Analytical and Bioanalytical Chemistry (2020) 412:4749–4760
https://doi.org/10.1007/s00216-020-02733-2

RESEARCH PAPER

Evaluating digital PCR for the quantification of human nuclear DNA: determining target strandedness

Margaret C. Kline[1] · David L. Duewer[2]

Received: 10 April 2020 / Revised: 13 May 2020 / Accepted: 20 May 2020 / Published online: 30 May 2020

Abstract

Digital polymerase chain reaction (dPCR) methodology has been asserted to be a "potentially primary" analytical approach for assigning DNA concentration. The essence of dPCR measurements is the independent dispersal of fragments into multiple reaction partitions, amplifying fragments containing a target nucleotide sequence until the signal from all partitions containing at least one such fragment rises above threshold, and then determining the proportion of partitions with an above-threshold signal. Should originally double-stranded DNA (dsDNA) fragments be converted into two single strands (ssDNA) prior to dispersal, the dPCR measurements could be biased high by as much as a factor of two. Realizing dPCR's metrological potential therefore requires analytical methods for determining the proportion of ssDNA in nominally dsDNA samples. To meet this need, we have investigated several approaches to this determination: A_{260} ratio, dPCR ratio, cdPCR staircase, and ddPCR enzyme. In our hands, only the endonuclease-based approach provides adequately accurate estimates for relatively small ssDNA proportions. We present four (enzyme, assay) pairs that provide self-consistent results for human nuclear DNA extracts. However, the proportion of ssDNA differs by as much as 50% between assays, apparently related to the guanine-cytosine (GC) content of the fragment near the assay's target sequence. While materials extracted by us have no more than 6% ssDNA content even after long storage, a commercially obtained PCR assay calibrant contains ≈18% ssDNA.

Keywords Digital polymerase chain reaction (dPCR) · Double-stranded DNA (dsDNA) · Human nuclear DNA (nDNA) · Metrological traceability · Single-stranded DNA (ssDNA)

Introduction

Efficient probative analysis of DNA evidence requires calibration and validation with reliable reference materials. We and our colleagues at the National Institute of Standards and Technology (NIST) have explored whether and how digital polymerase chain reaction (dPCR) technologies can be used to certify the mass concentration of human nuclear DNA, [nDNA], in higher order calibrants. We have validated fundamental dPCR distributional assumptions using real-time chamber dPCR (cdPCR) [1], developed a reliable technique for determining the volumes of droplet dPCR (ddPCR) reaction partitions [2]; and established a measurement equation for converting measurements of the number of PCR targets per reaction volume to nanograms of human DNA per microliter [3]. We have used these techniques to certify Standard Reference Material (SRM) 2372a Human DNA Quantification Standard for use by the forensic human identity communities [4, 5]. However, until recently we had not

DISCLAIMER: Certain commercial equipment, instruments, or materials are identified in this report to specify adequately experimental conditions or reported results. Such identification does not imply recommendation or endorsement by the National Institute of Standards and Technology, nor does it imply that the equipment, instruments, or materials identified are necessarily the best available for the purpose.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00216-020-02733-2) contains supplementary material, which is available to authorized users.

✉ David L. Duewer
david.duewer@NIST.gov

[1] Biomolecular Measurement Division, Material Measurement Laboratory, National Institute of Standards and Technology, Gaithersburg, MD 20899-8314, USA

[2] Chemical Sciences Division, Material Measurement Laboratory, National Institute of Standards and Technology, 100 Bureau Drive, Stop 8390, Gaithersburg, MD 20899-8390, USA

adequately investigated a potentially significant source of bias for dPCR analysis.

The presence of single-stranded DNA (ssDNA) fragments in a sample believed to be entirely double-stranded DNA (dsDNA) will cause dPCR results to be biased high [6, 7]. A dsDNA fragment that fully separates will disperse as two ssDNA fragments. A dPCR evaluation of a completely ssDNA sample prepared from a completely dsDNA parent will thus overestimate [nDNA] by a factor of 2. Based on the close agreement between conventional spectrophotometric absorption assessments [8, 9] of freshly extracted materials before and after NaOH-denaturing, our analyses have assumed that extracted DNA is entirely dsDNA and that the potential ssDNA-related bias need not be explicitly addressed.

Measurement results for a human-source material sold as a calibrant for use with (non-digital) quantitative PCR (qPCR) challenged this assumption. Using supplier-assigned values, we gravimetrically prepared 10 mL of a 50.0(1) ng/μL solution for use as a transfer standard, named here as "ComY." (Note: Values with associated uncertainties are stated as "xx(y)" where "xx" is the value and "y" is the standard uncertainty expressed in units of the least-significant decimal place.) Conventional spectrophotometric characterization [8] estimated the [nDNA] of this material as 50.2(1) ng/μL. Unexpectedly, determination of the NaOH-denatured analogue of this material [9] estimated [nDNA] as 47.6(1) ng/μL. Subsequent evaluation using ddPCR results and the conversion equation [3] estimated [nDNA] to be 53(2) ng/μL. While this ≈12% discordance would be unexceptional for routine samples, it is unacceptable for a primary calibrant.

To resolve whether this and other DNA extracts contained detectable proportion of ssDNA, we sought methods for quantifying ssDNA. Except for an indirect and platform-specific proposal [10], we are unaware of any current methods suitable for use with complex genomic mixtures.

This report summarizes our assessments of four methods for evaluating "strandedness" and our methods for producing materials with the range of ssDNA proportions that the assessments required. We have determined that comparison of spectrophotometric absorbance measurements at 260 nm of native and denatured sample (A_{260} Ratio Analysis) is not adequately quantitative due to changes induced by DNA fragmentation during denaturation. Somewhat similarly, comparison of dPCR results before and after heat denaturation (dPCR Ratio Analysis), is inadequately sensitive due to loss of target accessibility or amplifiability during denaturation. While it can provide direct visualization of the ssDNA proportion and be adequately sensitive, successful evaluation of cumulative crossing-threshold values from real-time cdPCR (cdPCR Staircase Analysis) requires an extremely efficient PCR assay and specialized instrumentation. However, given a restriction enzyme that efficiently cuts the dsDNA target of an efficient PCR assay while leaving ssDNA targets mostly intact, comparison of ddPCR results before and after enzymatic restriction (ddPCR Enzyme Analysis) provides accurate and adequately sensitive assessment of the proportion of ssDNA in a DNA extract.

Materials and methods

While all critical materials and methods used in this study are summarized here, background information and full details for many of the methods are provided in the freely available NIST Special Publication (SP) 1200–27 [11]. The methods described in SP 1200–27 were developed using DNA extracts from only 2 sources; we here use samples from many sources with quite different histories to demonstrate the general applicability of the methods for human nDNA.

Samples

Table 1 lists the 13 native extracts used in this study, representing 12 human donors or donor pools: 6 female, 4 male, and 2 mixed female and male. Eight of the samples are components of Standard Reference Materials (SRMs) 2390 DNA Profiling Standard (discontinued), 2372 Human DNA Quantitation Standard (replaced by SRM 2372a), or 2372a Human DNA Quantitation Standard [4]. Four of the extracts were commercially obtained, the rest were extracted

Table 1 Native samples

		Extracted/Stored [b]		
Name	Source DNA[a]	Year	ng/μL	°C
A	male, sd (SRM 2372 A) NIST	2006	57	4
B	female, md (SRM 2372 B) NIST	2006	61	4
C	mixed, md (SRM 2372 C) commercial	<2006	59	4
aA	male, sd (SRM 2372a A) NIST	2016	50	4
aB	female, sd (SRM 2372a B) NIST	2016	58	4
aC	mixed, md (SRM 2372a C) NIST	2016	48	4
CanB	female, md, NIST	2006	60	4
ComY	male, md, commercial	<2013	50	4
K562	female, cl (SRM 2390–13) commercial	<1990	230	−80
NIST1	male, sd (SRM 2390–16) commercial	<1990	200	−80
NIST4	female, sd, NIST	2015	170	4
NIST5	female, sd, NIST	2010	209	4
UB	female, md (SRM 2372 B stock) NIST	2006	100	4

[a] cl: cell line, md: multi-donor, sd: single donor, NIST = extracted at NIST, commercial = extracted elsewhere

[b] year extracted, approximate mass concentration of extract, extract storage temperature

at NIST from buffy coat cells using a salting out procedure [12], modified by use of ammonium acetate in place sodium chloride. These sources provide a diverse set of reference material-quality human DNA extracts stored under known conditions, concentrations, and time.

Table 2 lists 17 samples derived from the source materials to provide samples with known relative proportions of ssDNA. The manipulations used to produce these materials included heat denaturing to completely convert dsDNA into ssDNA, mechanically shaking to partially convert dsDNA into ssDNA, and volumetric combining of two materials to produce samples of known intermediate proportions of ssDNA. See the Electronic Supplementary Material (ESM) Fig. S1 for graphical representations of the relationships among these materials.

All of the native extracts are prepared for analysis and/or for use in the preparation of manipulated materials as ≈50 ng/μL nuclear DNA in 10 mmol/L tris-(hydroxymethyl)aminomethane HCl, 0.1 mmol/L ethylenediaminetetraacetic acid, pH 8.0 (TE^{-4}) buffer.

Table 2 Manipulated samples

Code [a]	Source DNA	Manipulation
CanBd	CanB,	Heat denatured
CanBd20	CanB, CanBd	Mixture, 80:20
CanBd40	CanB, CanBd	Mixture, 60:40
CanBd80	CanB, CanBd	Mixture, 20:80
NIST4d	NIST4	Heat denatured
NIST4s	NIST4	Shaken
NIST4sd	NIST4s	Heat denatured
UBd	UB,	Heat denatured
UBd20	UB, UBd	Mixture, 80:20
UBd40	UB, UBd	Mixture, 60:40
UBd80	UB, UBd	Mixture, 20:80
UBd50	UB, UBd	Mixture, 50:50
UBd50d	UBd50	Heat denatured
UBs	UB	Shaken
UBsd	UBs,	Heat denatured
UBs50	UB, UBs	Mixture, 50:50
UBs50d	UBs50	Heat denatured

[a] Codes used as suffixes to source material name

d: heat denatured

dxx: volumetric mixture of a native and its heat-denatured analog, xx denotes percentage of manipulated material

dxxd: heat denatured volumetric mixture of a native and its heat-denatured analog

s: shaken

sd: heat denatured shaken material

sxx: volumetric mixture of a native and its shaken analog, xx denotes percentage of manipulated material

sxxd: heat denatured volumetric mixture of a native and its shaken analog

Heat denaturation

30 μL samples are thermally denatured in 200 μL PCR tubes by heating at 96 °C for 60 s using a well-calibrated thermocycler, then rapidly cooled to 4 °C and stored on ice until ready for use.

Mechanically shaken

200 μL of sample material is aliquoted into a 1.5 mL polypropylene tube. The tube is mechanically shaken at 147 rad/s (1400 RPM) at 56 °C for 3 h in an Eppendorf Thermomixer F1.5 (Eppendorf North America, Hauppauge, NY, USA). The shaken sample is then stored at 4 °C until ready for use. When a larger sample volume is desired, the contents of multiple 200 μL aliquots are combined.

Volumetric mixtures

For 30 μL of a (100-*x*):*x* volumetric mixture of a native sample and its manipulated analog, combine 30(100-*x*)/100 μL of the native ≈50 ng/μL sample with 30(*x*)/100 μL of its freshly manipulated analog in a in 200 μL PCR tube; e,g, for a 20:80 mixture combine 30 × 20/100 = 6 μL of the native sample with 30 × 80/100 = 24 μL of the manipulated sample. Mix by vortexing then microcentrifuge. Store at 4 °C until ready for use.

Spectrophotometry

Absorbance spectra are acquired using a Cary 3500 spectrophotometer (Agilent, Santa Clara, CA, USA) and 70 μL microcuvettes. Data are collected at 0.5 nm increments from 220 nm to 360 nm, slit width of 1 nm, and a dwell time of 0.048 s per increment. Samples are diluted to ≈17 ng/μL.

Gel electrophoresis

Electrophoretic evaluations are acquired using Genomic DNA ScreenTape devices on a TapeStation 4150 (Agilent Technologies, Santa Clara, CA). Evaluations are performed using the device's sample loading buffer and the loading buffer augmented with 1X of SYBR Green II dye. The proprietary buffer is optimized for detection of dsDNA using the SYBR Green I dye. Addition of SYBR Green II enhances ssDNA detection sensitivity, although the augmented buffer is not necessarily optimum for either dye. These intercalating dyes bind to both dsDNA and ssDNA, but with different affinities and fluorescence intensities.

Chamber digital PCR (cdPCR)

We use a Fluidigm BioMark (South San Francisco, CA USA) real time cdPCR system with BioMark 48.770 digital arrays. Each analysis uses a disposable microfluidic device ("chip") that has 48 panels of 770 reaction chambers. The supplier states that all chambers have nominal volumes of 0.75 nL. Samples are prepared using the TaqMan Gene Expression Master Mix (Applied Biosystems, Foster City, CA). Analyses are performed according to the manufacturer's recommendations except for PCR assay-specific annealing between 56 °C and 61 °C, amplifying for 60 cycles, and using ≈1 ng/μL as our target nuclear DNA mass concentration in the reaction mixture. Chambers are considered positive when the fluorescence signal exceeds a user-set threshold at the end of the 60th amplification cycle. The threshold is typically set to be just above the background fluorescence. To enable consistent results, we use a minimum of six technical replicates (panels) per sample per chip. All samples are evaluated in at least two chips.

Droplet digital PCR (ddPCR)

We use QX100 and QX200 ddPCR systems (Bio Rad, Hercules, CA USA) where samples are dispersed as ≈0.74 nL droplets in a fluorinated oil using a disposable microfluidic cartridge. Samples are prepared for amplification using the manufacturer's "Supermix for Probes (No dUTP)". Analyses are performed according to the manufacturer's recommendations, except for annealing at 61 °C, amplifying for 60 cycles, accepting results when at least 5000 droplets are counted, and using 1 ng/μL as our target nuclear DNA mass concentration in the reaction mixture. Droplets are considered positive if the fluorescence signal exceeds a user-set threshold, typically set to be just above the signal for non-template control (NTC) samples. To enable consistent results, we use a minimum of four technical replicates (wells) per sample per plate. All samples are evaluated in at least two plates.

PCR assays

The ten human nuclear DNA PCR assays used to value-assign SRM 2372a are described elsewhere [3, 4, 11]. Four of these assays have been used in this study: D5, HBB1, NEIF, and POTP. The target sequence for D5 is located near the centromere of chromosome 5. The target for HBB1 is near the tip of the p-arm of chromosome 11. The target sequences for NEIF and POTP are located on chromosome 2, respectively near the centromere and tip of the p-arm.

Restriction enzymes

Type II restriction endonucleases cut dsDNA at sequence-specific locations. The enzymes *Hinf*I, *Nla*III, *Sty*I-HF, and *Xcm*I were purchased from New England BioLabs Inc. (Ipswich, MA USA) and were used according to the supplier's recommendations, with the exception that all digestions were carried out for 60 min.

Computation

Data are exported from the instrumental platforms into purpose-developed spreadsheet-based analysis systems.

Results and discussion

A_{260} Ratio analysis

DNA mass concentration for "pure" dsDNA expressed in units of nanogram per microliter is conventionally proportional to 50 times the absorbance at 260 nm using a 1 cm pathlength: [dsDNA] = 50 A_{260} ng/μL [8]. The usual proportionality for "pure" ssDNA is 40, but literature values range from 37 to 40: [ssDNA] = (37 to 40) A_{260} ng/μL [8, 13, 14].

If heat denaturation completely and permanently separates dsDNA entities into two ssDNA entities without fragmentation, the proportion of all ssDNA in the sample, p_{ss}, should be linearly related to the relative increase in A_{260} after heat-denaturation:

$$p_{ss} = (A_{260n}/A_{260d} - R)/(1 - R), \quad (1)$$

where A_{260n} is the absorbance of the native extract, A_{260d} is the absorbance after denaturation, and R is the ratio of the proportionality constants for dsDNA relative to ssDNA. If the conventional constants are metrologically true, $R = 40/50 = 0.80$. The derivation of Eq. 1 is detailed in the ESM.

Table 3 lists absorbance measurements for four {native, denatured} pairings: {UB, UBd}, {UB50, UB50d], [UBs, UBsd}, and {UBs50, UBs50d}. These materials were volumetrically prepared to have identical [nDNA]. Since {UB, UBd} is the least manipulated, the p_{ss} for the UB sample should be small; Eq. 1 yields values from −0.030(4) to 0.208(3) using the conventional range of constants: R = 40/50 = 0.800 to R = 37/50 = 0.740.

Assuming: 1) the only UV-active entities in the extracts are unfragmented dsDNA and ssDNA, 2) UBd is entirely ssDNA, and 3) p_{ss} for UB is zero then R = 0.794. Using this value for the more manipulated samples yields plausible estimates, although the assumption that all dsDNA and ssDNA are unfragmented is implausible: DNA is fragmented by

Table 3 A_{260} Ratio estimates of ssDNA proportion

Sample	A_{260} [a]	p_{ss} R = 0.800	R = 0.794	R = 0.740
UB UBd	0.3572(2) 0.4499(3)	−0.030(3)	0.000(3)	0.208(3)
UBd50 UBd50d	0.3972(1) 0.4556(5)	0.359(5)	0.378(5)	0.507(4)
UBs UBsd	0.3883(11) 0.4665(2)	0.162(12)	0.186 (12)	0.355(9)
UBs50 UBs50d	0.3817(51) 0.4686(2)	0.073(54)	0.100(53)	0.287(42)

[a] Values in parentheses report standard deviations between independent preparations, expressed in units of the least-significant decimal place of the associated absorbance

mechanical manipulation and heating. The consequences of manipulation are apparent in the increase in A_{260d} from UBd to UBs50d.

Denaturation and fragmentation are confirmed by the electrophoretic behavior of the native and heat-denatured UB, UBd, UBs, and UBsd samples in the as-supplied and SYBR Green II augmented buffers (ESM Fig. S2). In both buffers, the electropherograms of the non-denatured samples display a single peak with an average apparent size of (55,000 to 60,000) basepairs (bp). In the augmented buffer, the heat-denatured materials display a band with an apparent size much greater than 60,000 bp. The "smaller" band thus reports dsDNA, the "larger" ssDNA. The dsDNA band is nearly absent in heat-denatured materials, indicating near-complete strand separation. While the height of the dsDNA band in UBs is much reduced, the area under its left-hand (low bp size) tail is increased and there is no detectable ssDNA peak. While shaking fragments this material, consistent with the observed increase in A_{260}, it does not detectably induce strand separation.

The A_{260} measurand is the intrinsic absorption of the constituent nucleotide monomers modified by the "hyperchromicity" or π bond shielding provided by the higher-order structures of the polymer [14–16]. The less shielding, the larger A_{260}. Without a reliable estimate for R, A_{260d}/A_{260n} ratios do not provide reliable p_{ss}(A_{260} Ratio) estimates. Furthermore, the large uncertainty in A_{260n} for the UBs and UBs50 samples reflects differences in nominally identically treated aliquots of the same material. This suggests that small differences in treatment can induce relatively large changes in structure.

While our heat-denaturing protocol (96 °C for 60 s) appears to completely convert dsDNA into ssDNA, it also induces fragmentation which contributes to π-bond deshielding. In our hands, this assay does not provide reliable p_{ss} estimates for human [nDNA].

dPCR Ratio analysis

To the extent that the fragments containing a PCR assay's target sequence ("entities") are independently and randomly dispersed into equal-volume reaction partitions, the average number of entities per partition, λ, can be estimated as: $\lambda = -\ln(1 - F_{pos})$, where F_{pos} is the fraction of counted partitions that provide a positive signal and "ln" is the Napierian logarithm [1, 17]. If heat denaturation completely and permanently separates dsDNA entities into two ssDNA entities without changing the accessibility or amplifiability of assay targets, the proportion of ssDNA entities in the sample should be linearly related to the relative increase in the number of entities after denaturation:

$$p_{ss} = 2 - \lambda_d/\lambda_n, \quad (2)$$

where λ_n is number of entities per partition in the native extract and λ_d is the number after heat denaturation. The derivation of Eq. 2 is detailed in the ESM.

Figure 1 displays p_{ss} (dPCR Ratio) estimates for two series of volumetrically prepared mixtures of native and denatured samples: {CanB, CanBd20, CanBd40, CanBd80} and {UB, UBd20, UBd40, UBd80}. The estimates from Eq. 2 are plotted against the volume fraction of the denatured sample:

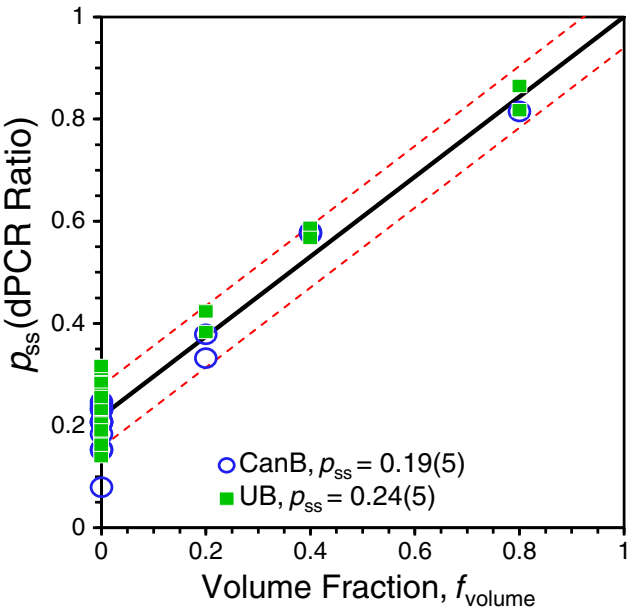

Fig. 1 dPCR Ratio Estimates of ssDNA Proportion. Open blue circles represent p_{ss} estimates for the CanB series of native and denatured sample, solid green squares for the UB series. The thick black line represents an average linear relationship between p_{ss} and the volume fraction of denatured sample in the mixture, f_{volume}; the dashed red lines bound approximately 68% of the estimates. The p_{ss} estimates for the native CanB and UB samples are estimated as the intercepts of linear fits that are constrained to provide $p_{ss} = 1$ at $f_{volume} = 1$

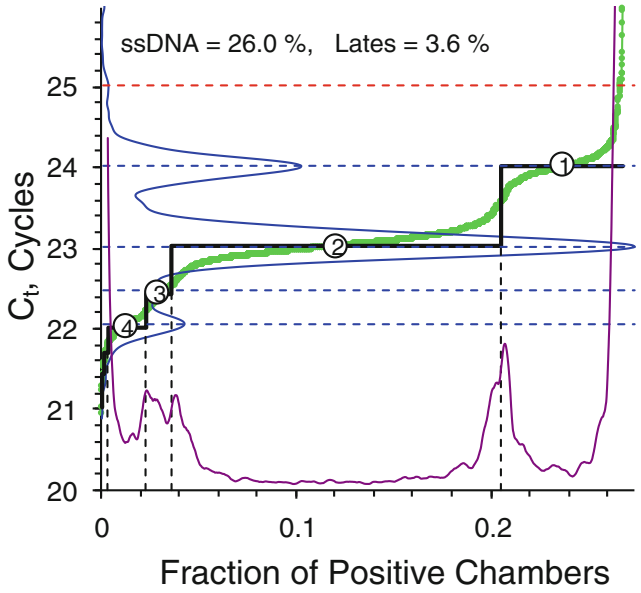

Fig. 2 Exemplar cdPCR Staircase Analysis. The green curve is the observed cumulative distribution (ogive) of crossing threshold (C_t) values as a function of the fraction of positive chambers (F_{pos}) for sample UBs50. This result was achieved with the POTP assay, an annealing temperature of 56 °C, and using the manufacturer's "linear derivative" analysis. The blue curve plotted along the C_t axis is a kernel-density representation of the number of positive chambers at given C_t. The purple curve along the F_{pos} axis estimates the ogive rate of change derivative, $d(C_t)/d(F_{pos})$. The thick black "staircase" curve is the idealized ogive for a sample consisting of 26% single-stranded DNA (ssDNA) in an extract that is otherwise double-stranded DNA (dsDNA). The middle of each tread is labeled with the number of entities per chamber that produce the tread. The horizontal blue dashed lines mark the location of kernel density peak maxima. The C_t location of the staircase is defined by alignment of the uppermost tread and the kernel density function's uppermost peak. The black vertical dashed lines connect the location of the stair risers with the horizontal axis; ideally, they bisect a peak in the derivative (purple curve). The horizontal red dotted line is one C_t above the one-entity tread; C_t values above this line are considered "late starters" and are excluded from the optimization process. The text reports the estimated percentage of ssDNA and late starters in the sample

$$f_{\text{volume}} = V_d/(V_n + V_d), \tag{3}$$

where V_n is the volume of the native sample in the mixture and V_d is the volume of the denatured sample. The CanB and UB native samples were assessed with all four PCR assays, those for the mixture samples were assessed just with HBB1 and POTP. While cdPCR can provide λ_d/λ_n ratios, those displayed in Fig. 1 were acquired as a component of the ddPCR Enzyme analyses. The replication variability prevents determining whether the linear model in f_{volume} is appropriate.

The long-term between-aliquot standard deviation, $s(p_{ss})$, of the dPCR Ratio estimates can be modeled as a parabolic function of p_{ss} with a maximum at p_{ss} of 0.5 declining to zero at p_{ss} of 0 and 1 (ESM Fig. S3). For displays and analyses that utilize uncertainties, the standard uncertainties for the dPCR Ratio estimates are assigned as $u(p_{ss}) = \max(0.21p_{ss}(1\text{-}p_{ss}))$, $s(p_{ss})$), where "max" is the function "take the larger value".

cdPCR Staircase analysis

Since our cdPCR system monitors fluorescence intensity in all reaction chambers at the end of each amplification cycle, it is possible to estimate the cycle at which each positive chamber exceeds the threshold, C_t. For PCR assays providing sufficiently coordinated and efficient amplification, analysis of the cumulative distribution (ogive) of C_t as a function of cycle number enables estimation of the fraction of chambers that originally contained $n = (0, 1, 2, \ldots)$ amplifiable entities [18]. Comparison of the observed ogive with idealized ogives for trial p_{ss} values enables estimation of a best-fit p_{ss} value. Fig. 2 illustrates the method; the method is described in detail in SP 1200–27 [11].

Figure 3 displays cdPCR Staircase estimates of p_{ss} for the two series of volumetrically prepared mixtures discussed above, with the addition of samples CanBd and UBd. These estimates were obtained using the NEIF assay. Fig. 3a plots p_{ss} as functions of f_{volume}, with the p_{ss} values for the native CanB and UB samples estimated as the intercepts of least-squares quadratic fits. The Staircase estimates are sufficiently precise to confirm that the p_{ss} values for the mixtures are not linear in f_{volume}. Fig. 3b plots p_{ss} as functions of the fraction of ssDNA entities relative to all accessible and amplifiable targets [19] in the sample, f_{entity}:

$$f_{\text{entity}} = \left(p_{ss,n}(1-f_{\text{volume}}) + p_{ss,d}\cdot\varphi\cdot f_{\text{volume}}\right)/(1 + f_{\text{volume}}(\varphi-1)), \tag{4}$$

where $p_{ss,n}$ is the proportion of ssDNA in the native sample, $p_{ss,d}$ is the proportion of ssDNA in the denatured sample, and φ is the effective number of ssDNA accessible and amplifiable entities produced by denaturing a dsDNA entity. In Fig. 3b, these parameters are estimated as the mean p_{ss} of the native and denatured samples, with $\varphi = 1.78$ (the mean of the observed λ_d/λ_n for these samples). Very similar values were obtained using unconstrained optimization.

When an ogive provides well-defined stairsteps, between-analyst interpretation differences are seldom more than ±0.005. The long-term between-chip standard deviation, $s(p_{ss})$, is an approximately constant 0.017 (ESM Fig. S4). For displays and analyses that utilize uncertainties, the standard uncertainties for the cdPCR Staircase estimates are assigned as $u(p_{ss}) = \max(0.017, s(p_{ss}))$.

ddPCR Enzyme analysis

Restriction with an enzyme that cuts dsDNA between an assay's forward and reverse primers can render those targets

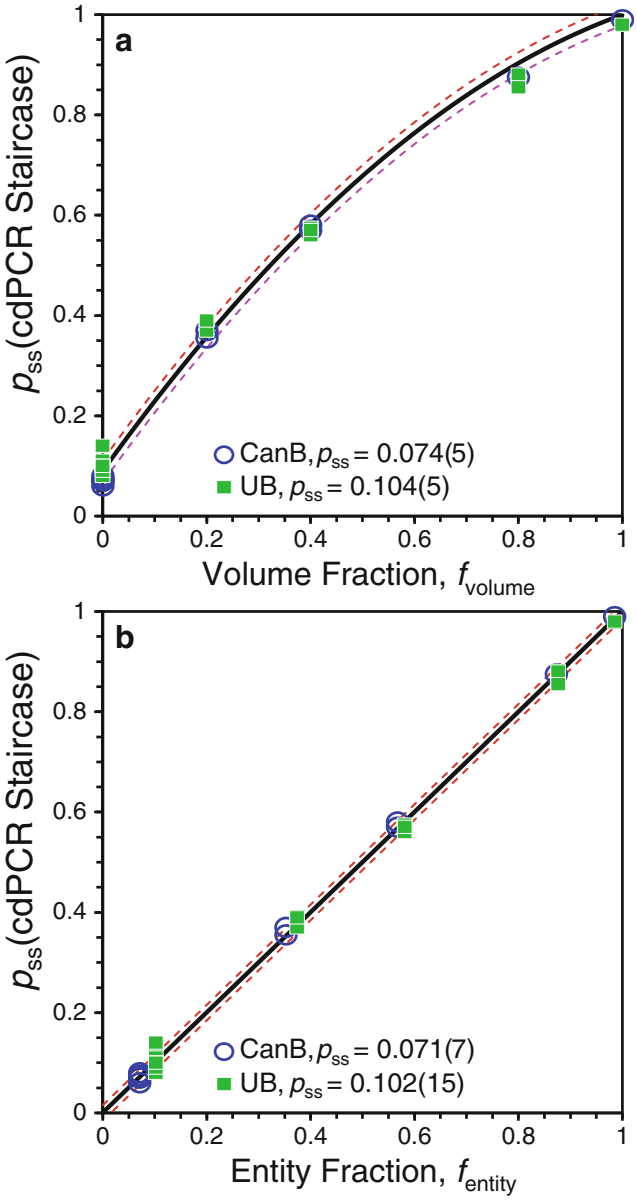

Fig. 3 cdPCR Staircase Estimates of ssDNA Proportion. Open blue circles represent NEIF-assay p_{ss}(cdPCR Staircase) estimates for the CanB series of native and denatured sample, solid green squares for the UB series. The thick black curves represent best-fit relationships that are constrained to provide $p_{ss} = 1$ at f_{volume} or $f_{entity} = 1$; the dashed red curves bound approximately 68% of the estimates. **a** p_{ss} plotted as functions of the volume fraction of denatured sample in the mixture, f_{volume}. The p_{ss} estimates for the native CanB and UB samples are estimated as the intercepts of quadratic fits. **b** p_{ss} estimates plotted as functions of ssDNA entity fraction, f_{entity}, using parameters defined from the {CanB, CanBd} and {UB, UBd} pairs

non-amplifiable. If the restriction completely blocks dsDNA amplification without damaging ssDNA targets, p_{ss} should be equal to the ratio of the number of entities in the enzyme-treated and native samples [20]. To account for fractional loss of ssDNA due to enzymatic cutting of "double-strand-like" transient conformations [21], this ratio needs to be corrected by the ratio of the number of entities in the denatured enzyme-treated and heat-denatured samples:

$$p_{ss} = (\lambda_e/\lambda_n)/(\lambda_{de}/\lambda_d), \quad (5)$$

where λ_e is number of entities per partition in the native extract after restriction and λ_{de} is the number in the denatured sample after restriction. Fig. 4 displays the enzyme-assay combinations used in this study.

The long-term between-aliquot standard deviation of the p_{ss} values, $s(p_{ss})$, follows a Horwitz-like uncertainty function [22] that increases with increasing p_{ss} (ESM Fig. S5). For displays and analyses that utilize uncertainties, the standard uncertainties for the ddPCR Enzyme estimates are assigned as $u(p_{ss}) = \max(0.02(p_{ss})^{0.63}, s(p_{ss}))$.

Figure 5 displays p_{ss}(ddPCR Enzyme) estimates from the Xcm:POTP assay for the two mixture-series. The p_{ss} values are plotted as a functions of f_{entity}, where the $p_{ss,n}$ and φ parameters are estimated as the mean measured values for the native sample and $p_{ss,d} = 1.065$ is estimated by optimization. The optimized $p_{ss,d}$ value is compatible with the 1.03(4) λ_{de}/λ_d distribution characteristic of the Xcm:POTP assay. These >1 values suggest that for the POTP assay, restriction reduces hairpins or other intrastrand pairings [23] that hinder amplification.

Comparisons among analysis methods

ddPCR Xcm:POTP Vs dPCR Ratio analysis

Figure 6 compares the p_{ss}(dPCR Ratio) and p_{ss}(Xcm:POTP) estimates for 13 natural (minimally manipulated) samples, 2

Assay	Forward Primer – **Probe** – Reverse Primer
NEIF	GC…TC**TCATG**|**CAGTTGTCAGAAGCTG***c*TG…AG CG…AG*a*|**GTACGTCAACAGTCTTCGAC***g*AC…TC *NlaIII*
POTP	CC…TT*t***CACCAACTGA**|**AATATG***g*CC…GT GG…AA*a***GTGGTTGAC**|**TTTATAC***c*GG…CA *XcmI*
D5	TT…AT*a***ATAATATCAGGGTAAACAGGG**|*a*ATCT…AG AA…TA*t***TATTATAGTCCCATTTGTCCC***t*TA|GA…TC *HinfI*
HBB1	GC…TC*cta***AGCCAGTGCCAGAAGAGC**|**CAAGGA***c*AG…CC CG…AG*gat***TCGGTCACGGTCTTCTCGGTTC**|**CT***g*TC…GG *StyI-HF*

Fig. 4 Enzymatic Inactivation of dsDNA Targets. Each row describes one dPCR assay and the restriction enzyme used to make its target non-amplifiable. The first and second lines in each row display the forward and reverse strands, where A = adenine, C = cytosine, G = guanine, and T = thymine. Cut sites are marked with "|", with lines connecting overhanging bases. Forward primer sequences are in red, reverse primers in orange, and probes in green. Bases located in the flanking regions between the primer and probe sequences are in *lowercase italic blue*, "…" represents two or more bases that are not pertinent to the cut site

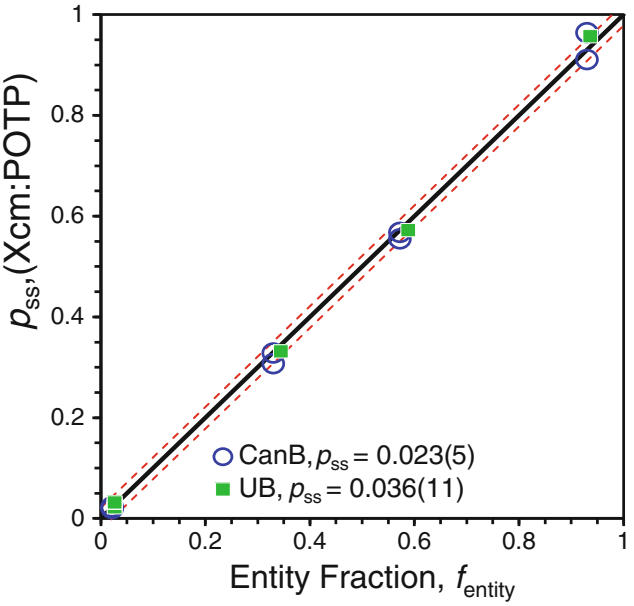

Fig. 5 ddPCR Enzyme Estimates of ssDNA Proportion. Open blue circles represent p_{ss}(Xcm:POTP) estimates for the CanB series of native and denatured sample, solid green squares for the UB series. The p_{ss} estimates are plotted as functions of ssDNA entity fraction, f_{entity}, using parameters defined by the CanB and UB samples and optimization. The thick black line represents ideal proportionality; the dashed red lines bound approximately 68% of the estimates

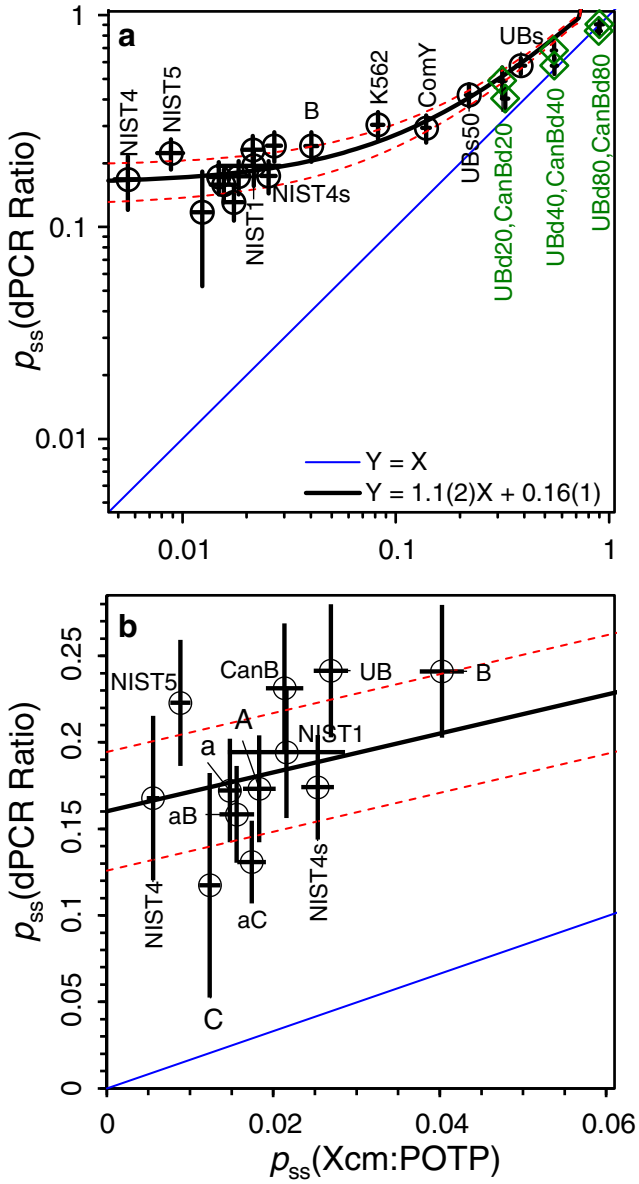

Fig. 6 Comparison of Enzyme and Ratio Assays. Symbols represent mean estimates from the dPCR Ratio as a function of the ddPCR Xcm:POTP analyses; black open circles denote results from unmodified and shaken samples; green open diamonds denote results for mixtures containing heat-denatured sample. The crosses represent standard uncertainties. The thin blue lines represent equality between the assays. The thick black lines represent an errors-in-variables best-fit linear relationship between the assays for the unmodified and shaken samples; the dashed red lines bound approximately 68% of the estimates. **a** Estimates displayed on $\log_{10}$ axes. **b** A higher-resolution display on linear axes

samples prepared by mechanical shaking, 1 mixture prepared using a shaken material, and the 6 mixtures prepared using denatured materials. Using the FREML error-in-variables technique that utilizes uncertainty estimates in both variables [24, 25], results for all but the heat-denatured mixtures are linearly related with a non-zero intercept of ≈0.17.

Estimating p_{ss}(dPCR Ratio) using Eq. 2 assumes that heat denaturation at 96 °C for 60 s does not reduce the number of amplifiable and accessible targets. If denaturation renders some ssDNA entities inaccessible or non-amplifiable, λ_d will be reduced and 2 - λ_d/λ_n will increase. The p_{ss}(dPCR Ratio) estimates thus may be biased high as well as being imprecise. If the denaturation process reduces the number of amplifiable and accessible targets by a consistent fractional dsDNA entity loss of 0.05, the "2" in Eq. 2 would become 1.9 and the intercept (and thus the bias) would be reduced to ≈0.07. Even if justified, this modification does not improve the precision of the estimates. We speculate that the imprecision of p_{ss}(dPCR Ratio) measurement is driven by variable target loss due to small differences in sample preparation (e.g., vortex and centrifuge duration).

The relationship between the assays for the 6 mixtures prepared using heat-denatured material appears to approach equality for the mixtures that contain 80% denatured sample. This suggests that heat-denaturation does not do much damage to the ssDNA entities produced by heat-denaturation.

While dPCR Ratio analysis successfully measures the proportion of ssDNA fragments at a given locus, our heat-denaturing protocol inconsistently reduces the number of fragments containing an accessible and amplifiable target. The method is thus useful mainly for identifying materials with very high (> 0.2) p_{ss}. The poor precision and high detection limit for human nDNA may not be representative for other forms of dsDNA that can be fully and reliably converted to ssDNA using gentler denaturation processes.

ddPCR Xcm:POTP Vs cdPCR Staircase analysis

Figure 7 compares the NEIF-assay p_{ss}(cdPCR Staircase) and p_{ss}(Xcm:POTP) estimates for the same 22 samples used above. As with p_{ss}(dPCR Ratio), for all but the heat-denatured mixtures p_{ss}(cdPCR Staircase) is linearly related to p_{ss}(Xcm:POTP) with a non-zero intercept. The FREML estimate of the intercept is ≈0.06.

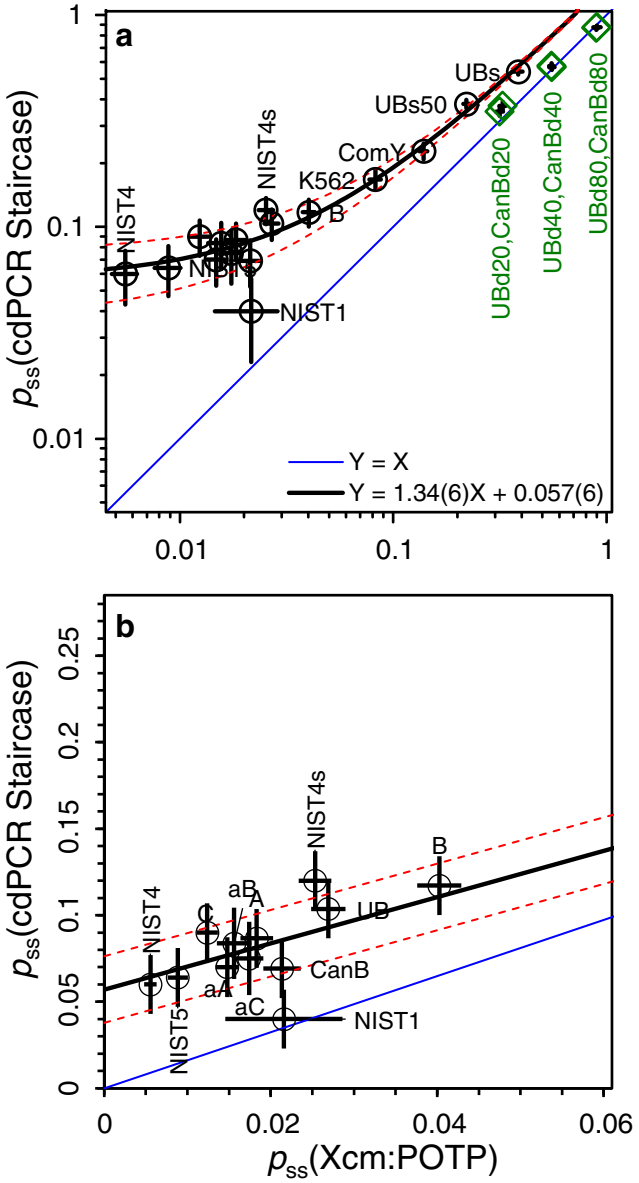

Fig. 7 Comparison of Enzyme and Staircase Assays. Symbols represent mean estimates from the NEIF-based cdPCR Staircase assay as a function of the ddPCR Xcm:POTP Enzyme analyses; black open circles denote results for unmodified and shaken samples; green open diamonds denote results for mixtures containing heat-denatured sample. The crosses represent standard uncertainties. The thin blue lines represent equality between the assays. The thick black lines represent an errors-in-variables best-fit linear relationship between the assays for the unmodified and shaken samples; the dashed red lines bound approximately 68% of the estimates. .**a** Estimates displayed on $\log_{10}$ axes. **b** A higher-resolution display on linear axes

This intercept may be an artifact of the Staircase method since ogives do not distinguish between chambers containing single ssDNA entities that amplify during the first cycle and those containing single dsDNA entities that start to amplify during the second cycle. The length of the ssDNA tread may thus be biased high; however, the greater length and complexity of the microfluidic piping used in cdPCR chips relative to ddPCR cartridges may induce more mechanical strand separation prior to dispersal into the respective reaction partitions. The near-equality between the assays for the 80% ssDNA mixtures is compatible with this speculation.

Given an ideal PCR assay, the cdPCR Staircase analysis directly estimates p_{ss}. However, an ideal assay would induce all target-containing fragments to start amplifying during the same cycle and produce the same number of copies per target per cycle. Unfortunately, none of our human nDNA assays are quite that efficient and so may provide results that are somewhat biased.

Xcm:POTP Vs other ddPCR Enzyme assays

Figure 8 compares the p_{ss}(Sty:HBB1) and p_{ss}(Xcm:POTP) estimates for the 22 samples. The results from these two assays are not only linearly related, they are proportional; the standard uncertainty in the FREML estimated intercept is nearly as large as the intercept's value. Using a zero-intercept FREML model, the p_{ss}(Sty:HBB1) estimates are ≈1.6-fold larger than p_{ss}(Xcm:POTP) for the natural and the shaken samples. However, there are sample-specific as well as target-specific factors impacting the ddPCR Enzyme analyses: p_{ss}(Sty:HBB1) is more than 3-fold larger than p_{ss}(Xcm:POTP) for NIST4 but is less than 1.2 larger for NIST1.

There are similar proportional relationships among other ddPCR Enzyme assays (ESM Fig. S6). This suggests that "single-strandedness" is not a simple global property of all human nDNA fragments but rather is local to each target sequence. The near-equality between p_{ss}(Sty:HBB1) and p_{ss}(Xcm:POTP) for the CanBd20, CanBd40, CanBd80, UBd20, UBd40, and UBd80 mixtures is compatible with this inference. To the extent that heat-denaturing irreversibly converts all dsDNA to ssDNA, it eliminates locus-specific differences.

For all samples other than those prepared as heat-denatured mixtures, p_{ss}(Xcm:POTP) is smaller than estimates from the other assays investigated: asymptotically 1.1(2)-fold smaller than p_{ss}(dPCR Ratio), asymptotically 1.34(6) smaller than p_{ss}(cdPCR Staircase), 1.36(4) smaller than p_{ss}(Nla:NEIF), 1.57(4) than p_{ss}(Hinf:D5), and 1.61(4) than p_{ss}(Sty:HBB1). While the small increase in λ_d after restriction with *Xcm*I may contribute to these differences, we speculate that the smaller results provided by p_{ss}(Xcm:POTP) signifies that

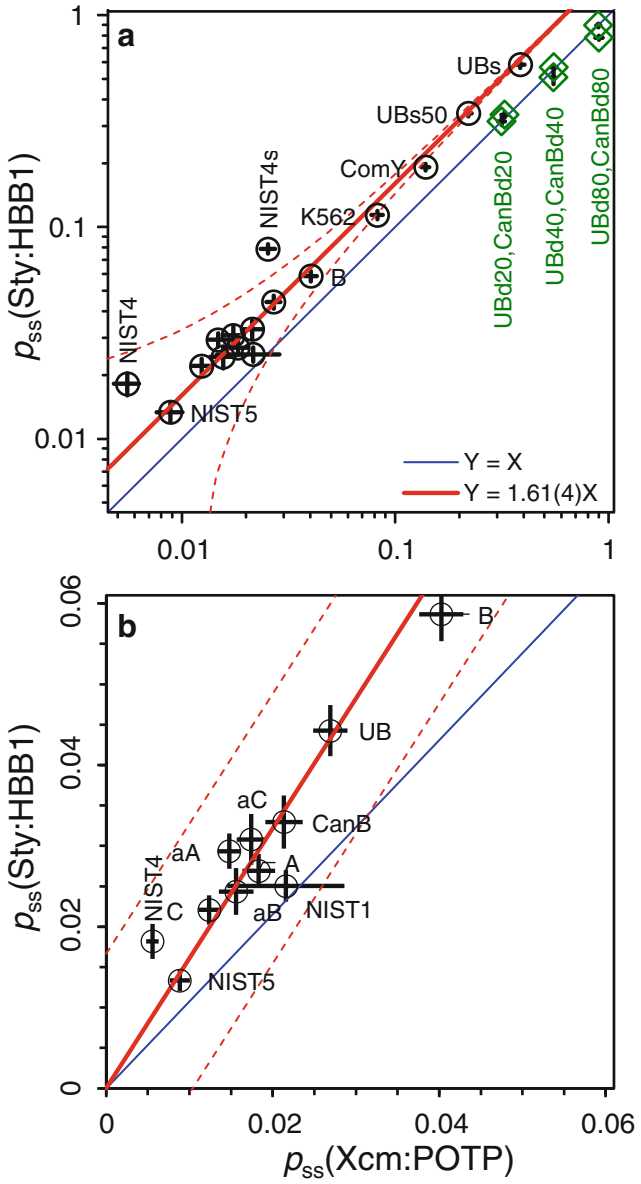

Fig. 8 Comparison of ddPCR Enzyme Assays. Symbols represent mean estimates from the ddPCR p_{ss}(Sty:HBB) Enzyme assay as a function of the ddPCR Xcm:POTP Enzyme analyses; black open circles denote results for unmodified and shaken samples; green open diamonds denote results for mixtures containing heat-denatured sample. The crosses represent standard uncertainties. The thin blue lines represent equality between the assays. The thick black lines represent an errors-in-variables best-fit proportional relationship between the assays for the unmodified and shaken samples; the dashed red lines bound approximately 68% of the estimates. **a** Estimates displayed on $\log_{10}$ axes. **b** A higher-resolution display on linear axes

there are fewer native ssDNA fragments containing the POTP target sequence than for the other PCR assays considered here. DNA stability increases with increasing guanine-cytosine (GC) content [26]. The GC content of the 2000 basepair interval centered on the midpoint of the POTP target sequence is almost 0.59 whereas it is at most 0.41 for the other assays (ESM Fig. S7). In addition to being related to GC content, the proportion of locally ssDNA structure within a DNA fragment may also be related to the storage history of the extract and/or its concentration: the same shaking process induced much less local ssDNA in NIST4 (stored at a relatively high concentration for 4 y) than in UB (stored at lower concentration for 13 y).

P_{ss} Vs λ_n

Direct evaluation of the relationship between the number of independently dispersing entities, λ_n, and the proportion of ssDNA in a sample, p_{ss}, is practical only for mixtures prepared using globally denatured materials. However, comparing relative changes in λ_n as a function of the relative changes in p_{ss} provides indirect evaluation. Fig. 9 displays the relative changes between selected pairs of samples or assays, where the pairs are denoted as p and q:

$$\Delta\lambda_n(p,q) = (\lambda_n(p) - \lambda_n(q))/\lambda_n(q) = \lambda_n(p)/\lambda_n(q) - 1 \quad (6)$$

and

$$\Delta p_{ss}(p,q) = p_{ss}(p) - p_{ss}(q). \quad (7)$$

While subject to comparison-specific biases, the observed λ_n increase as p_{ss}(ddPCR Enzyme) increases. The data are

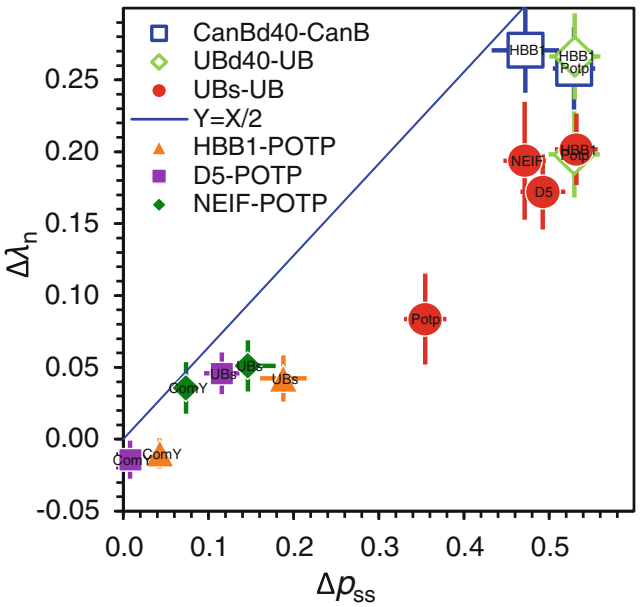

Fig. 9 Change in λ_n as a Function of Change in p_{ss}. Symbols represent the observed $\{\Delta p_{ss}, \Delta\lambda_n\}$ for representative samples and PCR assays. The larger symbols to the upper right represent differences induced by shaking (UBs) and heat-denaturing (CanBd40 and UBd40) as assessed with (POTP,HBB1,D5,NEIF). The smaller symbols to the lower left represent differences between the (HBB1,D5,NEIF) and POTP assays of the same aliquots of samples ComY and UBs. Crosses represent approximate ± standard uncertainties. The blue line represents: $\Delta\lambda_n = p_{ss}/2$

compatible with the ideal linear relationship of zero intercept and 0.5 slope.

Strandedness and the ComY-measurement discordance

If ssDNA content is the root cause for the ComY [nDNA] discordance, bias-correction should bring the estimates into substantial agreement. We believe that the mean of the four p_{ss}(ddPCR Enzyme) assays considered in this report provides an appropriate estimate: $p_{ss} = 0.18(2)$.

The four-assay mean λ_n for an $F = 1/50$ dilution of ComY is 0.239(4) entities per droplet. These assays have each been determined to amplify $r = 1$ target per human haploid genome. The droplet volume at the time of these measurements is believed to have been $V = 0.74(1)$ nL. From the conversion equation $[\text{nDNA}] = 3.301\lambda/(FVr) = 53(2)$ ng/μL [3]. Since [nDNA] is proportional to λ and the expected relative change in $\lambda = p_{ss}/2$, the expected bias is equal to $53(2) \times (0.18(2)/2) = 4.8(6)$ ng/μL and the bias-corrected [nDNA] is equal to 53(2) - 4.8(6) = 48(2) ng/μL. This encompasses the NaOH-denatured A_{260}-based estimate of 47.6(1) ng/μL.

Conclusion

Establishing the metrological traceability of dPCR results requires evaluating relevant sources of bias and, if necessary, correcting the value and/or expanding the uncertainty [27]. We believe that use of the p_{ss}(ddPCR Enzyme) assays described in this report will enable documenting the absence of significant ssDNA content bias - or correcting for it if required - in future ddPCR assessments of candidate human [nDNA] reference materials. In contrast to the ≈18% bias for the ComY commercial qPCR calibrant, the p_{ss}(ddPCR Enzyme) estimates for most of the NIST-extracted source materials studied are less than 4% and all are less than 6%. This suggests that the strandedness bias in ddPCR λ_n values for similarly extracted materials will be no more than 2% to 3%.

Funding information This work was supported in part by the NIST Special Programs Office project Forensic DNA.

Data availability The summary ddPCR and cdPCR data used are presented as Tables S1 and S2 in the ESM. The aA, aB, and aC materials are components A, B, and C of SRM 2372a and are available for purchase from NIST through https://www.nist.gov/srm. The spreadsheet-based cdPCR Staircase analysis system is available on request from the corresponding author.

Compliance with ethical standards

Conflicts of interest/Competing interests The authors declare that they have no conflict of interest nor competing interests.

Ethics approval All work presented has been reviewed and approved by the National Institute of Standards and Technology Human Subjects Protections Office. This study was determined to be "not human subjects research" (often referred to as research not involving human subjects) as defined in U. S. Department of Commerce Regulations, 15 CFR 27, also known as the Common Rule (45 CFR 46, Subpart A), for the Protection of Human Subjects.by the NIST Human Subjects Protection Office and therefore not subject to oversight by the NIST Institutional Review Board.

References

1. Kline MC, Duewer DL. Evaluating digital polymerase chain reaction for the quantification of human genomic DNA: lifting the traceability fog. Anal Chem. 2017;89(8):4648–54. https://doi.org/10.1021/acs.analchem.7b00240.
2. Dagata JA, Farkas N, Kramer JA. Method for measuring the volume of nominally 100 μm diameter spherical water-in-oil emulsion droplets. NIST Special Publication (NIST SP) 260–184. 2016. https://doi.org/10.6028/NIST.SP.260-184.
3. Duewer DL, Kline MC, Romsos EL, Toman B. Evaluating droplet digital PCR for the quantification of human genomic DNA: converting copies per nanoliter to nanograms nuclear DNA per microliter. Anal Bioanal Chem. 2018;410(12):2879–87. https://doi.org/10.1007/s00216-018-0982-1.
4. Romsos EL, Kline MC, Duewer DL, Toman B, Farkas N. Certification of standard reference Material® 2372a human DNA quantitation standard. NIST Special Publication (NIST SP) 260-189, 2018. https://doi.org/10.6028/NIST.SP.260-189.
5. NIST certificates of analysis are available through the NIST standard reference materials homepage. https://www.nist.gov/srm. Accessed 26 May 2020.
6. Bhat S, Curach N, Mostyn T, Bains GS, Griffiths KR, Emslie KR. Comparison of methods for accurate quantification of DNA mass concentration with traceability to the international system of units. Anal Chem. 2010;82:7185–92. https://doi.org/10.1021/ac100845m.
7. Sanders R, Huggett JF, Bushell CA, Cowen S, Scott DJ, Foy CA. Evaluation of digital PCR for absolute DNA quantification. Anal Chem. 2011;83:6474–84. https://doi.org/10.1021/ac103230c.
8. Sambrook J, Russell DW. Molecular cloning: a laboratory manual. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press; 2001.
9. ISO 21571:2005; Foodstuffs – Methods of Analysis for the Detection of Genetically Modified Organisms and Derived Products – Nucleic Acid Extraction, Annex B Methods for Quantitation of Extracted DNA; International Standards Organization: Geneva, Switzerland: 2005, pp. 34–36.
10. Wilson PJ, Ellison SLR. Extending digital PCR analysis by modelling quantification cycle data. BMC Bioinformatics. 2016;17:421. https://doi.org/10.1186/s12859-016-1275-3.
11. Kline MC, Duewer DL. Evaluation of methods for assessing the proportion of single stranded nuclear DNA in human blood extracts. NIST Special Publication (NIST SP) 1200-27, 2019. https://doi.org/10.6028/NIST.SP.1200-27.
12. Miller SA, Dykes DD, Polesky HF. A simple salting out procedure for extracting DNA from human nucleated cells. Nucleic Acids Res. 1988;16(3):1215. https://doi.org/10.1093/nar/16.3.1215.
13. Cavaluzzi MJ, Borer PN. Revised UV extinction coefficients for nucleoside-5′-monophosphates and unpaired DNA and RNA. Nucleic Acids Res. 2004;32(1):e13. https://doi.org/10.1093/nar/gnh015.
14. Nwokeoji AO, Kilby PM, Portwood DE, Dickman MJ. Accurate quantification of nucleic acids using hypochromicity measurements

in conjunction with UV spectrophotometry. Anal Chem. 2017;89(24):13567–74. https://doi.org/10.1021/acs.analchem.7b04000.
15. Volkov SN, Danilov VI. Study of 1st and 2nd absorption-band hypochromism in natural DNA. FEBS Lett. 1976;65(1):8–10. https://doi.org/10.1016/0014-5793(76)80609-6.
16. Tinoco I. Hypochromism in polynucleotides. J Am Chem Soc. 1960;82(18):4785–90. https://doi.org/10.1021/ja01503a007.
17. Sykes PJ, Neoh SH, Brisco MJ, Hughes E, Condon JL, Morley A. Quantitation of targets for PCR by use of limiting dilution. BioTechniques. 1992;13(3):444–9.
18. Duewer DL, Kline MC, Romsos EL. Real-time cdPCR opens a window into events occurring in the first few PCR amplification cycles. Anal Bioanal Chem. 2015;407(30):9061–9. https://doi.org/10.1007/s00216-015-9073-8.
19. Kline MC, Romsos EL, Duewer DL. Evaluating digital PCR for the quantification of human genomic DNA: accessible amplifiable targets. Anal Chem. 2016;88(4):2132–9. https://doi.org/10.1021/acs.analchem.5b03692.
20. Zhou Y, Paull TT. Direct measurement of single-stranded DNA intermediates in mammalian cells by quantitative polymerase chain reaction. Anal Biochem. 2015;479:48–50. https://doi.org/10.1016/j.ab.2015.03.025.
21. Nishigaki K, Kaneko Y, Wakuda H, Husimi Y, Tanaka T. Type II restriction endonucleases cleave single-stranded DNAs in general. Nucleic Acids Res. 1985;13(16):5747–60.
22. Thompson M. Uncertainty functions, a compact way of summarising or specifying the behaviour of analytical systems. TrAC. 2011;30(7):1168–75. https://doi.org/10.1016/j.trac.2011.03.012.
23. Bikard D, Loot C, Baharoglu Z, Mazel D. Folded DNA in action: hairpin formation and biological functions in prokaryotes. Microbiol Mol Biol Rev. 2010;74(4):570–88. https://doi.org/10.1128/MMBR.00026-10.
24. Ripley BD, Thompson M. Regression techniques for analytical bias. Analyst. 1987;112:377–83. https://doi.org/10.1039/AN9871200377.
25. Analytical Methods Committee. Linear functional relationship estimation by maximum likelihood. https://www.rsc.org/Membership/Networking/InterestGroups/Analytical/AMC/Software/FREML.asp. Accessed 26 May 2020.
26. Yakovchuk P, Protozanova E, Frank-Kamenetskii MD. Base-stacking and base-pairing contributions into thermal stability of the DNA double helix. Nucleic Acids Res. 2006;34(2):564–74. https://doi.org/10.1093/nar/gkj454.
27. Magnusson B, Ellison SLR. Treatment of uncorrected measurement bias in uncertainty estimation for chemical measurements. Anal Bioanal Chem. 2008;390(1):201–13. https://doi.org/10.1007/s00216-007-1693-1.