Latest Developments in Forensic DNA Profiling- the UK National DNA Database

Helen N. Dunbar, Rebecca L. Sparkes, Andrew J. Hopwood, Richard Pinchin, Stephanie K. Watson
Forensic Science Service, Priory House, Gooch Street North, Birmingham, UK. B5 6QQ

INTRODUCTION

The National DNA Database (NDNADB) of England and Wales utilises a heptaplex STR multiplex developed by the Forensic Science Service. The Second Generation Multiplex (SGM) co-amplifies Amelogenin, D21S11, D18S51, HUMTH01, HUMVWFA31/A, HUMFIBRA (FGA) and D8S1179 giving an estimated discriminating power of 1 in 50 million [1, 2].

The NDNADB was launched in 1995 with an estimated 135,000 Criminal Justice (CJ) samples to be submitted every year by the 41 UK police forces. CJ samples were defined by the Police and Criminal Evidence Act 1984 as samples from any person arrested for an indictable offence. To date, over 300,000 samples have been processed making it the largest of its kind in the world. Approximately 19,000 individual to crime scene and 4500 scene to scene matches have been identified. As a result of this success, sample submission rates have increased. The need for a high throughput STR system both in terms of efficiency and cost reduction has led to research into automated processes. We have developed, in conjunction with Process Analysis and Automation Ltd. and Hamilton (GB) Ltd., a microtitre plate-based robotic process for the manipulation of samples prior to electrophoresis.

To increase the discriminating power of the SGM, we intend to introduce a Third Generation Multiplex (TGM). The discriminating power of TGM is in the region of 1 in 20 million giving a SGM/TGM combined discriminating power of 1 in 10^{15}.

TGM

Loci for this multiplex were selected on the basis of several factors; (a) an observed discriminating power above 0.85, (b) heterozygosity above 0.7, (c) PCR product between 100 and 340 base pairs in length, (d) loci to be located on different chromosomes to avoid linkage, (e) low stutter formation and (f) low ratio of n to n+1 peaks. HUMTH01 was included in TGM as an internal control in common between SGM and TGM multiplexes. The final composition contains D3S1358, HUMTH01, D2S1338, D1S518, D10S516, D22S684 and D14S306 (Fig.1).

The reaction mix was optimised by evaluating research into individual component concentrations. Similarly, PCR cycling conditions were examined and a compromise reached on all parameters to allow adequate amplification of all loci.

Preliminary allelic ladders were created by the pooling of PCR products that contained a spread of alleles. Final allelic ladders were produced by purifying the required alleles and combining them [3]. The purified alleles were sequenced [4]. D14S306 and D1S518 are both simple GATA repeats, D22S684 is a complex repeat and D3S1358, D2S1338, and D10S516 are compound repeats. Allele distributions were generated for Caucasian, Afro-Caribbean and Asian populations.

AUTOMATION

Currently, CJ samples are uniquely identified by a barcode and manually extracted using a Chelex-100 based method [5], quantified by Fluoroskan pico green [6] and PCR'd with STR multimix [1,2] plus Amplitaq™ *Taq* DNA polymerase (Perkin Elmer, Warrington, UK). The amplification products are then electrophoresed on an acrylamide gel using ABD 377 automated sequencer technology and analyzed by GeneScan™ analysis software. Complete DNA profiles are submitted for inclusion onto the NDNADB and any partial profiles or sample failures are repeated. The majority of this process is labour intensive and repetitive, therefore automation is required.

The automated system had to satisfy the following criteria; (a) capable of high-throughput, (b) minimal operator involvement, (c) software capabilities for sample tracking, (d) cost-effective and (e) no cross-contamination. We have achieved this by the introduction of two robotic systems, the first for pre-PCR and the second for post-PCR product manipulation prior to gel running. Both of these systems are controlled by Overlord (Process Analysis and Automation), a supervisory software package which allows the execution of both DOS and Windows® based packages. Overlord also communicates with an in-house Laboratory Information Management System (LIMS) for sample tracking. High throughput is achieved by using 96-well microtitre plates permitting 72 CJ samples to be processed simultaneously.

The pre-PCR robotic system consists of the following components [Fig.2]: The plate hotel holds the microtitre

plates i.e. reagent plates, PCR plates, DNA plates; the plate stacks hold empty microtitre plates intended for use as dilution plates and quantification plates plus the bleach plate for probe washing; Fluoroskan Ascent (LabSystems, Finland) PicoGreen™ quantification equipment; a pipetting station containing a single probe and a multiprobe head of 8 cannulae; a barcode reader for sample and batch tracking; a water purification unit which supplies distilled water for washing probes; a waste bin; an IBM compatible computer for system control and a robotic arm for plate manipulation.

This system has the capability of performing several protocols dependent upon the type of run required. A standard run makes a standard quantitation curve using known concentrations of stabilised sonicated human placental DNA (Cambio Ltd., Cambs, UK) and then quantifies the extracted DNA samples. From this data, Overlord calculates the volume of extract required for PCR and instructs the robotics to create a PCR plate. In the event of a PCR failure, a whole batch rerun is necessary. In this case, Overlord instructs the operator to place the appropriate DNA plate in the specified position on the plate hotel from where a new PCR plate for that batch is made up. Individual samples from different batches can also be combined into a new PCR batch as the system has the capacity to select individual samples from up to 70 different DNA plates.

The post-PCR robotic system consists of the following components [Fig.3]: A plate hotel for PCR, denaturation, reagent and bleach plates; a pipetting station similar to that operating with the pre-PCR robotic system; a barcode reader for sample and batch tracking; a water purification unit for probe washing; an IBM compatible computer for system control and a robotic arm for plate manipulation.

This system prepares PCR products for loading onto an acrylamide gel. A standard run transfers DNA products from the PCR plate to a denaturation plate and adds the dextran blue/formamide/size standard mix. An operator then transfers the denaturation plate to a Perkin Elmer 9600 PCR block. The system can also re-run whole batches or mixed batches from up to 40 plates.

Currently, DNA extraction and gel loading are manual procedures. The automation of both depends directly on a method conducive to automation. Alternative extraction techniques to Chelex are being researched. The promise of 96 capillary electrophoresis technology e.g. Molecular Dynamics may also be another step towards a fully automated STR multiplex system.

STRess

Prior to the launch of STRess (STR Expert System Suite) [7, 8, 9] for routine use with CJ samples in 1997, a gel was genotyped by two independent operators (genotyper A and B) and a third operator compared the results before submitting a mediated profile for inclusion onto the NDNADB. This system was both time- and manpower-consuming, creating a “bottle neck” in sample processing. The aim of STRess was to release manpower for sample processing by replacing genotyper B. As such, the onus on STRess was to perform as well or better than a human operator. Interpretation guidelines, drawn from approximately 100,000 samples processed by the FSS and used by experienced operators, were incorporated into the programming.

The actual allelic ladders found on the gel are compared to a known ladder pattern and any shift from the ideal (acrylamide batch specific) at each peak is calculated. This shift value is then used to compensate for any missing peaks, resulting in the formation of a ‘virtual’ allelic ladder that is gel specific and is used to designate sample lanes on the gel. STRess adds customised comments according to a range of post designation rules for example, if a homozygote peak is below threshold level for reporting, it will be brought to the attention of the operator comparing outputs of STRess and genotyper A.

It has been estimated that the incorporation of STRess into routine analysis has achieved a time saving of around 40%.

SUMMARY

In summary then, the NDNADB has proved to be a success as shown by increasing submission and profile match rates. TGM is nearing full validation with its implementation for adventitious hits on the NDNADB imminent. Validation of the robotic systems is now underway and STRess has already made a substantial impact on time saving since it was introduced into the NDNADB processing system in the last quarter of 1997.

REFERENCES

1. Sparkes R., Kimpton C., Watson S., Oldroyd N., Clayton T., Barnett L., Arnold J., Thompson C., Hale R., Chapman J., Urquhart A., Gill P., (1996) The Validation of a 7-Locus Multiplex STR Test for Use in Forensic Casework (I) Mixtures, Ageing, Degradation and Species Studies. *Int. J. Legal Med.*, **109**:186-194

2. Sparkes R., Kimpton C., Gilbard S., Carne P., Andersen J., Oldroyd N., Thomas D., Urquhart A., Gill P., (1996) The Validation of a 7-Locus Multiplex STR Test for Use in Forensic Casework (II) Artefacts, Casework Studies and Success Rates. *Int. J. Legal Med.*, **109**:195-204

3. Watson S., Kelsey Z., Web R., Evans J., Gill P., The Development of a Third Generation STR Multiplex System (TGM). Unpublished.

4. Urquhart A., Kimpton C.P., Downes T J., Gill P., (1994) Variation in Short Tandem Repeat Sequences: A Survey of Twelve Micro-satellite Loci for Use as Forensic Identification Markers. *Int. J. Legal Med.*, **107**:13-20.

5. Walsh P.S., Metzger D.A., Higuchi R., (1991) Chelex® 100 as a Medium for Simple Extraction of DNA for PCR-Based Typing from Forensic Material. *Biotechniques*, **Vol.10** No.4.

6. Hopwood A., Oldroyd N., Fellows S., Ward R., Owen S.-A., Sullivan K., (1997) Rapid Quantification of DNA Samples Extracted from Buccal Scrapes Prior to DNA Profiling. *Biotechniques*, **80**:18-20.

7. Werrett D., Pinchin R., Hale R., Problem Solving Data Acquisition and Analysis. Unpublished.

8. Gill P., Urquhart A., Millican E., Oldroyd N., Watson S., Sparkes R., Kimpton C., (1996) A New Method of STR Interpretation Using Inferential Logic - Development of a Criminal Intelligence Database. *Int. J. Legal Med.*, **109**: 14-22.

9. Gill P., Sparkes R., Kimpton C., (1997) Development of Guidelines to Designate Alleles Using an STR Multiplex System. *Forensic Sci. Int.*, **89**:185-197

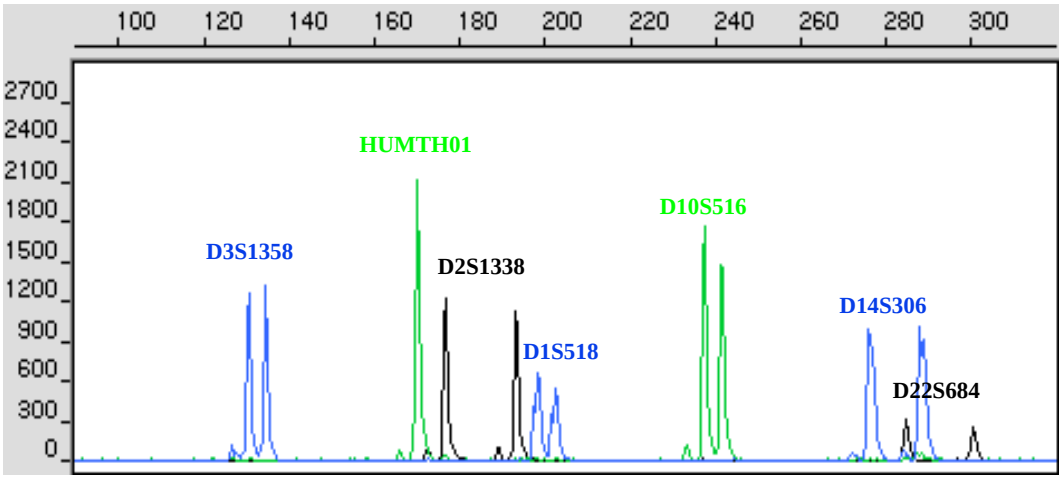

Figure 1: A Typical TGM STR Profile

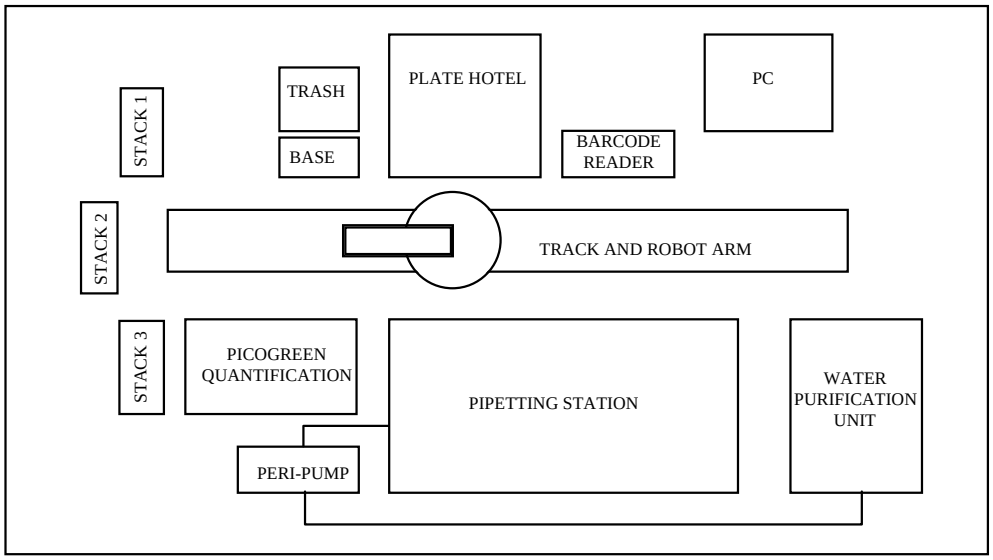

Figure 2: Schematic of the Pre-PCR Robotic System

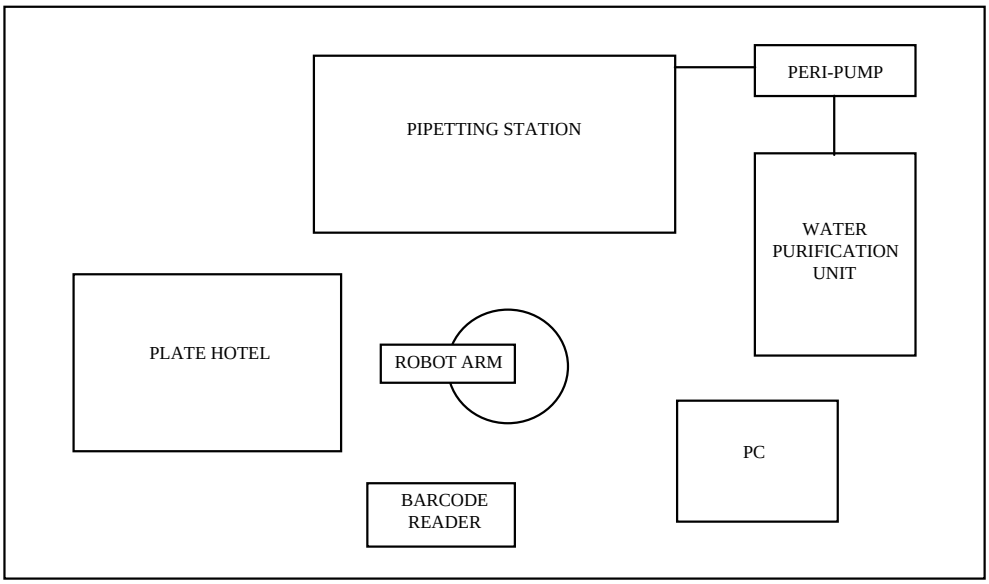

Figure 3: Schematic of the Post-PCR Robotic System