria:

(1) The study participants must be comprised of a statistically justified and demographically diverse population (determined using methods such as quota-based sampling) that is representative of the intended user population. Furthermore, the users must be

comprised of a diverse range of age and educational levels that have no prior experience with the test or its manufacturer. These factors shall be well-defined in the inclusion and exclusion criteria.

(2) All sources of bias (e.g., non-responders) must be pre-defined and accounted for in the study results with regard to both responders and non-responders.

(3) The testing must follow a format where users have limited time to complete the studies (such as an on-site survey format and a one-time visit with a cap on the maximum amount of time that a participant has to complete the tests).

(4) Users must be randomly assigned to study arms. Test reports given to users must: a) define the condition being tested and related symptoms; b) explain the intended use and limitations of the test; c) explain the relevant ethnicities regarding the variant tested; d) explain carrier status and relevance to the user's ethnicity; e) provide links to additional information pertaining to situations where the user is concerned about their test results or would like follow-up information as indicated in test labeling). The study shall assess participants' ability to understand the following comprehension concepts: the test's limitations, purpose, and results.

(5) Study participants must be untrained, naïve to the test subject of the study and be provided only the materials that will be available to them when the test is marketed.

(6) The user comprehension study must meet the predefined primary endpoint criteria, including a minimum of a 90% or greater overall comprehension rate (i.e. selection of the correct answer) for each comprehension concept to demonstrate that the education module and test reports are adequate for over-the-counter use.

(D) A summary of the user comprehension study must be provided and include the following:

(1) Results regarding reports that are provided for each gene/variant/ethnicity tested.

(2) Statistical methods used to analyze all data sets.

(3) Completion rate, non-responder rate and reasons for non-response/data exclusion, as well as a summary table of comprehension rates regarding comprehension concepts (purpose of test, test results, test limitations, ethnicity relevance for the test results, etc.) for each study report.

(4) Your 21 CFR 809.10 compliant labeling and any test report generated must include the following warning and limitation statements, as applicable:

(i) A warning that reads "The test is intended only for autosomal recessive carrier screening in adults of reproductive age."

(ii) A statement accurately disclosing the genetic coverage of the test in lay terms, including, as applicable, information on variants not queried by the test, and the

proportion of incident disease that is not related to the gene(s) tested. For example, where applicable, the statement would have to include a warning that the test does not or may not detect all genetic variants related to the genetic disease, and that the absence of a variant tested does not rule out the presence of other genetic variants that may be disease related. Or, where applicable, the statement would have to include a warning that the basis for the disease for which the genetic carrier status is being tested is unknown or believed to be non-heritable in a substantial number of people who have the disease, and that a negative test result cannot rule out the possibility that any offspring may be affected with the disease. The statement would have to include any other warnings needed to accurately convey to consumers the degree to which the test is informative for carrier status.

(iii) For prescription use tests, the following warnings that read:

(A) "The results of this test are intended to be interpreted by a board-certified clinical molecular geneticist or equivalent and should be used in conjunction with other available laboratory and clinical information."

(B) "This device is not intended for disease diagnosis, prenatal testing of fetuses, risk assessment, prognosis or pre-symptomatic testing, susceptibility testing, or newborn screening."

(iv) For over-the-counter tests, a statement that reads "This test is not intended to diagnose a disease, or tell you anything about your risk for developing a disease in the future. On its own, this test is also not intended to tell you anything about the health of your fetus, or your newborn child's risk of developing a particular disease later on in life."

(v) For over-the-counter tests, the following warnings that read:

(A) "This test is not a substitute for visits to a healthcare provider. It is recommended that you consult with a healthcare provider if you have any questions or concerns about your results."

(B) "The test does not diagnose any health conditions. Results should be used along with other clinical information for any medical purposes."

(C) "The laboratory may not be able to process your sample. The probability that the laboratory cannot process your saliva sample can be up to [actual probability percentage]."

(D) "Your ethnicity may affect how your genetic health results are interpreted."

(vi) For a positive result in an over-the-counter test when the positive predictive value for a specific population is less than 50% and more than 5%, a warning that reads "The positive result you obtained may falsely identify you as a carrier. Consider genetic counseling and follow-up testing."

(vii) For a positive result in an over-the-counter test when the positive predictive value for a specific population is less than 5%, a warning that reads "The positive result you obtained is very likely to be incorrect due to the rarity of this variant. Consider genetic counseling and follow-up testing."

(5) The testing done to comply with subparagraph (b)(3) must show the device meets or exceeds each of the following performance specifications:

(i) The accuracy must be shown to be equal to or greater than 99 percent for both positive percent agreement (PPA) and negative percent agreement (NPA). Variants that have a point estimate for PPA or NPA of <99% (incorrect test results as compared to bidirectional sequencing or other methods identified as appropriate by FDA) must not be incorporated into test claims and reports.

(ii) Precision (reproducibility) performance must meet or exceed 99% for both positive and negative results.

(iii) The user comprehension study must obtain values of 90% or greater user comprehension for each comprehension concept.

(6) The distribution of this device, excluding the collection device described in subparagraph (b)(2), shall be limited to the manufacturer, the manufacturer's subsidiaries, and laboratories regulated under the Clinical Laboratory Improvement Amendments.