

# The Integrity and Yield of Genomic DNA Isolated from Whole Blood Following Long-Term Storage at $-30^{\circ}\text{C}$

Wenlong Carl Chen,<sup>1-3</sup> Robyn Kerr,<sup>3,4</sup> Andrew May,<sup>3,4</sup> Babongile Ndlovu,<sup>1</sup>  
Anelisa Sobalisa,<sup>1,4</sup> Sanelisiwe T. Duze,<sup>1,5</sup> Lavana Joseph,<sup>1</sup>  
Christopher G. Mathew,<sup>2,6</sup> and Chantal Babb de Villiers<sup>3,7</sup>

Long-term storage of whole blood can affect the integrity of DNA if it is not done under optimal conditions. The aim of this study was to determine whether long-term storage (2–19 years) of whole blood samples at  $-30^{\circ}\text{C}$  had a negative effect on the quality or quantity of genomic DNA that could be recovered at extraction. Genomic DNA was isolated from 2758 whole blood samples collected in 4 mL EDTA vacutainers from 1997 to 2012. DNA was extracted using the Qiagen<sup>®</sup> FlexiGene<sup>®</sup> DNA kit. The average storage duration at  $-30^{\circ}\text{C}$  was 12 years. The quality and quantity of the isolated DNA were assessed using spectrophotometry (NanoDrop<sup>™</sup>), a fluorometric assay for double-stranded DNA (Qubit<sup>™</sup>), and agarose gel electrophoresis. The mean DNA yield per sample was found to be 114  $\mu\text{g}$  from whole blood volumes that ranged from 0.5 to 4 mL. The mean A260/280 ratio and median A260/280 ratios were both 1.8. No correlation was found between the duration of storage and the total yield or the quality of DNA extracted. These data suggest that high-quality DNA can be extracted from whole blood samples that are stored at  $-30^{\circ}\text{C}$  for up to 19 years.

**Keywords:** DNA storage, long-term storage, DNA quality

## Introduction

DNA CAN BE isolated from almost all body tissues, including bodily fluids.<sup>1</sup> Due to its relative ease of acquisition, whole blood provides one of the most common sources of high-quantity and high-quality DNA for molecular applications, including clinical and epidemiological studies.<sup>2</sup>

Large genetic epidemiological studies can involve the collection of tens of thousands of valuable whole blood samples that can be, or have been, stored for many years. The blood volume available for DNA extractions is an obvious indicator of DNA yield.<sup>3</sup> However, long-term storage conditions of whole blood can affect the integrity of DNA.<sup>4</sup> Blood samples stored at  $4^{\circ}\text{C}$  for a short period of time will still yield DNA of acceptable quality provided the correct blood collection tubes were used.<sup>1,5</sup> However, the desired temperature for whole blood samples for DNA isolation is at  $-80^{\circ}\text{C}$  for long-term storage.<sup>4</sup> Whole blood samples can also be frozen at  $-20^{\circ}\text{C}$  for long-term storage,<sup>4,6,7</sup>

however, other studies have found lower DNA yields following this approach.<sup>1,7</sup> A study by Di Pietro et al. found that whole blood samples stored for long periods at  $-20^{\circ}\text{C}$  yielded high-quality DNA for genotyping studies, although the sample size was small ( $n=82$ ).<sup>8</sup>

The current study sought to determine whether the duration of long-term storage of whole blood samples, using a large sample size ( $n=2758$ ), stored at a stable  $-30^{\circ}\text{C}$ , has a negative effect on the quality and quantity of the genomic DNA extracted.

## Materials and Methods

### Samples

Samples analyzed were sourced from the Johannesburg Cancer Case–Control Study (JCS) at the National Cancer Registry (NCR) of South Africa. The JCS is a large, recently (2016) ended case–control study established in 1995.<sup>9</sup> In

<sup>1</sup>National Cancer Registry, National Health Laboratory Service, Johannesburg, South Africa.

<sup>2</sup>Sydney Brenner Institute for Molecular Bioscience, University of the Witwatersrand, Johannesburg, South Africa.

<sup>3</sup>Division of Human Genetics, School of Pathology, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa.

<sup>4</sup>Division of Human Genetics, National Health Laboratory Service, Johannesburg, South Africa.

<sup>5</sup>Centre for Enteric Diseases, National Institute for Communicable Diseases, Johannesburg, South Africa.

<sup>6</sup>Division of Medical Biochemistry, University of Cape Town, Cape Town, South Africa.

<sup>7</sup>Primary Care Unit, Department of Public Health and Primary Care, University of Cambridge, Cambridge, United Kingdom.

brief, the JCS recruited newly diagnosed cancer patients of black African ancestry. These patients were interviewed by trained research nurses and their blood was collected with informed consent for future research as approved by the Research Ethics Committee at the University of the Witwatersrand and described by Urban et al.<sup>9</sup> Blood samples were drawn from patients in 4 mL EDTA vacutainers and were stored in  $-30^{\circ}\text{C}$  freezers until further use. More than 20,000 whole blood samples (representing various cancer types) have been collected since March 1995 for the JCS, of which a subset of 2758 blood samples were used for this study. DNA extractions were performed on request for specific studies in collaboration with the JCS. At the time of the study being reported here, DNA had been extracted from 2758 blood samples for four different studies. Ethical approval was obtained from the University of the Witwatersrand Human Research Ethics Committee (Medical), clearance numbers: M140271 and M120117.

### Sample storage

All blood samples collected under the JCS were kept at room temperature until collected for transport. The blood samples were then transported in a cooler box with an ice block from the three separate study sites to a centralized storage facility. On the rare occasion that there was an issue with the transportation, samples were frozen at the collection site and collected at the next available time. Average transit time was  $\leq 30$  minutes. All samples were frozen in their collection tubes in an upright position at  $-30^{\circ}\text{C}$  within a maximum of 24 hours of collection. Miele™  $-30^{\circ}\text{C}$  general freezers were used for sample storage. The freezers' temperatures were monitored with an alarm-based system and backup freezers were available. As part of our risk management strategy, when a freezer failed, the samples were transferred to another freezer before thawing. All the blood samples were frozen in an upright position and thereafter stored on their sides. Although the blood was collected in a 4 mL EDTA tube, not all were filled to capacity; therefore, the volume of the blood sample available for extraction was noted before DNA extraction. Sample information was managed using Microsoft Excel and then incorporated into STATA database.

### DNA isolation

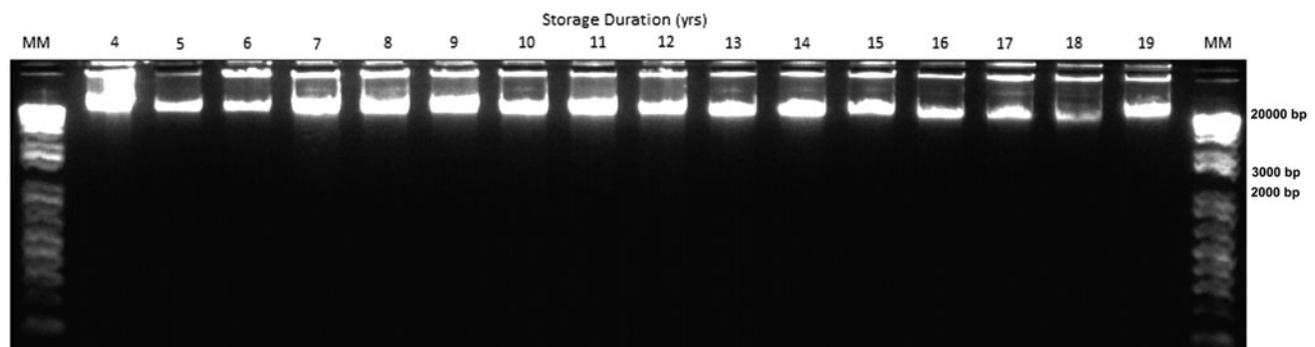
Whole blood samples collected between 1997 and 2012 ( $n = 2758$ ) were extracted. DNA extraction procedures occurred

in batches between May 2013 and June 2016. Six different technicians were involved over the 4-year period. The technicians were trained on the extraction protocol, blinded to the storage duration of the blood samples, and each technician extracted samples collected over the whole time period. Total genomic DNA was isolated using the FlexiGene® DNA Kit (Qiagen®). This kit was chosen initially for its ability to handle a diverse range of input blood volume and then used subsequently for consistency. Extractions were undertaken following the manufacturer's protocol, with the following minor modifications: the use of a dry heat incubator instead of a hot water bath, and resuspension of the DNA pellet in  $1\times$  Tris-EDTA buffer (Invitrogen™) instead of the supplied FG3 buffer. To ensure no wastage and optimum yield, the entire blood volume in a tube was used for DNA extraction. Two fixed protocols, one for samples with  $<2$  mL of blood and the other for  $>2$  mL, were optimized and used as per the manufacturer's procedures. The most noticeable difference between protocols was the use of 50 mL Nunc tubes for samples with blood volumes of  $>2$  mL, rather than the 15 mL Nunc tubes used for samples with blood volumes of  $<2$  mL. Reagent volumes were adjusted as per the sample volume (as per the manufacturer's protocol). The centrifugation conditions were unchanged. No RNase digestion was incorporated into the DNA isolation protocol.

### DNA quantity and quality

DNA quantity and purity were assessed spectrophotometrically using the NanoDrop™ 1000 (Thermo Scientific™) directly after extraction and before storage at  $-30^{\circ}\text{C}$ . DNA concentration was measured and DNA purity was calculated through the standard A260/280 and A260/230 ratios. This is a fast and accurate technique for determining DNA concentration of pure samples.<sup>10</sup> DNA/RNA and proteins have their maximum absorbance at 260 and 280 nm, respectively. An A260/280 ratio of  $\sim 1.8$  is generally accepted as a reflection of a pure DNA sample and a ratio of  $\sim 2.0$  is generally accepted as pure for RNA. A secondary measurement of the A260/230 ratio is used to determine the purity of DNA against other contaminants. A higher A260/230 is expected for pure DNA ( $\sim 2.0$ – $2.2$ ). Low A260/230 ratios indicate the presence of contamination or proteins that absorb light at the 230 nm wavelength.<sup>11</sup>

Genomic DNA integrity assessment was done, at a later stage, using gel electrophoresis for 218 samples selected at



**FIG. 1.** Gel image showing integrity of genomic DNA from 4 to 19 years of blood sample storage. One representative sample was used from each year of storage (4–19 years). Two microliters of genomic DNA was resolved on 1.0% agarose gel with GelRed™ staining. GeneRuler™ 1 kb plus DNA Ladder MM was used for sizing reference. Electrophoresis was done at 8 V/cm for 45 minutes. MM, molecular marker.

random (4–19 years storage, mean=12 years) (Fig. 1). Genomic DNA was resolved on 0.8% agarose gel, stained with GelRed™, at 8 V/cm for 45 minutes against the GeneRuler 1 kb Plus DNA Ladder (Thermo Scientific).

DNA quantitation was additionally assessed using a second methodology. Qubit fluorometry was performed using the Qubit™ 3.0 Fluorometer (Thermo Scientific). This method uses the double-stranded DNA (dsDNA) broad range assay to quantify dsDNA, as opposed to single-stranded DNA (ssDNA), present in a sample. Only a subset of samples ( $n=200$ ) was quantified using this chemistry owing to cost restraints. The 200 samples were randomly selected for year of collection (2–19 years of storage), with concentrations representative of the overall DNA concentration range. This measures dsDNA concentration only, which provides a more useful reading as contaminants, including degraded DNA and RNA, are not measured. The percentage of dsDNA (%dsDNA) was calculated by dividing the Qubit concentration by the NanoDrop concentration and multiplying by 100.

### Analysis

Descriptive statistics were compiled by years of storage. Means and medians were calculated for DNA yields, DNA concentrations, blood volume available per sample, DNA yield per milliliter of blood, A260/280 and A260/230. Both means and medians were compiled given the non-normally distributed nature of the data, which may limit the mean accuracy.

Spearman's ranked correlations were performed to determine whether duration of storage (years) of the samples, blood volume available for extraction, and the age of participants correlated with the total DNA yields. Samples ( $n=564$ ) without blood volume data were excluded.

Kruskal–Wallis  $H$  tests, with multiple group comparisons, were performed to determine whether DNA yields, A260/280 and A260/230 ratios, differed between the different years of storage duration. Kruskal–Wallis  $H$  tests were also used to determine differences in mean DNA yield per milliliter of blood across different storage durations as well as to determine differences in mean total DNA yields measured by Qubit and differences in mean %dsDNA across different storage durations. A nonparametric test for trend was performed to evaluate the relationship between %dsDNAs across the storage duration of the blood samples. Quantile regression estimating the conditional median was performed to model the effect of storage duration of the blood samples, blood volume available for DNA extraction, and age and gender of the participants had on the %dsDNA.

A two-sample  $t$ -test with equal variances was performed to assess the difference in mean DNA yield per 1 mL of blood between the two different DNA extraction protocols as well as mean DNA yield per 1 mL of blood between cancer samples and noncancer samples and between HIV-positive and HIV-negative samples.

### Results

Total genomic DNA was isolated from a total of 2758 samples collected between 1997 and 2012 for the JCS (Table 1). NanoDrop data were used to calculate DNA yield, DNA concentration, and to determine A260/280 and

A260/230 ratios. Blood samples were stored on average for 11.9 years. The longest storage period of samples before DNA extraction was 19 years ( $n=125$ ) and the shortest was 2 years ( $n=31$ ) (Table 1). Information on the volume of blood collected was available for 80% ( $n=2206$ ) of all samples (Table 1). Freezers were monitored closely so that no thawing of samples (resulting from freezer failures) happened during this study.

The quality (A260/280 and A260/230 ratios), concentration (ng/ $\mu$ L), and the quantity of isolated DNA ( $\mu$ g), and the mean and median concentration per year of storage for the samples are shown in Table 1. The mean A260/280 ratio across all samples was 1.79 and median A260/280 ratio was 1.8, suggesting that good-quality DNA was isolated. The mean A260/230 ratio across all samples was 1.71 and the median A260/230 ratio was 1.78,  $<2.0$ , suggesting the presence of proteins, contaminants, or RNA residues that absorb at 230 nm or less.

Genomic DNA from 218 samples was subjected to agarose gel electrophoresis to assess possible degradation. High-molecular-weight staining was observed for all 218 samples. Figure 1 shows a representation of the gel electrophoresis result for a representative subset of samples, which had been stored for different periods of time. A consistent low level of partial degradation (low intensity of smearing from the genomic band down to 3 kb) was noted across most samples analyzed on gel electrophoresis. The consistency of this smearing pattern suggests that this is likely an artifact from the DNA isolation protocol that caused DNA smearing rather than storage-related DNA degradation.

The DNA yields of samples per year of storage clustered tightly together, except for a few outlier samples with very high DNA yields or extremely low DNA yield (Fig. 2). The presence of the outliers illustrates the value of the median as a more informative reporting tool than the mean.

The average yield for all samples was 108.95  $\mu$ g as measured by NanoDrop, whereas the average yield for the 200 samples measured by Qubit was 86.19  $\mu$ g per sample. This indicates that  $\sim 79\%$  of total DNA measured by spectrophotometry was dsDNA, with the remaining proportion presumed to be ssDNA and/or denatured or degraded DNA. A plot of dsDNA percentage against the storage duration is shown in Figure 3.

A Kruskal–Wallis  $H$  test showed that no statistically significant differences exist between the mean total DNA yield by Qubit measurements and the number of years of storage ( $p=0.46$ ) and no statistically significant differences exist between the mean %dsDNA and the number of years of storage ( $p=0.10$ ). Nonparametric trend analysis for %dsDNA, quantified by Qubit across the storage durations, indicated a positive trend ( $z=6.24$ ,  $p<0.001$ ), with the %dsDNA increasing with increased storage duration of the blood samples. However, storage duration of the blood samples had a negligible effect on %dsDNA after adjusting for blood volume used for extraction, age, and gender ( $\beta=0.01$ ,  $p=0.045$ ).

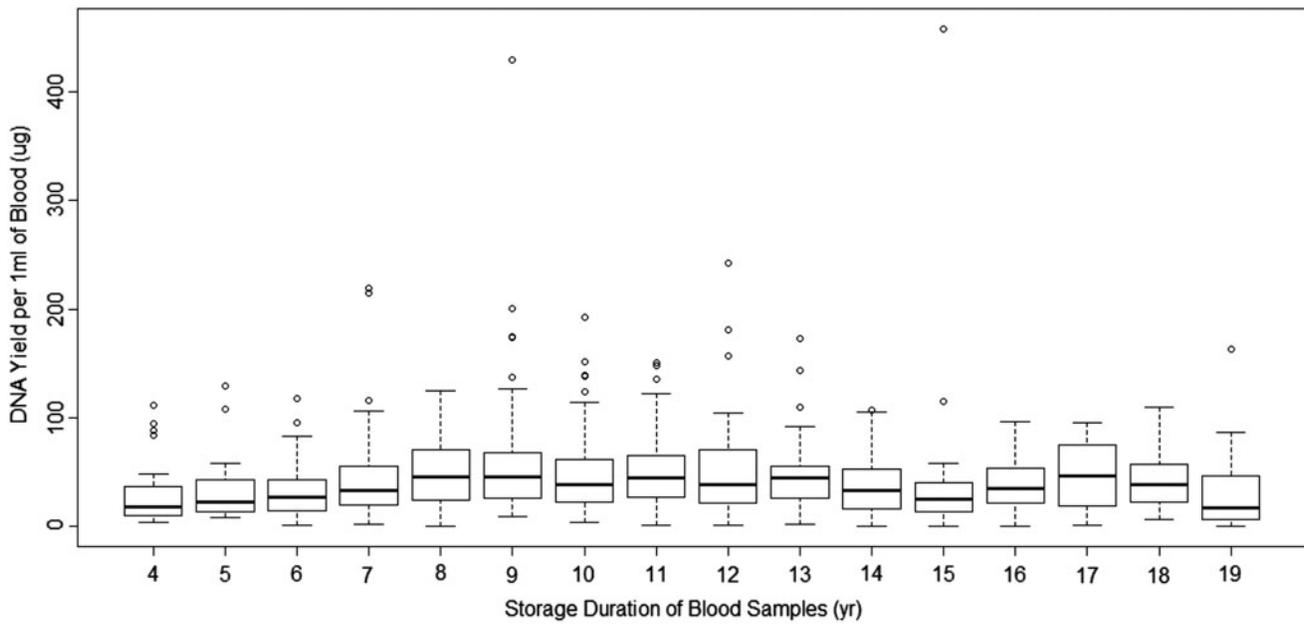
Spearman's ranked correlation was performed to assess the relationship between total DNA yields, storage duration of blood samples, blood volumes available for extractions, and age of participants at recruitment using 2194 samples. After Bonferroni correction, there was a strong positive correlation observed between the total blood volumes

TABLE 1. DESCRIPTIVE STATISTICS FOR THE SAMPLES USED FOR GENOMIC DNA ISOLATION PER YEARS OF STORAGE

Storage duration (years)	Sample size	Mean blood volume <sup>a</sup> (mL)	A260/280			A260/230			Mean DNA concentration (ng/ $\mu$ L) per sample <sup>b</sup>	Median DNA yield ( $\mu$ g) per sample <sup>b</sup>	Mean DNA yield/mL of blood ( $\mu$ g) <sup>b</sup>	Mean DNA yield ( $\mu$ g) per sample <sup>b</sup>	
			Median	Range	SD	Mean	Median	Range					SD
2	31	—	1.81	1.68–1.88	0.03	1.81	1.69	1.24–2.04	0.18	1.69	305.97	110.88	122.39
3	25	—	1.80	1.70–1.90	0.05	1.82	1.60	0.29–2.15	0.45	1.56	266.66	98.46	106.66
4	54	2.1	1.74	1.52–1.90	0.08	1.73	1.70	0.76–2.13	0.30	1.63	217.16	62.95	86.86
5	84	2.6	1.79	1.24–1.91	0.08	1.76	1.80	0.36–2.31	0.39	1.69	224.90	59.10	89.96
6	102	2.6	1.79	1.51–1.95	0.06	1.78	1.88	0.91–2.26	0.34	1.80	195.36	61.60	78.14
7	212	3.0	1.80	1.51–2.25	0.06	1.81	1.79	0.29–2.58	0.37	1.73	266.64	96.05	106.66
8	185	3.3	1.81	1.27–1.97	0.06	1.81	1.66	0.26–2.78	0.39	1.68	278.78	102.72	111.51
9	198	3.1	1.81	1.58–2.18	0.07	1.81	1.66	0.81–2.56	0.42	1.68	285.70	97.68	114.28
10	218	3.1	1.80	1.62–2.05	0.06	1.81	1.65	0.57–2.41	0.40	1.66	293.74	97.21	117.50
11	203	3.1	1.80	1.48–2.15	0.07	1.80	1.69	0.30–2.62	0.46	1.68	289.54	103.96	115.82
12	194	3.1	1.80	1.26–2.31	0.09	1.80	1.64	0.35–2.34	0.41	1.60	291.38	96.57	116.55
13	172	3.1	1.80	0.71–2.09	0.11	1.78	1.63	0.19–2.60	0.44	1.62	316.77	100.73	126.71
14	210	2.9	1.80	0.84–2.18	0.10	1.79	1.69	0.23–2.43	0.42	1.61	294.15	101.16	117.66
15	113	2.9	1.80	0.32–1.89	0.16	1.77	1.95	0.22–2.43	0.43	1.85	281.76	88.59	112.71
16	339	3.1	1.79	1.02–3.30	0.11	1.79	1.78	0.22–2.44	0.41	1.71	353.03	109.14	141.21
17	103	3.0	1.80	0.24–1.91	0.16	1.78	2.07	0.51–2.44	0.33	1.98	252.63	78.42	101.05
18	192	3.1	1.80	1.10–1.96	0.06	1.80	2.08	0.81–2.51	0.25	2.04	273.24	91.02	109.30
19	124	2.9	1.80	0.56–1.96	0.15	1.77	1.81	0.08–2.49	0.55	1.67	235.57	75.86	94.23

<sup>a</sup>Mean blood volume per sample.

<sup>b</sup>Mean DNA concentrations and median and mean DNA yields are calculated using NanoDrop™ data. SD, standard deviation.



**FIG. 2.** Range and distribution of DNA yields per 1 mL of blood extracted from frozen whole blood stored over 4–19 years. Blood volume data were not available for 2 and 3 years of storage. No statistically significant difference in yield was observed across the duration of blood storage. DNA yields were measured by NanoDrop.<sup>TM</sup>

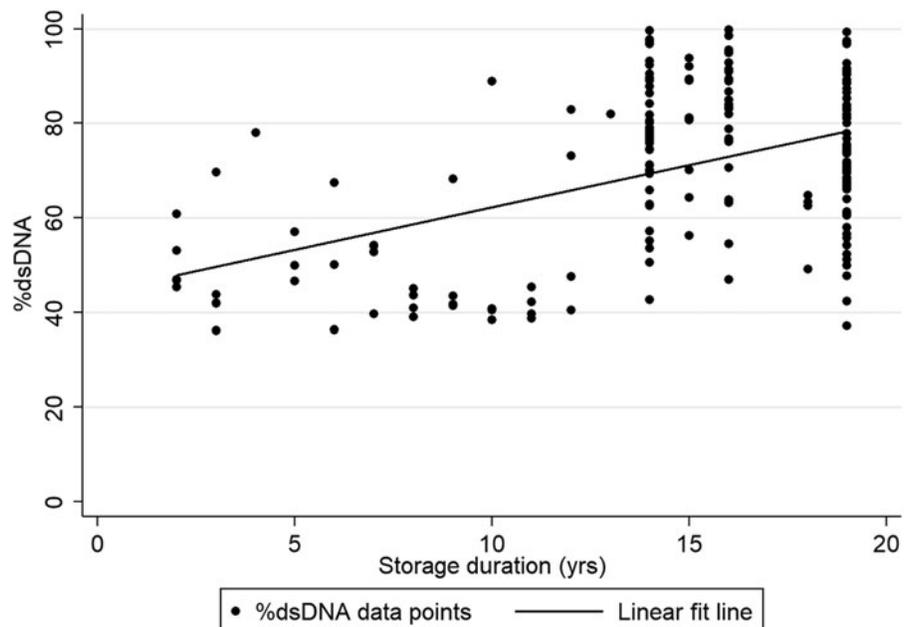
available for extractions and the total DNA yields, which was statistically significant,  $r_s = 0.2258$ ,  $p < 0.0001$ . So, the more total volume of blood available the more total DNA yield there was. There was no evidence of correlation between storage duration of blood samples or age of participants at recruitment with total DNA yields. There was no evidence of correlation between total blood volumes with A260/280 or A260/230 ratios.

A Kruskal–Wallis  $H$  test, after Bonferroni correction and correcting for technician variability, showed that no statistically significant differences exist between the DNA yields, A260/280 or A260/230 ratios, and the number of years of storage. As expected, total blood volume is a direct indicator of

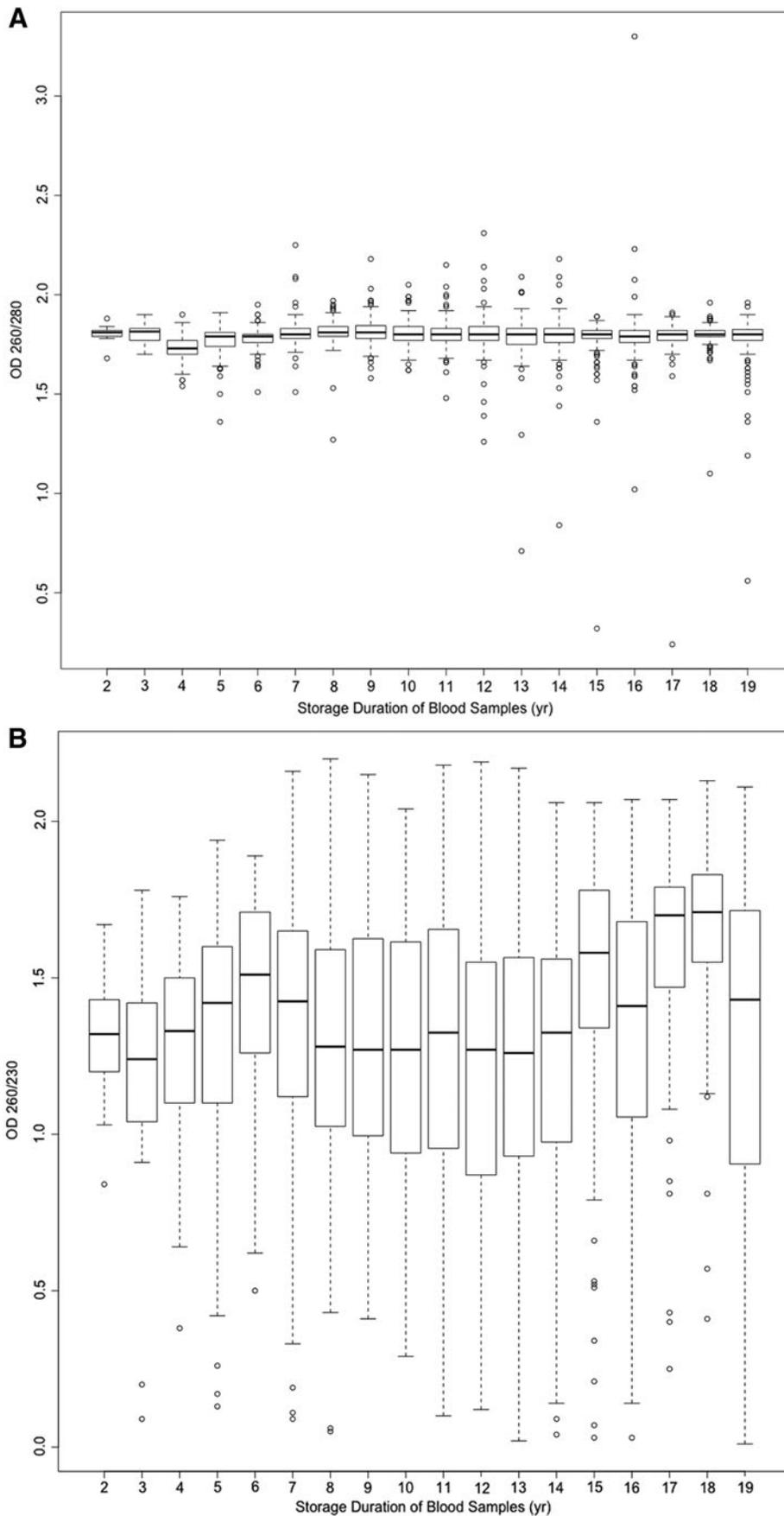
the amount of total DNA available ( $r_s = 0.2258$ ,  $p < 0.0001$ ). Similarly, no statistically significant differences were observed between the mean DNA yield per milliliter of blood and the storage duration of the blood sample ( $p = 0.25$ ).

The FlexiGene DNA Kit (Qiagen) gave an expected DNA yield of 37.5  $\mu\text{g}$  per 1 mL of blood assuming a normal white cell count of  $7 \times 10^6$  cells/mL blood.<sup>12</sup> The JCS sample yield was between 25.83 and 46.54  $\mu\text{g}$  (mean = 37.77  $\mu\text{g}$ ) of DNA per 1 mL of blood.

The effect of the optimized protocols for different sample blood volumes was assessed using a two-sample  $t$ -test with equal variances. Samples with blood volumes of 2 mL or less ( $n = 403$ ) yielded more DNA per 1 mL of blood



**FIG. 3.** %dsDNA versus storage duration. The linear fit line indicates that a positive trend exists between %dsDNA and the increase of storage duration. %dsDNA, percentage of double-stranded DNA.



**FIG. 4.** (A) A260/280 ratio distribution of all samples across duration of blood sample storage. A260/280 ratios were measured by NanoDrop. (B) A260/230 ratio distribution of all samples across duration of blood sample storage. A260/230 ratios were measured by NanoDrop.

(44.12  $\mu\text{g}$ , 95% CI: 39.73–48.51  $\mu\text{g}$ ) when compared to DNA yielded per 1 mL of blood for samples with blood volumes of  $>2\text{ mL}$  ( $n=1792$ ) (36.79  $\mu\text{g}$ , 95% CI: 35.42–38.16  $\mu\text{g}$ ) ( $p=0.0001$ ). The typical DNA yields of the FlexiGene kit for a 2 mL volume would be 33–45  $\mu\text{g}/\text{mL}$ . No difference in mean DNA yield per 1 mL of blood was observed between cancer samples ( $n=1482$ ) and noncancer samples ( $n=707$ ) ( $p=0.50$ ). A difference in mean DNA yield per 1 mL of blood was observed between HIV-positive ( $n=1851$ ) and HIV-negative samples ( $n=313$ ), with HIV-negative samples (38.70  $\mu\text{g}$ , 95% CI: 37.23–40.17  $\mu\text{g}$ ) yielding  $\sim 5\ \mu\text{g}$  DNA more per milliliter than HIV-positive samples (33.36  $\mu\text{g}$ , 95% CI: 29.39–37.32  $\mu\text{g}$ ) ( $p=0.0076$ ).

## Discussion

This study was undertaken to answer questions about the expected quality and quantity of DNA isolated from whole blood samples that have been stored at  $-30^\circ\text{C}$  between 2 and 19 years. The concern was that long-term storage of whole blood at  $-30^\circ\text{C}$  may have a negative effect on the quality and/or quantity of the DNA.<sup>7</sup> Genomic DNA was isolated from 2758 blood samples collected between 1997 and 2012 using the Qiagen FlexiGene DNA kit. To our knowledge this is the largest study of its kind assessing the quality and quantity of DNA isolated from frozen blood samples stored for long periods.

The results from our study are unique from previous findings<sup>1,7</sup> as we used a different DNA extraction method (kit vs. salting out or phenol/chloroform), stored samples at  $-30^\circ\text{C}$  temperature (compared to room temperature and  $4^\circ\text{C}$ ), and used a much larger sample size. The large sample size of our study affords greater statistical power to evaluate the relationship between DNA yields and storage duration of the blood samples. A recent publication looking at various storage conditions on forensic examinations of blood samples and blood stains ( $n=6$ ) concluded that to prevent DNA degradation during long-term (20 year) storage, they should be stored at below  $-20^\circ\text{C}$ ,<sup>13</sup> as our samples have been. We have a much larger sample size with varying lengths of storage to assess and have found that the quantity of DNA is not affected, and are fairly confident that the same can be said about the quality.

This study found that long-term storage of whole blood samples at  $-30^\circ\text{C}$  does not have a negative effect on the quality and quantity of the DNA extracted. The average duration of storage was  $\sim 12$  years. The mean A260/280 ratios were found to be  $\sim 1.8$  across all the years of collection, suggesting high purity of the DNA regardless of the duration of blood storage (Fig. 4A). The lower than expected A260/230 ratios indicate the presence of other organic residues (Fig. 4B). The source of the organic residues is likely artifacts resulting from the inadequate removal of organic chemicals used in the extraction chemistry.

The DNA yield was expected to decrease for samples stored for longer periods. Alrokayan reported a 30%–40% decrease in total DNA yield for whole blood samples frozen at  $-20^\circ\text{C}$ .<sup>7</sup> A decrease in DNA yield was not observed in the JCS samples extracted, as no correlation between storage duration and total DNA yields was observed. Assuming the blood samples are stored within the desired temperature range at stable conditions, DNA can be extracted from these samples after prolonged storage. DNA yields of samples per year of collection clustered tightly together for the 2–19 years, except for a few outliers (Fig. 2).

High-quality dsDNA is preferred for downstream applications such as polymerase chain reaction (PCR), microarray analysis, or next-generation sequencing analysis. Qubit is well known to be a more sensitive method to quantify DNA, compared to the NanoDrop.<sup>14,15</sup> Qubit readings are lower than NanoDrop readings as they do not measure single-stranded, degraded DNA.<sup>14,15</sup> Our results showed that, on average, 79% of total DNA measured by NanoDrop was dsDNA. From the subset of samples analyzed on Qubit, the %dsDNA did not decrease with increased years of storage ( $p=0.10$ ); instead, a positive trend was seen with increasing storage duration but the effect size was minimal ( $\beta=0.01$ ,  $p=0.045$ ). This supports our hypothesis that storage duration of the blood samples does not influence DNA quality. It is worth noting that although DNA was extracted from these archived blood samples, no assessments were done for RNA, micro-RNA, cell-free circulating DNA, or proteins.

The DNA yields are in line with expected yields cited by the manufacturers. With every additional 1 mL of blood volume available for extraction, the DNA yield would be expected to increase by 20.94  $\mu\text{g}$  ( $p<0.01$ ). The DNA yield per milliliter of blood does not increase when more blood was processed. Rather, optimization of the DNA extraction protocol used for different volumes of blood affects the DNA yield per milliliter of blood. Two fixed protocols were used in this study, the protocol for sample volumes of  $<2\text{ mL}$  of blood yielded  $\sim 7\ \mu\text{g}$  ( $p=0.0001$ ) more DNA per milliliter of blood than the protocol used for samples of  $>2\text{ mL}$  of blood. This is possibly due to the extra time and care given per extraction for the small volumes of these samples because of the limited volume. It is possible that the potential yield could be improved by optimizing the extraction protocol to the exact blood volume available. This would, however, require extra laboratory time and staff.

The patients' age may also influence the number of white cells available.<sup>5,16</sup> In older patients there is often a decline in the quantity of DNA due to a reduction in the number of leukocytes and lymphocytes with age progression.<sup>17</sup> However, this correlation was not observed in this study. As anticipated, samples from individuals who were HIV positive at the time of sampling had a lower DNA yield as a likely consequence of fewer peripheral blood mononuclear cells due to infection. However, no difference in mean yield was observed between cancer samples and noncancer samples.

One limitation of this study was that gel electrophoresis and fluorometry were not done for all 2758 samples to assess the integrity of the extracted DNA at the beginning of DNA storage. Hence the data lack robustness to definitively conclude that long-term storage of blood samples does not negatively affect DNA integrity. However, from the representative subset of samples ( $n=200$ ) that were assessed by gel electrophoresis, the authors are confident that the DNA integrity of DNA extracted from long-term stored blood samples is sound. A comprehensive analysis would include doing a quantitative PCR or a long-range PCR and we have applied for funding to do this. We have used these samples in other genotyping studies and they have produced adequate genotyping call rates (unpublished).

From the results of this large study, the authors are confident that long-term storage (average 11.9 years) of whole blood samples stored in 4 mL EDTA vacutainers at a stable  $-30^\circ\text{C}$  does not have a negative effect on the quality and quantity of the DNA. The authors are also confident that the DNA integrity assessment done on a representative sample

is sufficient to represent the larger cohort. Despite the limitations of this study, we found no evidence that DNA integrity worsened over a prolonged period of blood sample storage. Archived whole blood samples can still be utilized in genetic epidemiology studies provided the blood samples are stored correctly with constant temperature monitoring to prevent freeze/thaw episodes.

### Acknowledgments

The authors acknowledge Sisters Gloria Mokwatle, Patricia Rapoho, and Pheladi Kale who carried out the interviews and collected blood specimens from the patients included in this analysis, Mrs. Lettie Bester who prepared the specimens for testing and storage, and Mrs. Margaret Urban and Mr. Tonicah Maphanga for maintaining the JCS database. They thank Dr. Mazvita Sengayi for statistical analysis support and Ms. Cassandra Soo and Ms. Natalie Smyth for fluorescent DNA quantification. They also thank the oncology clinicians and the administration at Charlotte Maxeke Johannesburg Academic Hospital (formerly Johannesburg General Hospital) for assistance and for access to patients in their care, and the patients who gave freely of their time. Research reported in this publication and ongoing extraction of DNA from JCS blood samples are supported by the Grants, Innovation and Product Development (GIPD) Unit of the South African Medical Research Council with funds received from the South African National Department of Health and the U.K. Medical Research Council from the U.K. Government's Newton Fund. The authors thank the DST/NRF internship program for supporting B.N., A.S., S.T.D., and L.J. while interns at CERG. They thank the DAAD/NRF in country scholarship for a fellowship to W.C.C., and the Cancer Association of South Africa (CANSA) for research cost support to C.G.M. The contents are solely the responsibility of the authors and do not necessarily reflect the views of the funding bodies.

### Authors' Contributions

W.C.C.: substantial contributions to conception and design, acquisition of data (DNA extraction), analysis and interpretation of data, and final approval of the version to be published. R.K.: interpretation of data, revising the article critically for important intellectual content, and final approval of the version to be published. A.M.: analysis and interpretation of data, revising the article critically for important intellectual content, and final approval of the version to be published. B.N., A.S., S.T.D., and L.J.: acquisition of data (DNA extraction), maintenance of the database and summary report (base of manuscript), and final approval of the version to be published. C.G.M.: interpretation of data, revising the article critically for important intellectual content, and final approval of the version to be published. C.B.d.V. conceived the project, substantial contributions to conception and design, interpretation of data, revising the article critically for important intellectual content, and final approval of the version to be published.

### Author Disclosure Statement

No conflicting financial interests exist.

### References

1. Nederhand RJ, Droog S, Klufft C, et al. Logistics and quality control for DNA sampling in large multicenter studies. *J Thromb Haemost* 2003;1:987–991.
2. Miller SA, Dykes DD, Polesky HF. A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res* 1988;16:1215.
3. Chacon-Cortes D, Griffiths LR. Methods for extracting genomic DNA from whole blood samples: Current perspectives. *J Biorepos Sci Appl Med* 2014;2:1–9.
4. Holland NT, Smith MT, Eskenazi B, et al. Biological sample collection and processing for molecular epidemiological studies. *Mutat Res* 2003;543:217–234.
5. Richardson AJ, Narendran N, Guymer RH, et al. Blood storage at 4 degrees celsius—Factors involved in DNA yield and quality. *J Lab Clin Med* 2006;147:290–294.
6. Steinberg KK, Sanderlin KC, Ou CY, et al. DNA banking in epidemiologic studies. *Epidemiol Rev* 1997;19:156–162.
7. Alrokayan SAH. Effect of storage temperature on the quality and quantity of DNA extracted from blood. *J Biol Sci* 2000;3:392–394.
8. Di Pietro F, Ortenzi F, Tilio M, et al. Genomix DNA extraction from whole blood stored from 15- to 30-years at  $-20^{\circ}\text{C}$  by rapid phenol-chloroform protocol: A useful tool for genetic epidemiology studies. *Mol Cell Probes* 2011;25:44–48.
9. Urban M, Banks E, Egger S, et al. Injectable and oral contraceptive use and cancers of the breast, cervix, ovary, and endometrium in Black South African women: Case-control study. *PLoS Med* 2012;9:e1001182.
10. Desjardins P, Conklin D. NanoDrop microvolume quantitation of nucleic acids. *J Vis Exp* 2010. DOI: 10.3791/2565.
11. Thermo Scientific™. Nanodrop: Assessment of nucleic acid purity. Technical Note 52646, 2015;1–3.
12. Qiagen®. FlexiGene® DNA Handbook. 2010. Available at: [www.qiagen.com](http://www.qiagen.com) (accessed March 21, 2016).
13. Hara M, Nakanishi H, Yoneyama K, et al. Effects of storage conditions on forensic examinations of blood samples and bloodstains stored for 20 years. *Leg Med (Tokyo)* 2016;18:81–84.
14. Nakayama Y, Yamaguchi H, Einaga N, et al. Pitfalls of DNA quantification using DNA-binding fluorescent dyes and suggested solutions. *PLoS One* 2016;11:e0150528.
15. Robin JD, Ludlow AT, LaRanger R, et al. Comparison of DNA quantification methods for next generation sequencing. *Sci Rep* 2016;6:24067.
16. Mychaleckyj JC, Farber E, Chmielewski J, et al. Buffy coat specimens remain viable as a DNA source for highly multiplexed genome-wide genetic tests after long term storage. *J Transl Med* 2011;9:91.
17. Erkeller-Yuksel FM, Deneys V, Yuksel B, et al. Age-related changes in human blood lymphocyte subpopulations. *J Pediatr* 1992;120:216–222.

Address correspondence to:  
 Wenlong Carl Chen, MSc (Med)  
 National Cancer Registry  
 National Health Laboratory Service  
 Room F50 STIRC Building  
 1 Modderfontein Drive  
 Sandringham  
 Johannesburg 2010  
 South Africa

E-mail: [wenlongc@nicd.ac.za](mailto:wenlongc@nicd.ac.za)