Bar coding objects with DNA†

Jonathan P. L. Cox

Department of Chemistry, University of Bath, Bath, UK BA2 7AY. E-mail: E-mail: chsjplc@bath.ac.uk; Tel: +44 1225 826548; Fax: +44 1225 826231

Received 27th February 2001, Accepted 20th March 2001
First published as an Advance Article on the web 4th April 2001

A bioanalytical scheme is described for bar coding objects with DNA. Several DNA templates of different lengths are mixed together and the mixture used to label the object of interest. To disclose the bar code, the DNA mixture is eluted from the object, amplified by the polymerase chain reaction with a universal pair of primers, and then the amplified templates resolved by agarose gel electrophoresis. In its present state, the bar coding scheme allows for over a million different labels. The reliability of the scheme has been demonstrated in a typical situation by labelling a letter with a template mixture (~1.5 fmol of DNA), sending the letter through the post and successfully reading the bar code on receipt of the letter.

Introduction

Using DNA to mark objects, for example to protect a document against fraud or for tracing purposes, is an extremely attractive and indeed well-known concept (recall the online demonstration of projectile tagging in the film Judge Dredd). Recently it has been put into practice.1,2 The advantages of DNA as a label are that information can be stored in the nucleotide sequence and only tiny amounts are required (theoretically, just a single molecule) effectively making the label invisible. However, a DNA fingerprint3,4 is another valuable source of information which can be extracted from DNA without recourse to sequencing. This communication describes a scheme for generating artificial DNA fingerprints to chemically bar code objects.

The scheme is straightforward (Fig. 1). A number of DNA templates of different lengths are selected from a master set and combined. The template mixture is then used to label the object of interest. The “object” could be either a solid or liquid. In Fig. 1 it is the page of a document. To read the label, the templates in the mixture are simultaneously amplified in a single polymerase chain reaction (PCR)5 with one pair of primers. The amplified templates are then analysed by agarose gel electrophoresis.6 The result is a set of bands constituting the object’s bar code. The scheme is similar to genomic DNA fingerprinting techniques (e.g. ref. 6). The key difference being that the composition of the artificial fingerprint is strictly controlled by the labeller. This difference has a practical advantage: the DNA mixtures which give rise to the artificial fingerprints are easy to make; preparing samples of genomic DNA which produce a similar number of sharply defined fingerprints would be much harder.

In this communication it is shown that at least a million different DNA bar codes are attainable by agarose gel electrophoresis. The mixtures used to generate the bar codes have been successfully employed in labelling objects (letters) and are tolerant to handling. Furthermore, the bar coding scheme is quick to implement and requires only routine bioanalytical equipment. It should be particularly useful in forensic science and in matters of security.

Experimental

Templates

The twenty DNA templates were amplified from a vector derived from pYX131 (R & D Systems). Vector solutions were prepared from E. coli using a S.N.A.P. kit (Invitrogen). After elution from the resin provided with this kit, care was taken to handle all vector solutions with Gilson Pipetmen fitted with Aerosol-Resistant Tips (Molecular BioProducts, Merck) to avoid contamination.

Details of the DNA templates and the vector housing them are available as Electronic Supplementary Information (ESI).† The ESI also contains details of primer sequences, template concentrations, optimisation of PCR conditions and further comments on the PCRs.

PCR analysis

Polymerase chain reactions were performed on a PTC-100 Programmable Thermal Controller (MJ Research) in a 50 μl volume containing the Expand High Fidelity PCR System (Roche) enzyme mix (0.75 μl) and buffer (with MgCl2), 200 μM of each dNTP (Ultrapure, Amersham Pharmacia Biotech) and 25 pmol of each primer. The two primers were obtained from Perkin-Elmer Applied Biosystems and used as supplied. The amount of each template varied from 0.01 to 1 fmol. The PCR programme was 94 °C, 2 min; (94 °C, 15 s; 55 °C, 30 s; 72 °C, 1 min) × 15; (94 °C, 15 s; 55 °C, 30 s; 72 °C, 1 min + 5s per cycle) × 17; 72 °C, 7 min.

DNA bar coding a letter

To label a letter, an aqueous 10-template mixture was diluted by half with ethanol (AnalR, BDH) and 0.5 μl of this new solution spotted onto a designated region of the letter. The total amount

Fig. 1 Bar coding an object with DNA. Here the object is the page of a document. See “Introduction” for full explanation. Clear rectangles: double-stranded DNA templates.

† Electronic Supplementary Information available. See http://www.rsc.org/suppdata/an/b1/b101850g/
‡ Although the resolution achievable by polyacrylamide gel electrophoresis (PAGE) would allow a far greater number of fingerprints, PAGE is expensive and time-consuming to do, and would therefore be at odds with the intended simplicity of the scheme.

DOI: 10.1039/b101850g

This journal is © The Royal Society of Chemistry 2001

Analyst, 2001, 126, 545–547 545
of DNA spotted was about 1.5 fmol (0.1–0.2 fmol of each template). The ethanol acted as a wetting agent: the paper (white A4 CopyPrintPlus) absorbed aqueous solutions of the template mixture very poorly, but with the ethanol the solution was absorbed rapidly (seconds). On absorption, the solution spread over an area of about 2 mm\(^2\). The paper was dry in approximately 5 min. Typically letters spent three days in transit. On return of the letter, the area which had been spotted with the DNA solution was excised with a clean razor blade and immersed in 30 \(\mu\)l of water for 15 min. Then, after gently tamping the paper with a pipette tip, the water was transferred to a PCR tube, the reagents detailed above added and the PCR performed.

### Daughter fingerprint

The products from a 6-template reaction were purified with a Wizard PCR Preps kit (Promega), eluting with 30 \(\mu\)l. Ten \(\mu\)l of the eluate were digested with 10 U of Ear I (New England Biolabs) for 4 h.

### Agarose gel electrophoresis

Samples from PCRs and the Ear I-digestion were run on 2.5% agarose (Bio-Rad Laboratories) gels at 150 V for about 1 h 30 min using an Anachem Midi electrophoresis unit. The fingerprints may also be read using mini-gel apparatus.

In Figs. 2 and 3, 15 \(\mu\)l from each PCR were loaded on to the gel. Lane 1 of Fig. 4 contains 3 \(\mu\)l from the Wizard PCR Prep of a 6-template PCR. Lane 2 of Fig. 4 contains 10 \(\mu\)l of this same Wizard PCR Prep, but treated with Ear I.

![Figure 2](image1.png)

**Fig. 2** Twenty- and ten-band fingerprints. Two separate examples of the simultaneous amplification of 20 templates (lanes labelled “20”) and five examples of 10-template amplifications (“10s”, lanes A–E) are shown. M: 50 base pair DNA Ladder (Life Technologies). Figures alongside the gel photograph are in base pairs. Although one template was not selected for the 10-template PCRs (the one giving rise to the third lowest band in the 20-template PCRs), it was successfully amplified in other multiple template PCRs (e.g. lowest band in lane 1, Fig. 4, in addition of course to the 20-template PCRs).

![Figure 3](image2.png)

**Fig. 3** Fingerprint obtained from a letter labelled with a 10-template mixture (the same mixture that had given rise to the fingerprint shown in lane C of Fig. 2). The letter had been sent through the post. Lane 1: PCR on labelled section of letter. Lane 2: PCR on unlabelled section of letter. Figures alongside the gel photograph are in base pairs.

### Results and discussion

Twenty templates were chosen to carry out the proposed DNA bar coding scheme. These make possible a large number of fingerprints (2\(^{20}\), just over a million). The templates vary in size from about 80 to 800 base pairs. Each template is contained in a vector and consequently amplified from this vector.

The first important goal was to show that all 20 templates can be simultaneously amplified in a single PCR to give detectable, resolved bands. Inspection of fingerprints obtained from genomic DNA\(^7,8\) suggested that this would be possible although it was by no means a foregone conclusion. Fig. 2 (lanes labelled “20”) confirms that all 20 templates can be simultaneously amplified in a single PCR to give detectable, resolved bands. In building up to the 20-template reaction, 6-, 8-, 9-, 11-, 12- and 18-template reactions were also performed successfully (data not shown).

Next it was necessary to show that a range of fingerprints could be obtained from subsets of randomly chosen templates. Several 10-template reactions were attempted. For each reaction, the ten expected bands were observed. The results from five such reactions are depicted in Fig. 2 (lanes A–E).

To show that the DNA templates are able to act as a label, a small amount (0.5 \(\mu\)l, containing about 1.5 fmol total DNA) of the template mixture which produced the fingerprint in lane C, Fig. 2, was spotted onto a printed letter (to cover an area of about 2 \(\text{mm}^2\)). After allowing the paper to dry for 10 min (on drying, the mixture was undetectable in visible or UV light), the area which had been spotted with DNA was excised, soaked in water for 15 min, the aqueous solution removed and subjected to a PCR. The templates were amplified successfully (data not shown).

In their imaginative work on DNA microdots, Clelland et al. tested the microdots by attaching them to a printed letter and sending them through the US mail.\(^2\) A very similar strategy was adopted here to demonstrate the durability of the templates as a label. A letter marked in the same manner as described in the previous paragraph was posted to a colleague who read the letter and then posted it back. A PCR was then performed on the labelled section of the letter, and also on an unlabelled section of the letter, as outlined in the previous paragraph. The results are shown in Fig. 3. Pleasingly, the templates were amplified successfully from the labelled section but not from the unlabelled one. This experiment was repeated several times with the same positive outcome.

In addition to surviving the rigours of the postal system, the template mixtures are stable on paper for at least six months. While the stability of the template mixtures is still under scrutiny, given the almost geological longevity of cell-free, hydrated DNA,\(^9\) one might expect the template mixtures to remain intact for some considerable time (years).

One of the principal attractions of the DNA bar coding scheme is that it requires only very basic bioanalytical apparatus (agarose gel electrophoresis equipment and a thermal cycler) and should therefore be cheap and convenient to implement. Results can also be obtained fairly rapidly: making up the label (the most tedious part of the scheme, which in any case would
lend itself readily to automation) takes about half an hour; acquiring the fingerprint from a labelled letter takes just under four hours.

There is also considerable opportunity for developing the scheme further. For instance, although 20 templates potentially give rise to over a million fingerprints, it is clear from Fig. 2 (lanes “20”) that there is plenty of room for additional bands, above (in particular), below and inbetween existing ones. Thirty-five to forty bands is probably close to the limit of what can be resolved on a single agarose gel (optically-read bar codes incorporate up to about 50 bars$^{10}$). Therefore assuming that $35–40$ templates can be amplified in a single PCR and subsequently detected on an agarose gel, the number of possible fingerprints is vast ($10^{10}–10^{12}$).

With only a little extra effort, the complexity of the fingerprints may be increased further still by restriction digestion of the amplified mixture, to give a daughter fingerprint. For example, four of the six templates which have been simultaneously amplified in lane 1 of Fig. 4 have a single recognition sequence for the restriction enzyme $Ear$ I. Thus, on digestion with $Ear$ I, the susceptible templates are cleaved to give the daughter fingerprint in lane 2. To make effective use of this derivative scheme, the master set of templates could be expanded to include a small number of new templates which are the same size as, but different in sequence from (and therefore having varying degrees of sensitivity to a nominated restriction enzyme) the original templates. By switching between templates of the same size it would then be a simple matter to create apparently identical fingerprints distinguishable by subsequent enzymatic digestion, in the same way that genomic restriction fragment length polymorphisms are used as genetic markers.$^{11}$

Finally, throughout this communication emphasis has been placed on resolving the different DNA fragments comprising the fingerprint. There is, however, really no reason why resolution should be so important — two or more unresolved bands could act like the thick elements of conventional bar codes. The two bands (which in fact arise from four separate fragments of 151 to 171 base pairs) just above the 150 base pair mark in lane 2 of Fig. 4 have almost coalesced to give a feature of this kind.

Acknowledgements

Reading ref. 2 and The Code Book by Simon Singh (Fourth Estate, London, 1999) prompted this work. I am grateful to Pooja Kumar and Derek Woolfson for participating in the postal experiments, Bryn Davies for technical assistance and Cameron Neylon and Simon Palmer for comments on the manuscript. Financial support was kindly provided by the University of Bath and the Wellcome Trust.

References