View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by LJMU Research Online

Developmental validation of the ParaDNA® Intelligence System—A novel approach to DNA profiling

Stephen Blackman [a,*], Nick Dawnay [a], Glyn Ball [a], Beccy Stafford-Allen [a],

Nicholas Tribble [a], Paul Rendell [a], Kelsey Neary [b], Erin K. Hanson [b], Jack Ballantyne [b,c],

Beatrice Kallifatidis [d], Julian Mendel [d,e], DeEtta K. Mills [d,e], Simon Wells [a]

[a] Product Development Group, LGC Ltd., Culham Science Centre, Abingdon OX14 3ED, UK

[b] National Center for Forensic Science, PO Box 162367, Orlando, FL 32816-2367, USA

[c] Department of Chemistry, University of Central Florida, PO Box 162366, Orlando, FL 32816-2366, USA

[d] International Forensic Research Institute, Florida International University, Miami, FL, USA

[e] Department of Biological Sciences, Florida International University, Miami, FL, USA

Abstract

DNA profiling through the analysis of STRs remains one of the most widely used tools in human identification across the world. Current laboratory STR analysis is slow, costly and requires expert users and interpretation which can lead to instances of delayed investigations or non-testing of evidence on budget grounds. The ParaDNA1 Intelligence System has been designed to provide a simple, fast and robust way to profile DNA samples in a lab or field-deployable manner. The system analyses 5-STRs plus amelogenin to deliver a DNA profile that enables users to gain rapid investigative leads and intelligent prioritisation of samples in human identity testing applications. Utilising an innovative sample collector, minimal training is required to enable both DNA analysts and nonspecialist personnel to analyse biological samples directly, without prior processing, in approximately 75 min. The test uses direct PCR with fluorescent HyBeacon1 detection of STR allele lengths to provide a DNA profile. The developmental validation study described here followed the Scientific Working Group on DNA Analysis Methods (SWGDAM) guidelines and tested the sensitivity, reproducibility, accuracy, inhibitor tolerance, and performance of the ParaDNA Intelligence System on a range of mock evidence items. The data collected demonstrate that the ParaDNA Intelligence System displays useful DNA profiles when sampling a variety of evidence items including blood,

saliva, semen and touch DNA items indicating the potential to benefit a number of applications in fields such as forensic, military and disaster victim identification (DVI).

Keywords: ParaDNA, Direct PCR, Rapid DNA, STR analysis, Triage, Crime scene

1. Introduction

The analysis of short tandem repeats (STRs) is the primary means used today for human identification and forensic DNA testing [1,2]. Current STR analysis involves extraction, purification and quantification of DNA from forensic evidence, polymerase chain reaction (PCR) amplification of STRs and detection of alleles through size separation in capillary electrophoresis (CE). These lab processes require highly trained personnel and are both time consuming and expensive [3]. The high demand for DNA evidence in criminal investigations can lead to large backlogs of samples requiring STR analysis, and budgetary requirements may mean that only a limited number of samples are processed for each case [4]. This inevitably leads to delays in arrests and convictions, and potential DNA evidence can often be overlooked, meaning criminals may be left free to commit further offences [5].

There have been a number of attempts to reduce DNA evidence backlogs through use of direct PCR [6], automation [7] and rapid DNA technologies [8,9]. Rapid DNA technologies as defined by the FBI [10] currently involve systems that essentially miniaturise and automate the existing processes (DNA extraction, PCR and size separation) [11]. However, these systems are expensive to purchase and operate and also still require technical staff to run and interpret the results, making them inappropriate for screening of large numbers of evidence items, especially at a crime scene. The tactical use of novel forensic platforms and processes that allow DNA analyses of human tissue and forensic samples would allow investigators to rapidly act on investigative leads and prioritise samples for downstream analyses. Such solutions would offer benefits in a wide variety of fields including traditional forensic investigations, military investigations and DVI [12].

One proposed approach utilises the ParaDNA system of instruments, software and tests for the genetic analysis of biological samples designed for non-expert users. The ParaDNA Intelligence Test has been developed to allow direct DNA analysis of a range of sample types interrogating 5 STR loci, plus amelogenin to deliver a DNA profile. It uses PCR amplification followed by HyBeacons probe detection of allele lengths in a closed tube system [13]. The ParaDNA Intelligence Test is designed to deliver actionable intelligence in a lab or at a crime scene by laboratory analysts or even

appropriately trained non-expert users (e.g. crime scene investigators) in approximately 75 min. This paper describes the developmental validation of the ParaDNA Intelligence Test performed at LGC Ltd., and independent studies at University of Central Florida (UCF) and Florida International University (FIU). Validation studies were performed according to the revised guidelines published by the Scientific Working Group on DNA Analysis Methods (SWGDAM) [14] and tested the sensitivity, reproducibility, accuracy, inhibitor tolerance, and performance of the ParaDNA Intelligence System on a range of mock evidence items. The results demonstrate that the ParaDNA Intelligence System generates useful DNA profiles when sampling a variety of evidence items including blood, saliva, semen and touch DNA items that could conceivably support operational and investigational improvements.

2. Materials and methods

2.1. ParaDNA Intelligence System

The ParaDNA Intelligence System comprises two parts: an instrument designed for DNA amplification and fluorescent detection including software to control the process and interpret the results, and a consumable kit comprising of a sample collector and a ready-to-go reaction plate containing the biochemistry required for amplification and detection of the DNA profile. Many of the components are the same as those described for the ParaDNA Screening Test [15] and are shown in Supplementary material Fig. 1. The sample collector, reaction plate and instrumentation are common to both the ParaDNA Screening and Intelligence Systems. The only differences are the tests (biochemistry) in the plates and the way the software interprets the results.

The ParaDNA Intelligence Test (LGC, PARA-030) is a custom-designed 4-well plate pre-loaded with all the reagents required for PCR and melt curve detection in four independent chambers. The test uses HyBeacon Technology [13,16] to amplify and detect 5 STR loci: D3S1358, D16S539, D8S1179, D18S51 and TH01 and also the amelogenin gender marker. Concordance of the ParaDNA Intelligence Test to a CE-based system (AmpFlSTR1 SGM Plus1, Life Technologies, Paisley, UK) has been previously described [17]. During this validation, different studies examine either the whole ParaDNA Intelligence System including sampling (e.g. case-type study) or only the ParaDNA Intelligence Test, bypassing the potential variability introduced by sampling (e.g. sensitivity study). The LGC validation studies were performed combining data from multiple instruments and operators, UCF and FIU used one instrument and two operators each.

In addition to the lab-based ParaDNA Screening Instrument described previously [15], LGC have developed a mobile ParaDNA Field Portable Instrument (LGC, PARA-020; Supplementary material Fig. 1d). Both versions of the instrument contain four thermal cyclers and fluorescence detection

systems to process up to four samples independently at any time. The Field Portable Instrument is robust, can operate on battery power and has a built-in PC to control the instrument, store data or search/compare profile data. Both instruments can run either the ParaDNA Screening Test or the ParaDNA Intelligence Test.

The ParaDNA software controls the instrument, analyses the data and displays the profile results (Supplementary material Fig. 1e). The analysis protocol associated with each ParaDNA test is automatically selected in preparation for the run; the operator is only required to provide a sample name and load the reaction plate prior to starting the analysis. The analysis protocol is comprised of PCR amplification and melt curve detection. During the melt, changes in fluorescence are recorded as the sample is heated from 20 C to 70C. A sharp drop in fluorescence is expected as the HyBeacon probe dissociates from the target STR (Supplementary material Fig. 2a). The temperature at which this transition occurs is referred to as the melting temperature (TM) and reflects the length of the amplified allele that is bound to the probe.

Deriving a profile from the ParaDNA Intelligence Test melt curve data is a fully automated process that involves quantifying the fluorescence change associated with each of the possible allele contributions. This analysis is based on an understanding of how the fluorescence of bound and unbound probe varies with temperature. The quantitative allelic data are compared to a series of thresholds that are designed to reject system noise and stutter and thereby only retain genuine allelic contributions (Supplementary material Fig. 2b). Note that the thresholds are derived from an extensive set of training data that consist of more than 4000 samples.

On occasion, the largest spurious contributions due to system noise and stutter are bigger than the smallest genuine contributions. This is especially true when smaller quantities of DNA are introduced to the reaction plate and naturally leads to the creation of an uncertainty band spanning the overlap region. Data points falling in an uncertainty band will prevent confident allele calls from being made. Finally, the allele calls from each melt curve in the four separate wells are combined to generate a five STR plus amelogenin profile for the sample analysed. Unless otherwise stated, all data in this study were analysed using ParaDNA software version 1.1.2.8.

2.2. Samples

A summary of samples used during the ParaDNA Intelligence Test developmental validation is shown in Supplementary material Table 1, the studies they relate to are summarised in Supplementary material Table 2. Panels of human DNA samples were obtained from Public Health, England (PHE, Salisbury, UK) as DNA purified from lymphoblast transformed cell lines (HRC-1 to HRC-5 and EDP-1).

DNA samples were quantified in-house using Plexor1 HY (Promega, Madison, WI) following the manufacturer's protocol. Additional control DNA samples were obtained from National Institute of Standards and Technology (NIST, Gaithersburg, MD) (2391c) and were used according to the manufacturer's quantification values. DNA samples were chosen to cover as many allele combinations in the ParaDNA intelligence test range as practicable. Animal DNA samples were obtained from PHE and microbial DNA obtained from ATCC (LGC Standards, Teddington, UK); both used the manufacturer's quantification values. Animal and primate blood samples were purchased from commercial sources: Hemostat Laboratories (Dixen, CA), Bioreclamation (New York), and Innovative Research, Inc. (Novi, MI). Ferret blood samples were obtained from Marshall Farms (New York). Mock evidence samples were provided by volunteers from LGC, FIU or UCF staff and all body fluid samples were collected in accordance with procedures approved by each institution's ethics/review board. Mock evidence samples were prepared on a number of substrates as detailed in the relevant study outlines below and in Supplementary material Table 1. Negative controls comprised DNA-free low-EDTA TE (IDT, Leuven, Belgium) or DNA-free tissue culture water (Sigma, Gillingham, UK) or DNA-free (Ethylene Oxide treated; Synergy Health, Bradford, UK) evidence items.

2.3. Species specificity

To characterise the ParaDNA Intelligence Test, a range of DNA from non-human sources (cat, dog, horse, rabbit, rat, pig, chicken, cow, chimpanzee, *Escherichia coli, Staphylococcus aureus* and *Candida albicans*) was tested in triplicate at 1 ng input DNA per well (4 ng total per plate). In addition, blood samples from alligator, cat, cow, deer, dog, ferret, frog, guinea pig, horse, mouse, rabbit, turtle, Rhesus Monkey, African Green Monkey, Chimpanzee and Baboon were tested by UCF by sampling dried blood stains on either 903 paper (GE, Little Chalfont, UK), filter paper or cotton cloth using the ParaDNA Sample Collector.

2.4. Sensitivity

To assess the sensitivity of the ParaDNA Intelligence Test, a serial dilution of saliva samples was tested. Five separate saliva samples were collected and, before testing on ParaDNA, 200ml was extracted with the QIAamp1 DNA mini kit (QIAGEN, Manchester, UK) using the blood protocol with a single 200ml elution and the DNA quantified as above to enable normalisation of sample concentration. Saliva samples were then serially two-fold diluted in PBS and applied directly onto the nibs of ParaDNA Sample Collectors to give approximate final DNA amounts per reaction plate of

2 ng–31.25 pg (500–7.8 pg per well). Five replicates of each saliva sample at each DNA amount were analysed using ParaDNA intelligence test (25 data points at each DNA amount).

To enable individual testing laboratories to compare their own controlled sensitivity studies to this validation, six human genomic DNA stock solutions were two-fold serially diluted and applied to each of the four nibs of ParaDNA Sample Collectors to give final DNA amounts per reaction plate of 2 ng–31.25 pg (500–7.8 pg per well). Six replicates of each DNA sample at each DNA amount were analysed using ParaDNA intelligence test (36 data points at each DNA amount). Similar studies were performed at UCF who tested two DNA samples in triplicate between 4000 pg and 250 pg, and FIU who tested one DNA sample in triplicate between 8000 pg and 62.5 pg DNA per plate. All of the above sensitivity studies applied DNA or saliva to the nibs directly using a pipette. This is not the normal method of adding samples to the ParaDNA Intelligence Test, where samples are scraped from substrates, and care must be taken to ensure the dispensed volume is retained on the nib until it dries.

The sensitivity of the ParaDNA Intelligence Test was assessed using the number of accurate alleles displayed and how often full or partial profiles were obtained.

2.5. Repeatability and reproducibility

The repeatability of the ParaDNA Intelligence Test process was assessed by measuring the variance of results obtained by single users. The reproducibility of the ParaDNA Intelligence Test process between users was assessed by measuring the variance between results obtained by multiple (four) users with various experience levels (novice, forensically-trained, routine/expert user). Each user ran eight replicates across a range of three saliva concentrations (neat, 1:10 and 1:100; 50ml each) dispensed onto cotton swabs (Fisher Scientific: 23-400-114).

The inter-laboratory reproducibility of the ParaDNA sampling process was assessed by comparing data generated by staff at Florida International University (FIU) and the University of Central Florida (UCF). Two sets of ten replicate swabs spiked with 50ml of standard saliva dilutions (neat, 1:10 and 1:100) in PBS were prepared by LGC using the same saliva sample and one set shipped to each of the participating laboratories for testing.

2.6. Robustness

The ParaDNA Intelligence System uses samples directly, without purification, so is potentially vulnerable to differences in substrates or chemical insults. It may also be used on degraded DNA samples. The robustness of the ParaDNA Intelligence System was challenged using various mocked-up inhibited and aged samples. Six replicates of buccal swabs collected from volunteers who had consumed various drinks were ParaDNA sampled. The beverages chosen (red wine, grape juice, green tea, beer, coffee) were all known to contain different levels of polyphenolic compounds [18], such as tannic acid as well as other substances believed to inhibit PCR reactions. Additionally, six replicates of swabs containing 50ml of a neat saliva sample were contaminated with different levels of soil representing a light, medium and heavily contaminated sample before ParaDNA sample collection. Triplicate samples of blood and saliva stored at room temperature for 1 week on polyester, carpet and denim were also tested to assess the impact of these various substrates.

The ParaDNA Intelligence Test inhibitor tolerance was further tested using the following model systems to simulate compromised samples that the system may encounter. In each case, the template was purified DNA at a total of 1 ng per well and was added to the mix during preparation of chemistry as were the inhibitors (final concentrations in the reaction; n = 3). Tannic acid (12.5–150 ng/ml), humic acid (2.5–25 ng/ml) and hemin (5–50mM) are often used in validation studies to simulate natural substances such as tea and fruit, soil and degraded blood, respectively [21,6]. In addition, NaCl (5–100mM) and sodium phosphate buffer (5–100mM) were used to simulate substrates which could possibly affect the melt temperature of the HyBeacon probes in the ParaDNA Intelligence System.

To assess the performance of the ParaDNA Intelligence Test with aged samples, blood and saliva samples dried on glass were stored at room temperature for 18 months and derived profiles compared against similar samples stored for just 24 h. Blood and saliva samples on cotton were also tested after storage at 37 C for 1 week and 1, 6 and 12 months. To test some real life scenarios, blood samples on cotton were also tested after 1–2 weeks stored outdoors (protected from rain) or in a car (exposed to high temperatures).

2.7. Mock case samples

The ParaDNA Intelligence Test is designed to amplify human DNA from an evidence item or swab and give an accurate profile call which may be used for sample triage or profile comparisons. The types of samples encountered by the end-user may be very diverse and involve a number of case scenarios. Mocked-up case samples were chosen to represent those typically expected to cover high,

medium and low amounts of template and were collected from LGC, UCF and FIU staff members with the donor's consent (n=9–40 depending on sample type and testing site). Samples tested comprised of four categories: (i) blood (on glass and denim); (ii) saliva (on FTA1 cards (GE, Little Chalfont, UK), buccal swabs, smoked cigarettes and plastic drinks bottles); (iii) semen (on cotton); (iv) touch DNA (fingerprints, screwdrivers and mobile phones).

Sampling of these items followed LGC's ParaDNA User Guide [22]. Most mock samples were sampled directly, using the ParaDNA Sample Collector applied directly to the sample. The exception was blood on glass, which was sampled indirectly from a wet cotton swab that had been used to collect the sample from the glass to prevent the distribution of flakes generated through sampling such a friable substance on a non-porous substrate.

Developmental validation samples (both those subjected to ParaDNA testing and unsampled) were processed with SGM Plus in the LGC Forensics Scene of Crime (SoC) DNA laboratories to enable comparison of a standard, validated lab DNA profiling process with the ParaDNA Intelligence Test. Samples (blood and saliva) typically yielding high levels of DNA template were not progressed through laboratory profiling as these samples were assumed to deliver full DNA profiles during this study. Additionally, UCF and FIU performed PowerPlex® 16HS (Promega) and Identifiler Plus® (Life Technologies) analysis on fingerprint and mobile phone samples to compare the performance of the ParaDNA intelligence test against kits routinely used in their laboratories.

Plexor HY quantification data from ParaDNA sampled bottles, cigarettes, mobile phones, fingerprints and tools and corresponding replicate samples which had not undergone ParaDNA analysis were compared to assess the effect of using the ParaDNA Intelligence System on downstream laboratory processing.

2.8. Mixture studies

As DNA samples from multiple contributors are often encountered at a scene of crime it is important to understand how the ParaDNA Intelligence Test performs with mixed samples. There is no functionality in the ParaDNA Intelligence System for determination of the number of contributors, the major and minor contributor profiles or the contributor ratios. However, the ParaDNA Intelligence System is able to detect when mixtures are present, indicating as such to a user while still displaying a profile if enough alleles are detected. To characterise the ability of the ParaDNA Intelligence System to detect mixtures, single source human DNA was tested alongside mixed human DNA. Two quantified DNA samples were used to create mixtures in the following ratios—10:1, 7:1,

5:1, 3:1, 2:1, 1:1, 1:2, 1:3, 1:5, 1:7, 1:10. It was thought likely that mixtures would become more difficult to determine where the mixed components were at low levels of DNA, due to stochastic effects causing allele drop-out or unequal balance between alleles. Therefore mixtures at two different total DNA amounts (4 ng and 1 ng total per plate) were assessed. A similar mixture ratio experiment was performed using saliva samples to investigate if mixtures are interpreted differently with real cellular samples where drop-out rates may be different. To ensure similar amounts of DNA were combined in the saliva mixture study, first a range finder experiment was performed using ParaDNA Intelligence Tests and the allele counts obtained used to enable both saliva samples to be normalised to the 250 pg/ml level estimated from the sensitivity study.

2.9. PCR-based studies

The ParaDNA Intelligence Test detection system is different from traditional STR kits in that it detects allele length using melt analysis. To do so it requires not only primers but additional detection oligonucleotides (blockers and probes). To test the reliability and robustness of the amplification and detection components, DNA samples were run with chemistries where concentrations of components were used at ±20% from the standard formulation. Current manufacturing specifications make it unlikely that actual concentrations of mix components will ever be more than ±6% from specified values.

2.10. Accuracy

In a separate study, individual DNA samples from 381 UK Caucasian individuals were analysed using SGM Plus and the ParaDNA Intelligence Test with the derived STR profiles compared. The concordance demonstrated between the two systems, with reference to allele frequencies and the discriminatory power offered by the ParaDNA Intelligence Test, was shown to be 99.8% [17]. UK Caucasians are typically less variable than individuals of African origin and it might be suspected that slightly more discordance between ParaDNA and regular STR typing will be seen when individuals from more diverse populations are examined. The discordant allele calls were observed to come from some microvariants which ParaDNA cannot resolve, from rare alleles below the callable range of the test or due to resolution of the ParaDNA melt curve detection system [17]. The ParaDNA Intelligence Software searching and comparison functions use a series of rules to widen the search criteria (wildcards) in the case of microvariants [19]. Work to extend the working ranges of the ParaDNA Intelligence Test assays and detect alleles below the current test ranges is planned. To

assess accuracy of the ParaDNA Intelligence Test during the LGC validation studies, all ParaDNA profiles obtained were compared to expected profiles either provided by the supplier or assessed by SGM Plus analysis at LGC.

2.11. Data analysis

The ParaDNA measurements taken during the course of the developmental validation studies focused on the user outcome (profile displayed). Results are delivered to the user as recorded profile calls for each sample without displaying the underlying DNA melt data. An allele call may be shown in green (confident) or downgraded to a no call (by default this is displayed as (-) but is configurable). Samples where fewer than seven alleles are detected are more vulnerable to miscalls due to stochastic events in low template samples (data not shown); therefore the ParaDNA software downgrades any profiles where fewer than seven alleles were detected and displays “Insufficient DNA to determine a profile”. In determining whether a DNA mixture is present, the ParaDNA software detects the number of alleles at each locus including melt transitions that are close to the threshold values. If three or more alleles are detected at one locus the software does not display the alleles detected, instead displaying (- , -) and the locus is tagged as “more than 2 alleles exist”. If a third peak is detected which falls just short of the allele designation thresholds, then the locus is tagged as a “reduced confidence mixture”. If two or more loci are tagged as “more than 2 alleles exist” or “reduced confidence mixture” then the sample as a whole is declared a mixture by the software, while still displaying remaining allele calls at loci with only one or two alleles identified. Throughout this study, the metric “usable profiles” has often been used to define ParaDNA Intelligence Test performance. This metric is based on the software only displaying a profile if seven or more alleles are detected by the software as described above. These data can then be used for intelligence or triage purposes. However, loci where three or more alleles are detected (mixtures) do not have their alleles displayed but are included as the software counts these alleles towards the seven or more allele threshold. So it is possible to have a profile displayed in ParaDNA with fewer than seven alleles. Such profiles were included when tallying samples using the “Usable profiles” metric as they are still displaying potentially useful information. So a “usable profile” is defined as a profile where seven or more alleles are detected (not necessarily displayed). In theory, this could be only one allele displayed, but in practice mostly means seven alleles or more being displayed.

The melt curve detection system used in the ParaDNA Intelligence Test means a limited range of alleles can be detected for each STR [17]. The allele range at each locus detected and reported by the ParaDNA intelligence test is anticipated to detect and designate 94.13% of expected alleles at these loci, based on allele frequencies from sample populations [17]. A further 5.76% more rare alleles with repeat numbers above these ranges are also detected but cannot be resolved. Therefore the highest allele reported in each range is designated n+, indicating that it may represent the n allele or any larger undesignated alleles at that locus

3. Results and discussion

In many countries, DNA profiling is increasingly being used to generate databases not only of serious violent crime cases, but also non-violent offenders or suspects such as those arrested for drug possession, burglary or robbery. Current methods of DNA profiling involve DNA extraction/purification, PCR amplification of multiple loci followed by STR length determination through CE and expert interpretation of results. Although highly discriminatory, the process is slow, resource intensive and expensive. Many DNA labs have a large backlog of samples due to the high demand and low throughput leading to inevitable delays in criminal investigations [20]. The validation data presented here demonstrate that the ParaDNA Intelligence System can provide a robust and sensitive means of generating STR profiles rapidly and highlights the limits under which reliable results can be obtained.

3.1. Species specificity

Several non-human genomic DNAs and blood samples were tested for cross-reactivity with the ParaDNA Intelligence Test. With the genomic DNA samples, cross-reactivity was observed in all replicates of the chimpanzee samples, with between seven and ten alleles displayed (data not shown). The non-human blood samples only yielded a profile with one replicate of baboon blood, with 7 alleles displayed. This cross-reactivity with higher primate samples was expected as there is significant shared homology with human DNA sequences [23]. The remainder of the tested species gave no profiles with the ParaDNA Intelligence Test (cat, dog, horse, rabbit, rat, pig, chicken, cow, alligator, deer, ferret, frog, guinea pig, turtle, Rhesus Monkey, African Green Monkey, *E. coli, S. aureus* and *C. albicans*). Purified human DNA and blood positive controls gave full profiles.

3.2. Sensitivity

The case-type study in Section 3.7 assessed the sensitivity of the whole system with selected mocked-up sample types/substrates. However, more controlled measures of sensitivity of the ParaDNA Intelligence Test were required to allow users to compare to their own validations/verifications. Two studies were performed by directly dispensing samples onto the nibs of Sample Collectors (i.e. excluding sampling variation).

In the first study, a series of dilutions of purified genomic DNA samples were tested both at LGC and independently in labs at FIU and UCF. The results of these studies are presented in Fig. 1.

Results from the three labs were broadly similar and showed that all three labs achieved greater than 80% of samples giving a usable profile at 250 pg total DNA (62.5 pg per well). The lowest DNA amount where a usable profile was returned by the ParaDNA software, was 31.25 pg total (7.8 pg per reaction well).

Full ParaDNA profiles (12 alleles) were obtained down to 500 pg DNA total (125 pg per well) in the LGC data but one full profile was obtained at 250 pg total at FIU. More than 65% of samples at all three sites gave full profiles at 4 ng or greater total DNA (1 ng per well).

The ParaDNA Intelligence Test has not been optimised for use with purified DNA, so a second sensitivity study, using saliva as a more realistic sample type (albeit being added to nibs through pipetting rather than scraping a sample),was also tested. The saliva dilution data (Fig. 2) suggest that the ParaDNA Intelligence Test can deliver profiles down to 31.25 pg total input DNA and reliably give profiles down to 250 pg total input DNA. Full profiles were obtained in more than 90% of samples at 2 ng total input DNA whereas the lowest DNA amount where full profiles were observed in more than a single replicate was 250 pg total input DNA. No DNA extraction method is 100% efficient and the figures quoted in Fig. 2 are likely to be overestimations and so cannot be directly compared to the purified DNA values. The DNA quantification figures are based on the extraction method detailed in Section 2.4 and users should follow the method exactly if they wish to compare directly to this study.

3.3. Reproducibility/repeatability

The ParaDNA Intelligence Test process was tested by assessing the number of alleles displayed from replicate (n = 8) samples processed by single operators of a series of saliva dilutions on swabs (Fig. 3a). The low variance observed demonstrated the repeatability of the process. The reproducibility of

the ParaDNA Intelligence Test process was assessed by measuring the variance of results gained between four users with various experience (novice, SOCO/CSI, expert and routine user) (Fig. 3a) and, in a separate set of experiments, by comparing results from two independent labs (Fig. 3b). Reproducibility was demonstrated as no significant differences in the number of allele calls ($p > 0.05$ using Kruskal–Wallis test) observed between operators and between independent labs at any of the input DNA levels.

3.4. Robustness

The ParaDNA Intelligence Test is designed to be used on case type samples directly, without purification. To assess the robustness of the ParaDNA Intelligence Test, a number of studies were performed using mocked-up inhibited and degraded samples as well as model inhibitors often used in validation studies of other STR kits. The mean number of displayed alleles in the ParaDNA profiles for each mock inhibited sample (Fig. 4) suggests that the ParaDNA Intelligence Test has good tolerance to the type of inhibitors commonly found at the crime scene.

The ParaDNA Intelligence Test is not affected by any residual amounts of “real-world” polyphenolic compounds on buccal swabs collected immediately after donors had drunk the above beverages, with a t-Test showing none gave a statistically significantly different mean number of allele calls from the water positive control ($p > 0.05$) (Fig. 4a). This indicates that saliva deposits from individuals who have recently consumed one of these drinks are unlikely to suffer from inhibition. The ParaDNA Intelligence Test also showed no statistically significant difference between the positive control and light soil contamination, and still provided usable profiles (7 alleles detected) in some cases even when the sample was moderately contaminated with soil (Fig. 4b). In addition, there was no statistically significant difference between positive controls and blood and saliva samples stored at RT on polyester, carpet or denim for one week (not shown).

The data from the model inhibitor study indicated that the ParaDNA Intelligence Test yielded usable profiles in the presence of final concentrations of inhibitors in the assay up to 100 ng/ml (14mg total per plate) tannic acid, 10 ng/ml (1.4mg total) humic acid, and 10mM hemin (0.9mg total) (Supplementary material Fig. 3). Full profiles were obtained up to 50 ng/ml tannic acid, 5 ng/ml humic acid and 10mM hemin (Supplementary material Fig. 3).

To investigate the possibility of increased melt temperatures due to ionic strength of samples, a study was performed assessing the effects on allele calls of NaCl or sodium phosphate buffer (pH 7) on the ParaDNA Intelligence Test reactions. The study used a DNA sample homozygous in three out

of the five STR loci to increase the chance of observing miscalls rather than spurious mixtures. The results indicated that miscalls due to increased melt temperature were observed at salt concentrations of between 35 and 45mM NaCl (0.3 and 0.4mg total per plate) in the final reaction, just at the point where some samples fail to give a profile. Similar results were obtained in a study using a multivalent ion (sodium phosphate buffer), with some miscalls at 25mM (0.5mg total) whereas no sample gave a profile at the next concentration up (35mM, 0.65mg total) (Supplementary material Fig. 4).

The important thing to consider is how much salt is going to be introduced into the ParaDNA reaction mixes during normal use. This comes down to two things, what amount of salt is found in forensic sample types and how much contaminating salt might be transferred by the ParaDNA Sample Collector from the substrate. Some sample types commonly encountered might be expected to contain a high salt content, such as blood and semen, and so possibly vulnerable to peak shift and miscalls. However, these, and other sample types, are examined in the case-type samples section below and were unaffected by peak shift. In addition, small studies looking at sweaty clothing and buccal swabs from volunteers who had just eaten salty peanuts also showed no evidence of peak shift (not shown).

Another study assessing how much contaminating salt might be transferred in extreme circumstances indicated that it is possible to get miscalls and no profiles by sampling from dried cotton swabs previously saturated with 1MNaCl or 1Mphosphate buffer (these concentrations are about 50% greater than the total salt concentration found in sea water) or dried salt solutions on glass where crystals were clearly observed (not shown). However, there was zero transfer of the 1M NaCl or 1M phosphate buffer from dried cotton cloth or wet swabs. It seems highly unlikely that sampling using the ParaDNA Sample Collector from real forensic samples will lead to miscalls due to salt contamination as the sample would have to be contaminated with salt to an extent not expected to be found in crime scenes and just the right amount of salt needs to be transferred to fall in the region where miscalls occur. However, the possibility remains, so it is recommended that users show caution when sampling from substrates that are believed to be heavily contaminated by salt.

The performance of the ParaDNA Intelligence Test on aged samples was tested with blood and saliva samples on glass kept at room temperature for 18 months. A t-test of the data showed that the test performs equally well on both aged and fresh samples in terms of the mean number of alleles displayed ($p > 0.01$, Supplementary material Fig. 5). Experiments performed at UCF with blood and saliva samples on cotton aged for up to one year at 37 C also showed no loss of performance

(Supplementary material Fig. 6). UCF also tested exposure to real environments with no loss of performance after samples had been stored for two weeks in the trunk of a car or outside exposed to sun, temperature and humidity (but protected from rain).

3.5. Case-type samples

The profile call success rates of the ParaDNA Intelligence Test was assessed by LGC, UCF and FIU against a range of sample types donated from volunteers with known DNA profiles (reference buccal samples analysed with SGM Plus or Identifiler Plus).

All samples, including semen, were analysed with the ParaDNA Intelligence Test without the need for prior lysis and extraction. Fig. 5 shows the percent of samples displaying allele calls for each sample type group processed at LGC. The system performed consistently well when sampling from items expected to contain high amounts of DNA. The blood and semen samples generated profiles containing seven or more alleles in 100% of samples tested. The saliva samples (buccal swabs, FTA, drinks bottles and cigarettes) were also 100% successful, except in the case of cigarettes where there is likely to be a substantially lower level of DNA, depending on the donor of the cigarettes.

As expected, items that were found to contain a low level of DNA (termed 'touch DNA'; fingerprints, screwdrivers and mobile phones), yielded a lower number of samples with a profile of seven or more displayed alleles (38%). Mobile phones typically yielded more profiles than the mocked-up fingerprint and screwdriver samples (shown in Fig. 6).

Fig. 6 shows the percent of samples displaying a usable ParaDNA profile for the individual sample types through ParaDNA Intelligence Test analysis at the different testing sites (only LGC analysed screwdrivers and saliva on FTA; FIU did not analyse semen). The success rates achieved with ParaDNA Intelligence Test on different sample types was broadly similar between labs with high DNA samples (blood, buccal, semen, drinks bottles and FTA card) yielding profiles in 90–100% of samples. There was some variability observed between the sites when testing drinks bottles, cigarettes and fingerprints, probably indicating variability in donors and sample mock-up. The touch items gave poorer results as would be expected from these low DNA samples, particularly when the sample was spread over a larger area, such as with screwdrivers.

The ParaDNA outcomes were compared with results from analysing a replicate set of mocked-up evidence items using the SGM Plus kit in a forensic DNA laboratory (Fig. 7). It was shown that the ParaDNA Intelligence Test provided useful information in most cases where a laboratory result was expected. The exceptions were at very low DNA levels where the evidence was spread over a larger

area and not easily located (for example, tools). Similar observations were made at UCF and FIU when comparing mobile phones and fingerprint performance to Identifiler Plus and PowerPlex 16HS results from the same samples (Supplementary material Fig. 7). There were no discordant alleles observed between ParaDNA and SGM Plus in this study.

3.6. Impact of ParaDNA sampling on downstream lab analysis

Plexor HY quantification was used to measure the amount of DNA recovered using standard cotton swab sampling from low level evidence types. One set of evidence samples were previously sampled and analysed by the ParaDNA process and a replicate set of samples not processed by ParaDNA were also quantified. The results show that, apart from fingerprints, it was not possible to detect a significant negative impact from use of the ParaDNA Sample Collector on the amount of DNA subsequently extracted from these items (Supplementary material Fig. 8).

Fingerprints were the only sample type where a significant amount of the DNA present was apparently lost through ParaDNA sampling. These samples have a low amount of DNA, but are readily located in a small area. It is likely that the ParaDNA Sample Collector is removing proportionally more DNA from this sample type. This result was expected, leading to the sampling recommendation in the ParaDNA Intelligence Test user manual: "*CAUTION: Investigators should use their own judgement to assess when direct sampling may, or may not be appropriate. If there is little other evidence available, the investigator may choose to indirectly sample the evidence item in order to preserve as much material as possible, should further analyses be required*."

3.7. Mixture studies

The ParaDNA software detects the designated alleles at each locus and also any peaks that almost passed thresholds but were not confidently called as alleles as described in Section 2.11.

Fig. 8 shows the mixture calling capability of the ParaDNA Intelligence Test with mixed high and low concentrations of purified DNA and saliva samples. At 1:1, 2:1 and 3:1 ratios, 100% of purified DNA samples were identified as mixtures. The saliva mixtures were less successfully identified with one sample at 1:1 not being identified as a mixture. At 10:1, 20–40% of purified DNA samples are declared mixtures.

No single source saliva or DNA sample was declared a mixture in this experiment. Analysing the case-type samples in Section 3.5 reveals that 0.79% of samples gave a spurious mixture declaration (1

sample out of 127 returning a profile was spuriously called a mixture, 4 samples appeared to be genuine mixtures but were not detected).

3.8. PCR-based studies

The number of alleles displayed by the ParaDNA Intelligence Test did not change significantly when assay components (primers, blockers & probes, non-oligo reagents, $MgCl_2$, dNTPs and polymerase) were increased or decreased by 20% from the standard concentrations (not shown). The manufacturing maximum variance of 6% is well within this figure.

3.9. Accuracy

A specific study looking at concordance of the ParaDNA Intelligence Test to SGM Plus profiling under optimal conditions has been published separately [17]. The samples analysed with the ParaDNA Intelligence Test by LGC in this validation study were also compared to expected profiles (either provided by supplier or through SGM Plus analysis of optimal samples at LGC). Less than 1% of samples analysed with the ParaDNA Intelligence Test throughout the LGC study showed instances of genuine drop-in (i.e. not in stutter positions) being called in the ParaDNA Intelligence Test software, all from low level DNA samples. These occurred either when the expected profile was homozygote with the drop-in leading to a heterozygote call, or instances of drop-out of one expected heterozygote allele and drop-in of an unexpected. The level of drop-in experienced by users will vary depending on quality and source of sample being analysed. Spurious homozygote calls were also observed when one heterozygote peak had dropped out. However, the vast majority of these were in purified DNA samples which are more prone to drop-out than cellular samples. Only one cellular sample gave a spurious homozygote call out of 553 total (0.2%). There were two more samples displaying miscalled alleles in the total data set (0.2%) that were due to performance issues of the ParaDNA Intelligence System. None of the above issues were due to systematic non-concordance with SGM Plus as good quality samples from the same source DNA gave fully concordant profiles. The total miscall rate (excluding purified DNA drop-out but including drop-in) falls within the reported level for the ParaDNA Intelligence Test (1.2%).

4. Summary

The data presented here demonstrate that the ParaDNA Intelligence System can provide similar success rates to those demonstrated by commonly used STR profiling products. No prior lysis step was required for any sample type including semen, confirming previous observations with the ParaDNA Screening Test [24]. The production of DNA profiles directly from a variety of substrates and in the presence of inhibitors was also observed, indicating the potential of the ParaDNA Intelligence Test to be used directly on a variety of evidence items. In addition, the ease of use of the ParaDNA Intelligence System by specialist and non-specialist users in several labs was demonstrated.

The ParaDNA Intelligence System is not designed to replace existing technology or processes, but provides an additional tool for developing investigative intelligence from expert or appropriately trained non-expert users. This may allow investigations to be accelerated and help inform decisions regarding resource allocation. Despite the reduced discrimination power of a ParaDNA profile compared to most modern lab-based STR kits, ParaDNA profiles can be used to interrogate national or local databases to identify unknown suspects or narrow down the list of suspects, particularly in conjunction with geographical filtering. Further pilot studies with external groups assessing the beneficial impact of such early intelligence are currently underway. All users intending to utilise the ParaDNA Intelligence System are recommended to perform their own operational or internal validation/verification studies prior to implementation.

Acknowledgements

The authors would like to thank Jim Thomson for reviewing the manuscript before submission and the following staff members for their contribution to the development of the ParaDNA intelligence test; Monika Panasiuk, Nicola Duxbury, Romana Ahmed, Sarah Naif, Daniel Leonard, Daren Clark, and Angela Cacioppo.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.fsigen.2015.04.009.

References

[1] J.M. Butler, Genetics and genomics of core short tandem repeat loci used in human identity testing, J. Forensic Sci. 51 (2) (2006) 253–265.

[2] J.M. Butler, Forensic DNA Typing: Biology, Technology, and Genetics of STR Markers, Academic Press, 2005.

[3] J.A. Nicklas, T. Noreault-Conti, E. Buel, Development of a fast, simple profiling method for sample screening using high resolution melting (HRM) of STRs, J. Forensic Sci. 57 (2012) 478–488.

[4] Florida Department of Law Enforcement, Crime Laboratory Evidence Submission Manual, Florida Department of Law Enforcement, 2009 http://www.iape.org/resourcesPages/IAPE_Downloads/Submission_Manuals/FDLE%202009%20EvidSubManual.pdf (accessed 13.11.2014).

[5] C. Asplen, W. McVey, Taking control. Less taxpayer funding, but more DNA testing, Forensic Mag. 9 (2) (2012) 27–30.

[6] K. Oostdik, J. French, D. Yet, B. Smalling, C. Nolde, P. Vallone, E.L.R. Butts, C.R. Hill, M.C. Kline, T. Rinta, A.M. Gerow, S.R. Allen, C.K. Huber, J. Teske, B. Krenke, M. Ensenberger, P. Fulmer, C. Sprecher, Developmental validation of the PowerPlex® 18D System, a rapid STR multiplex for analysis of reference samples, Forensic Sci. Int. Genet. 7 (2013) 129–135.

[7] J.Y. Liu, C. Zhong, A. Holt, R. Lagace, M. Harrold, A.B. Dixon, M.G. Brevnov, J.G. Shewale, L.K. Hennessy, AutoMate express™ forensic DNA extraction system for the extraction of genomic DNA from biological samples, J. Forensic Sci. 57 (2012) 1022–1030, doi:http://dx.doi.org/10.1111/j.1556-4029.2012.02084.x.

[8] S. Gangano, K. Elliott, K. Anoruo, J. Gass, J. Buscaino, S. Jovanovich, D. Harris, DNA investigative lead development from blood and saliva samples in less than two hours using the RapidHIT™ human DNA identification system, Forensic Sci. Int. Genet. 4 (1) (2013) e43–e44. http://dx.doi.org/10.1016/jfsigss.2013.10.022.

[9] K. Pretz (2013, February 4). New system to speed up DNA analysis [Press release].

[10] The Federal Bureau of Investigation. Rapid DNA or Rapid DNA Analysis. http://www.fbi.gov/about-us/lab/biometric-analysis/codis/rapid-dna-analysis (accessed 13.11.2014).

[11] D. Harris, New advances make DNA even more powerful as a crime-fighting tool, Evidence Mag. 11 (1) (2013) 6–10.

[12] M. Halpern, J. Gerdes, J. Habb, A. Kiavand, J. Ballantyne, E. Hanson (2011). Rapid STR prescreening of forensic samples at the crime scene. National Institute of Justice, document number 236434.

[13] D.J. French, R.L. Howard, N. Gale, T. Brown, D.G. McDowell, P.G. Debenham, Interrogation of short tandem repeats using fluorescent probes and melting curve analysis: a step towards rapid DNA identity screening, Forensic Sci. Int. Genet. 2 (2008) 333–339.

[14] Scientific Working Group on DNA Analysis Methods. Validation Guidelines for DNA Analysis Methods (2012). http://swgdam.org/SWGDAM_Validation_Guidelines_APPROVED_Dec_2012.pdf. (accessed 13.11.2014).

[15] N. Dawnay, B. Stafford-Allen, D. Moore, S. Blackman, P. Rendell, E.K. Hanson, J. Ballantyne, J. Kallifatidis b. Mendel, D.K. Mills, N. Nagy, S. Wells, Developmental validation of the ParaDNA1 Screening System – a presumptive test for the detection of DNA on forensic evidence items, Forensic Sci. Int. Genet. 11 (2014) 73–79.

[16] D.J. French, C.L. Archard, T. Brown, D.G. McDowell, HyBeacon probes: a new tool for DNA sequence detection and allele discrimination, Mol. Cell. Probes 6 (2001) 363–374.

[17] G. Ball, N. Dawnay, R. Stafford-Allen, M. Panasiuk, P. Rendell, S. Blackman, N. Duxbury, S. Wells, Concordance study between the ParaDNA1 Intelligence Test, a rapid DNA profiling assay, and a conventional STR typing kit (AmpFlSTR® SGM Plus®), Forensics Sci. Int. Genet. 16 (2015) 48–51.

[18] K. Mullen, S.C. Marks, A. Crozier, Evaluation of phenolic compounds in commercial fruit juices and fruit drinks, J. Agric. Food Chem. 55 (2007) 3148–3157.

[19] ParaDNA Application Note. Speculatively searching an external database with exported ParaDNA profiles. [Internet] 2014. [cited: 2014, Aug 14th]. http://paradna.lgcforensics.com/faqa/.

[20] D. Cantillon, K. Kopiec, H. Clawson. Evaluation of the impact of the Forensic Casework DNA Backlog Reduction Program, National Institute of Justice, (2008), grant number 2005TO191/2005PR550.

[21] J.J. Mulero, N.J. Oldroyd, M.T. Malicdem, L.K. Hennessy, Developmental validation of the AmpFISTR® SEfiler Plus™ PCR amplification kit: an improved multiplex with enhanced performance for inhibited samples, Forensic Sci. Int. Genet. 1 (2008) 121–122.

[22] ParaDNA user manual v1.3. Available from LGC.

[23] K. Prüfer, K. Munch, I. Hellmann, K. Akagi, J.R. Miller, B. Walenz, S. Koren, G. Sutton, C. Kodira, R. Winer, J.R. Knight, J.C. Mullikin, S.J. Meader, C.P. Ponting, G. Lunter, S. Higashino, A. Hobolth, J. Dutheil, E. Karakoç, C. Alkan, S. Sajjadian, C. R. Catacchio, M. Ventura, T. Marques-Bonet, E.E. Eichler, C. André, R. Atencia, L. Mugisha, J. Junhold, N. Patterson, M. Siebauer, J.M. Good, A. Fischer, S.E. Ptak, M. Lachmann, D.E. Symer, T. Mailund, M.H. Schierup, A.M. Andrés, J. Kelso, S. Pääbo, The bonobo genome compared with the chimpanzee and human genomes, Nature 486 (7404) (2012) 527–531.

[24] N.D. Tribble, J.A.D. Miller, N. Dawnay, N.J. Duxbury, Applicability of the ParaDNA1 Screening System to seminal samples, J. Forensic Sci. 60 (2015) 690–692.

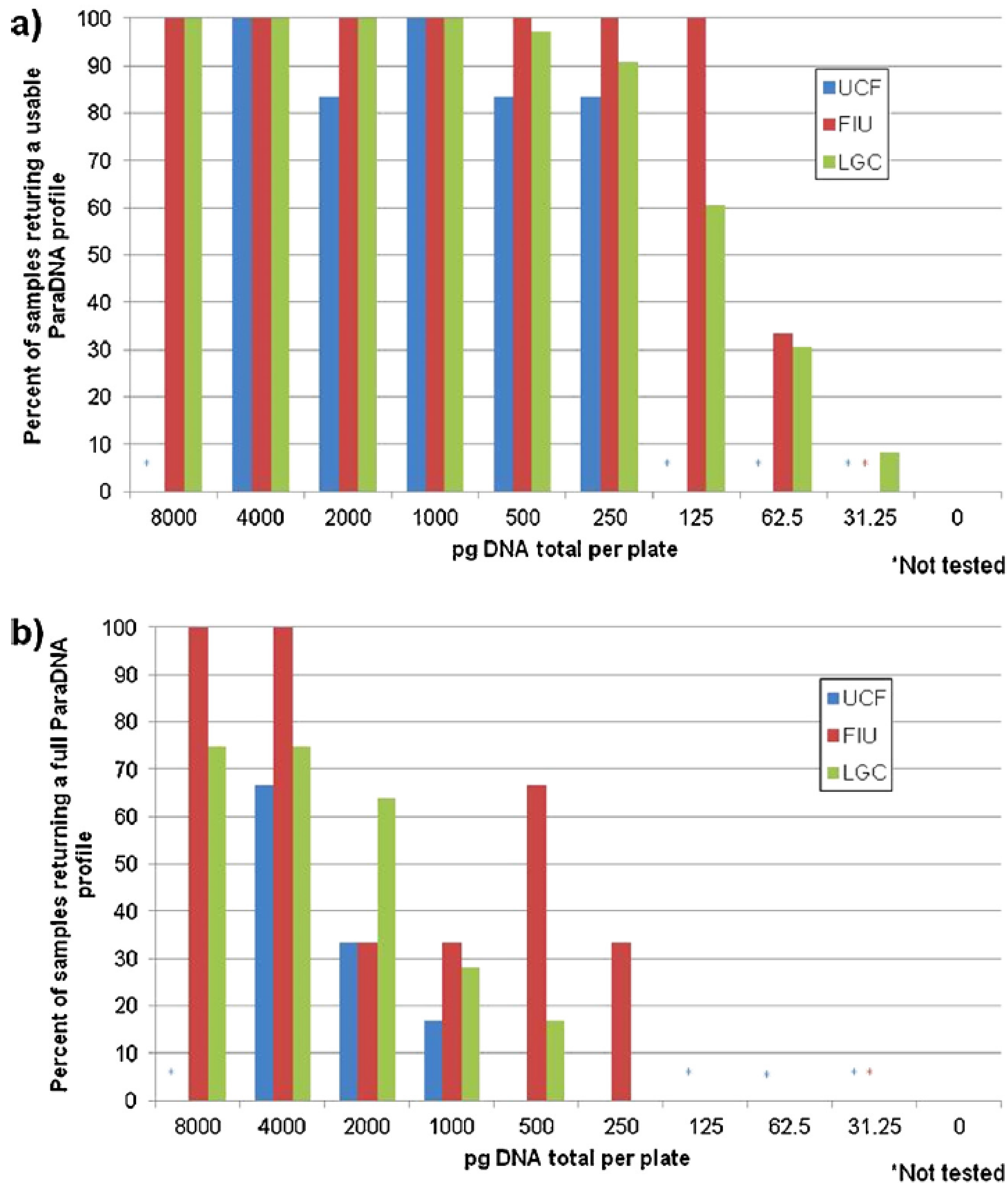

Fig. 1. Sensitivity plot of the ParaDNA Intelligence Test with a dilution series of purified human genomic DNAs performed by LGC and two independent laboratories. The graphs show; (a) percent of samples returning a usable profile (≥7 alleles detected) and (b) percent of samples returning a full ParaDNA profile (12 alleles). UCF tested two DNA samples in triplicate between 4000 pg and 250 pg; FIU tested one DNA sample in triplicate between 8000 pg and 62.5 pg and LGC tested six DNA samples with six replicates between 8000 pg and 31.25 pg total DNA per plate (* = not tested).

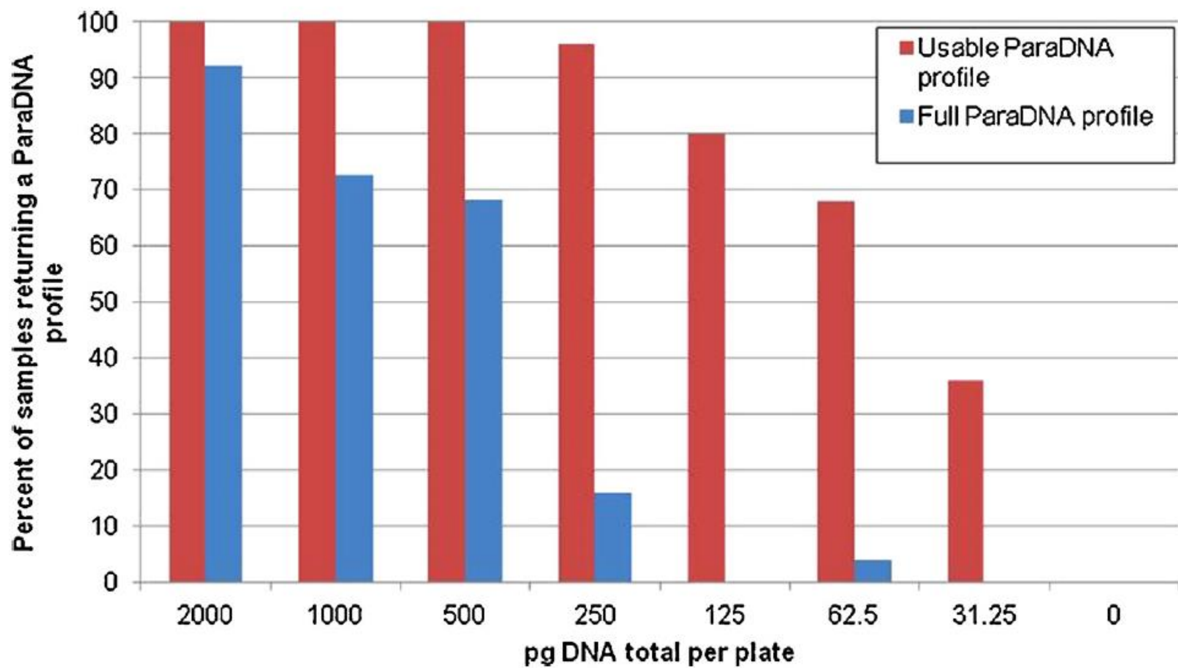

Fig. 2. Sensitivity plot of the ParaDNA Intelligence Test with a dilution series of human saliva samples performed by LGC. The DNA quantity in each of the saliva samples used was estimated by prior purification and quantification of DNA from an equivalent amount of saliva. Percent of saliva samples giving usable (≥7 alleles) and full (12 alleles) at each saliva dilution (DNA amount).

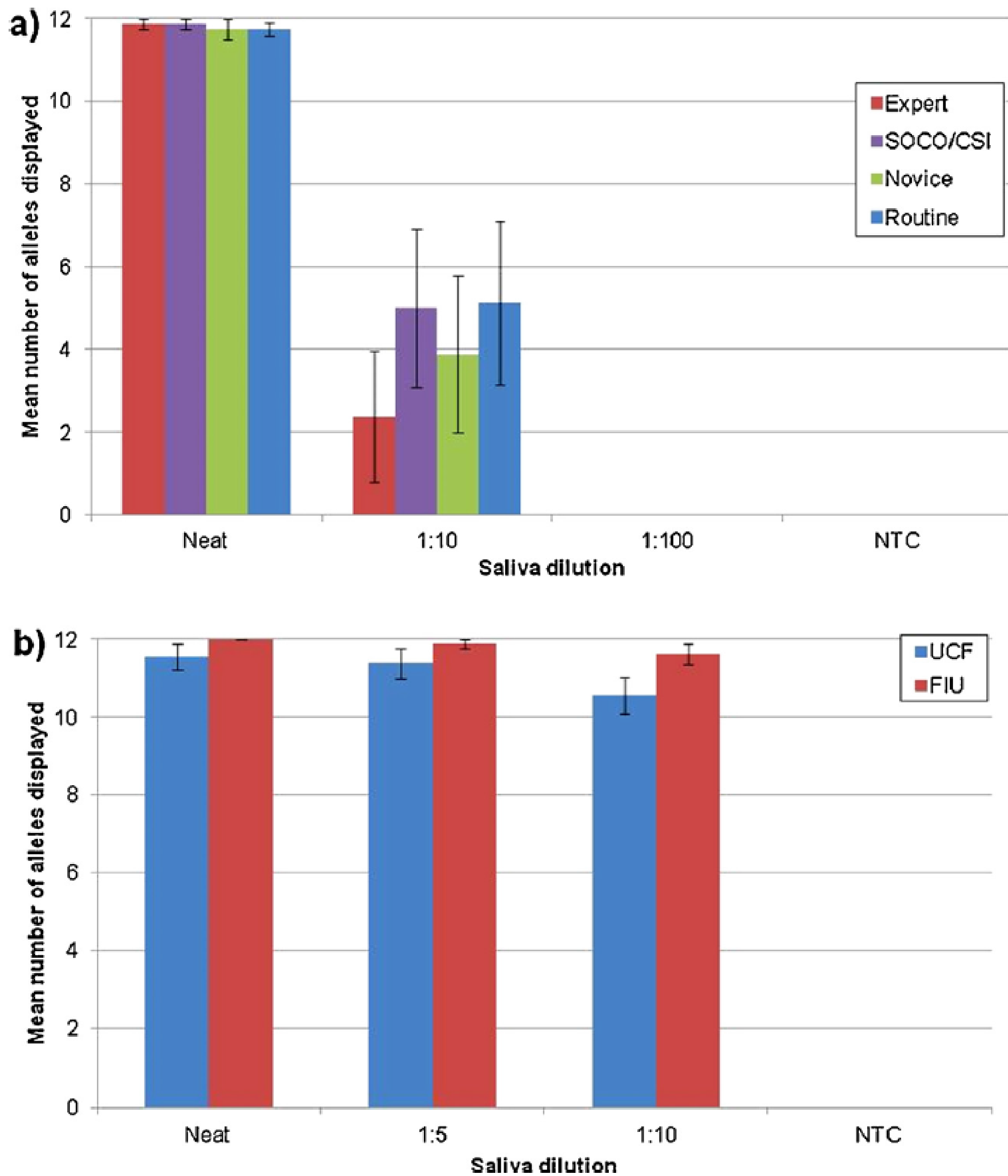

Fig. 3. Repeatability and reproducibility of allele calls between; (a) four different operators with different levels of experience and (b) two independent labs of recently trained operators. Expressed as mean number of displayed alleles (±SEM, n = 8).

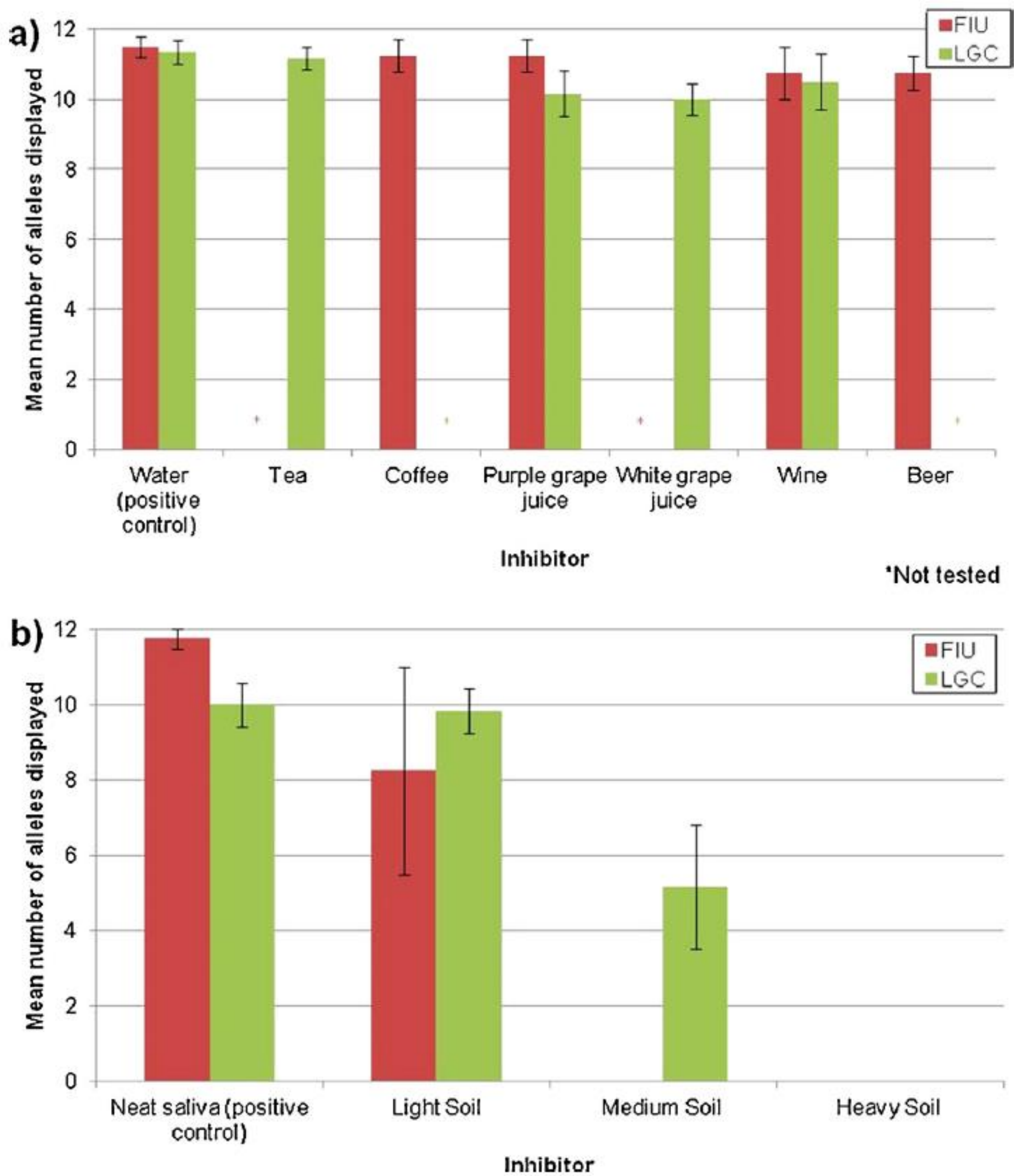

Fig. 4. Effect of "real-world" inhibitors of, (a) beverages containing polyphenolic compounds (FIU performed analyses on coffee, purple grape juice, wine and beer; LGC analysed the effect of tea, purple/white grapes juices and wine); (b) soil, on the ParaDNA Intelligence Test, expressed as mean number of alleles displayed (±SEM, n=6; *=not tested).

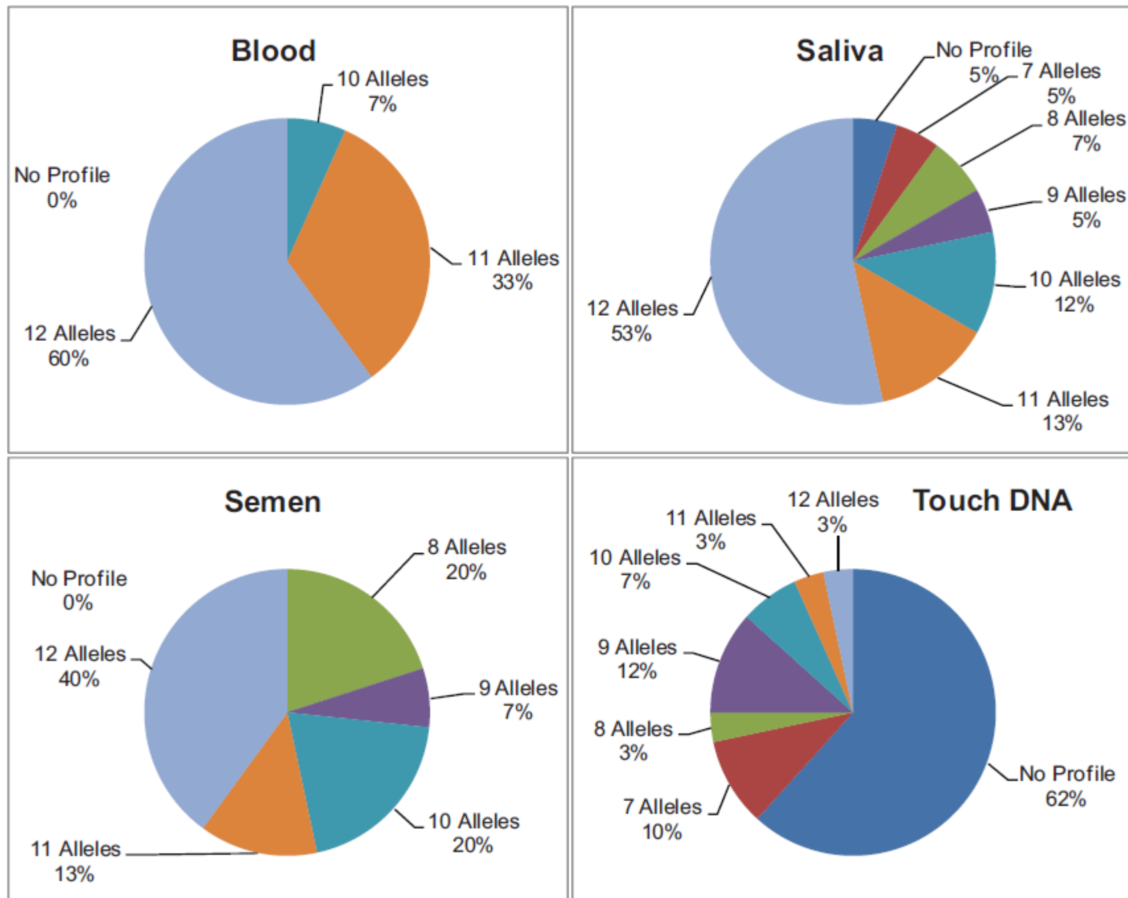

Fig. 5. Performance (percent of samples displaying number of allele calls) of the ParaDNA Intelligence Test with different categories of “case-type” samples.

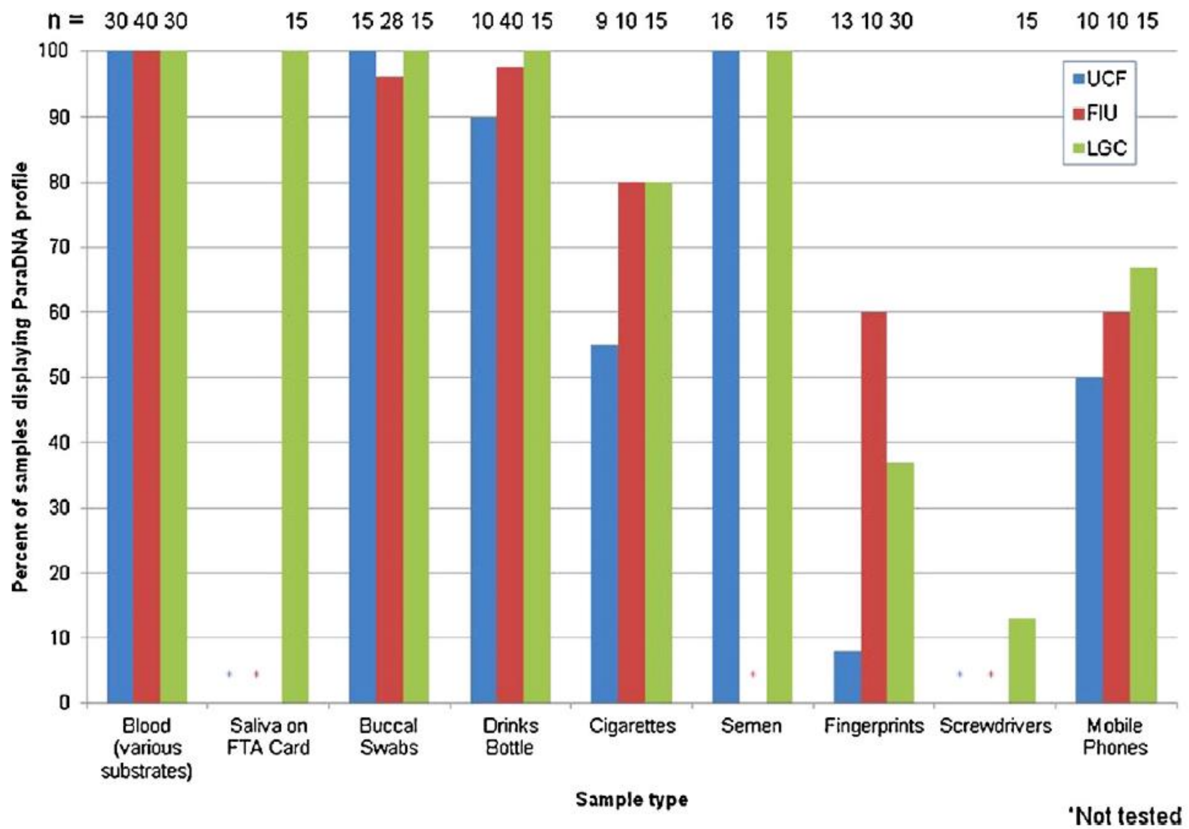

Fig. 6. Performance (percent of samples displaying a usable ParaDNA profile) of the ParaDNA Intelligence Test for different categories of "case-type" samples. UCF did not test screwdrivers or FTA card. FIU did not test semen, screwdrivers or FTA card (n=9–40 depending on sample type and testing site and is displayed above chart) (* = not tested).

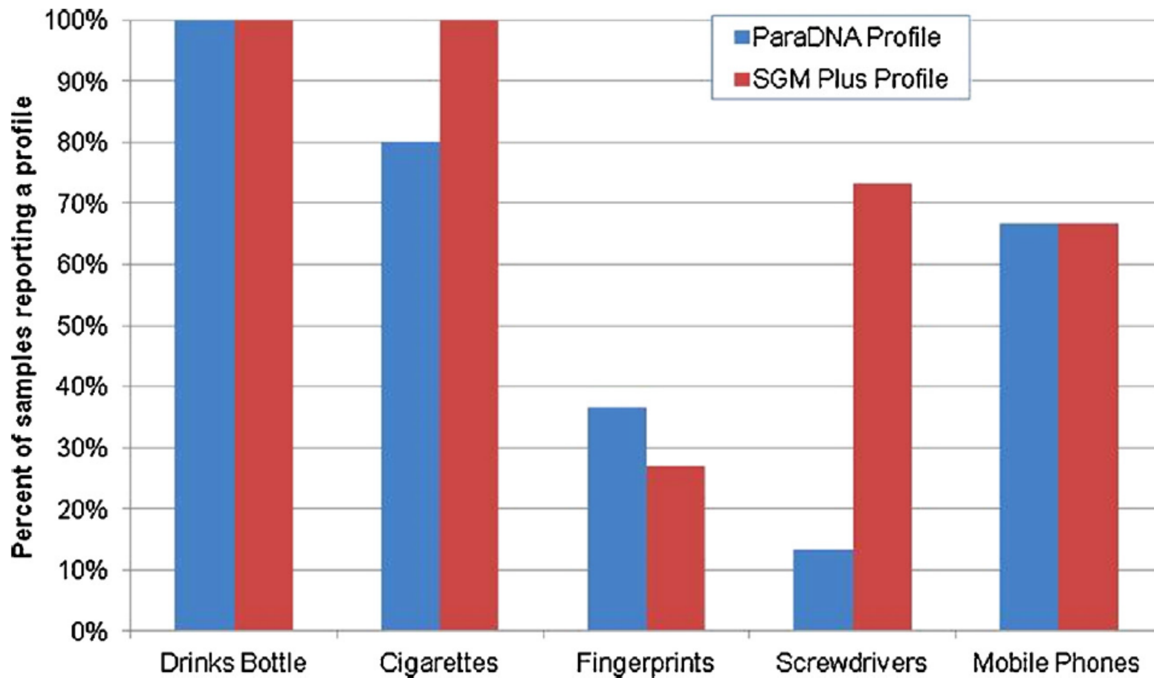

Fig. 7. Comparison of ParaDNA Intelligence Test performance on different sample types with SGM Plus. ParaDNA reported profiles where seven or more alleles were detected. SGM Plus profiles were deemed successful if seven or more alleles from the ParaDNA STRs were called. For both analysis methods, n = 15 for each sample type except for fingerprints where n = 30.

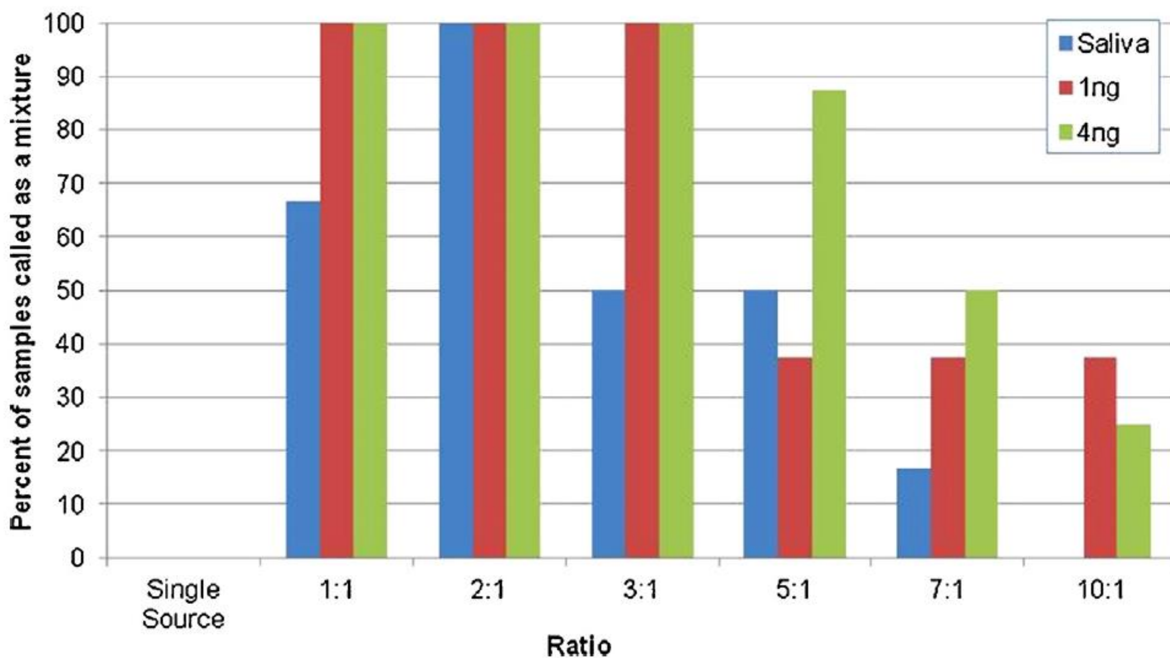

Fig. 8. Percent of samples called as mixture at different mixture ratios. n = 8 for DNA samples (n = 4 for 1:1); n = 6 for saliva samples (n = 3 for 1:1).