This is an Accepted Manuscript, which has been through the Royal Society of Chemistry peer review process and has been accepted for publication.

Accepted Manuscripts are published online shortly after acceptance, before technical editing, formatting and proof reading. Using this free service, authors can make their results available to the community, in citable form, before we publish the edited article. We will replace this Accepted Manuscript with the edited and formatted Advance Article as soon as it is available.

You can find more information about Accepted Manuscripts in the Information for Authors.

Please note that technical editing may introduce minor changes to the text and/or graphics, which may alter content. The journal’s standard Terms & Conditions and the Ethical guidelines still apply. In no event shall the Royal Society of Chemistry be held responsible for any errors or omissions in this Accepted Manuscript or any consequences arising from the use of any information it contains.
We report the labelling of dideoxy nucleotides (ddNTPs) for use in electrochemical array based primer extension for the detection of single nucleotide polymorphisms (SNPs). The results confirm extension of the immobilised primers for each of the four ddNTPs, representing a significant advance in achieving a cost-effective platform for screening for disease-specific SNPs.

The completion of the human genome project (HGP) has paved the way for mapping diversity of the genome, thus helping to understand the genetic causes of inherited diseases and susceptibility to drugs or environmental toxins. Over the past decade the development of new strategies for genotyping has attracted increasing interest, driven by the need for cost effective and efficient methods to take advantage of the knowledge acquired during the HGP in order to assess a broad range of biological phenomena (e.g., genetic variation, RNA expression, protein-DNA interactions and chromosome conformation). Finally, the advance of technology across diverse fields, including nucleotide biochemistry, polymerase engineering and computation, has facilitated the realisation of alternative strategies. In the genome sequence there are individual variations that include single nucleotide polymorphisms (SNPs), insertions and deletions (indels), microsatellites (MSs), and differences in the methylation status of important regions (e.g. CpG islands). However, the majority of the variations are attributable to SNPs. SNPs are single base pair mutations in a genome that occur in at least 1% of the total population. SNPs are attributable for 90% of the genetic variations and the rest is attributable to insertions or deletions of one or more bases, repeat length polymorphisms and rearrangements. The decoding of the human genome has revealed the presence of around 10 million SNPs (roughly 1 every 300-1000 bases). SNP genotyping is of fundamental importance and vast international effort is currently being made using next-generation sequencing to identify the location of SNPs in specific populations (e.g. to identify disease-associated SNPs), which is predicted to result in patient stratification, and a more personalised approach to medicine.

Arrayed primer extension (APEX) is a high throughput genotyping method that exploits dideoxy nucleotides for scanning known mutations over large regions of a DNA sequence. Typically, the APEX procedure involves locus specific PCR amplification, followed by fragmentation using uracil-N-glycosylase. The fragmented PCR products are denatured and hybridised to complementary capture probes that are surface-tethered on a glass array. Once hybridised they serve as primers for template-dependent DNA polymerase extension reactions using four fluorescently labelled dideoxynucleotides. Imaging is followed by data analysis to convert the fluorescence information into sequence data. The developed fluorescent APEX platform can simultaneously interrogate many SNPs in a single multiplexed assay. Motivated by the high accuracy and specificity of the fluorescent APEX, we have demonstrated electrochemical solid phase single base extension for SNP detection in the cardiomyopathy associated MYH7 gene. Electrochemical primer extension has several advantages including cost-effective, simple-to-easy and portable instrumentation that does not suffer from background light, as well as relatively inexpensive electrode arrays and well-established surface chemistries for automated probe immobilisation via spotting. The approach detailed here considerably simplifies previous reports of electrochemical detection of single base extension, where SBE and detection are carried out separately, or using dUTPs which are problematic for homopolymers. Metallic nanoparticles (NPs) have also been used, where probes hybridise DNA containing the SNP to be interrogated, and following hybridisation mismatched bases are hybridised to metallic NPs specific for each base, which is then detected using stripping voltammetry.

In the work reported here, ddNTPs used were covalently linked with four different redox active compounds with distinguishable electrochemical signals (anthraquinone, phenothiazine, methylene blue and ferrocene), at positions favourable for enzymatic incorporation.
The redox active dideoxy nucleotide triphosphates used for electrochemical primer extension assays. (AQ: anthraquinone (-0.40 V), MB: methylene blue (-0.20 V), Fc: ferrocene (0.50 V) and PTZ: phenothiazine (0.60 V)) All potentials vs Ag/AgCl.

The four different redox labels were chosen based on their distinctive redox potentials, allowing clear discrimination between the labels (Fig. 1). Carboxyl functionalised ferrocene (Fc), anthraquinone (AQ) and methylene blue (MB) were purchased from commercial sources, while phenothiazine was functionalised with an alkyl arm bearing a carboxyl terminus following the stepwise reaction protocol described in Figure S11. These carboxyl derivatives were used for functionalising propargyl amine bearing ddNTPs via amide bond formation, to achieve compounds for electrochemical solid phase primer extension (éPEX) reaction. The labelled ddNTPs were characterised using NMR and ATR FT-IR (Supporting Information). As can be seen in Fig. 2, éPEX consists of hybridisation of the template to a surface-tethered probe, followed by enzymatic incorporation of a label modified ddNTP, which terminates elongation, thus only extending the primer by one single base at the known SNP location. Subsequent to primer extension, electrochemical interrogation of the incorporated labels was performed by monitoring the redox reaction using differential pulse voltammetry (DPV) in potential windows corresponding to each of the labels.

Fig. 1 The redox active dideoxy nucleotide triphosphates used for electrochemical primer extension assays. (AQ: anthraquinone (-0.40 V), MB: methylene blue (-0.20 V), Fc: ferrocene (0.50 V) and PTZ: phenothiazine (0.60 V)) All potentials vs Ag/AgCl.

Fig. 2 Top: Surface-tethered DNA probes hybridise to target sequences one base prior to the SNP-site to be interrogated, followed by incorporation of redox-labelled ddNTP. Bottom: Dependent on the specific base present at the SNP site, different ddNTPs will be incorporated, and the incorporated ddNTP can be identified via the DPV signal of the redox label. [Note: The schematic above shows the possibility for each potential base present at the SNP site, but only one of the four labelled ddNTPs will be incorporated at each SNP site.]
Modified polymerases have been shown to be highly tolerant to nucleotide modifications with various groups at position 5 of pyrimidine bases (C and U) and position 7 of purine bases (A and G). Various reports have detailed the enzymatic incorporation of dNTPs labelled with ferrocene, anthraquinone, phenothiazine and gel electrophoresis has been used to confirm the incorporation of the modified dNTPs. In the work reported here, we exploit surface tethered probes and primer extension reactions, the incorporation of the modified ddNTP can be demonstrated using cyclic voltammetry. The main signal in each case corresponds to the specific incorporation of (a) AQ-ddATP, (b) MB-ddUTP, (c) Fc-ddGTP and (d) PTZ-ddCTP, thus confirming the fidelity of the polymerase incorporating the correct modified ddNTP. Differential pulse voltammetry (DPV) was used to study the accuracy and specificity of their incorporation. As shown in Fig. 4, when the éPEX reaction was performed in the presence of the four labelled ddNTPs, the DPV signal corresponding specifically to the incorporated ddNTP was significantly higher than the signals from the non-specific labelled ddNTPs, thus confirming the specificity of their incorporation. As shown in Fig. 2, the electrostatic interactions between the redox label and the incorporated ddNTPs are slight higher non-specific signals as compared to ferrocene. Methylene blue is known to interact with free guanines DNA. However, the signal attributed to the incorporated ddNTP base is significantly higher than the non-specific background.

**Fig. 3** Left: Cyclic voltammograms (in 10 mM Tris buffer + 0.5 M NaCl, pH 7) for surface tethered DNA duplexes incorporated with the four different redox-labelled ddNTP as follows: a) AQ-ddATP, b) MB-ddUTP, c) Fc-ddGTP and d) PTZ-ddCTP. Right: The corresponding plots of the cathodic peak current, (lp) vs different scan rate, (υ).

The results obtained thus confirm the surface confinement of the incorporated labelled ddNTPs. Following the representation of Fig. 2, exploiting the electroactive properties of the modified dNTPs, differential pulse voltammetry (DPV) was used to study the accuracy and specificity of their incorporation. As shown in Fig. 4, when the éPEX reaction was performed in the presence of the four labelled ddNTPs, the DPV signal corresponding specifically to the incorporated ddNTP was significantly higher than the signals from the non-specific labelled ddNTPs, thus confirming the specificity of the polymerase incorporating the correct modified ddNTP. Due to electrostatic / groove interactions with DNA, anthraquinone, methylene blue and phenothiazine show slightly higher non-specific signals as compared to ferrocene. Methylene blue is known to interact with free guanines DNA. However, the signal attributed to the incorporated ddNTP base is significantly higher than the non-specific background.

**Fig. 4** DPV signals recorded after four éPEX reactions performed each in the presence of the four labelled ddNTPs. The main signal in each case corresponds to the specific incorporation of (a) AQ-ddATP, (b) MB-ddUTP, (c) Fc-ddGTP and (d) PTZ-ddCTP.

We have demonstrated the design, synthesis and characterisation of four electroactive labelled dNTPs with distinctive redox potentials. The synthesised labelled dNTPs were subsequently used for electrochemical solid-phase primer extension assays. These éPEX assays can be adopted for the detection of multiple mutations/SNPs using arrays of electrodes, offering an excellent platform for the detection of disease-specific SNPs, addressing the future medical paradigm of patient stratification for personalised medicine. Future work will focus on the use of alternate redox labels to reduce the background signal and expanding to simultaneous interrogation of multiple SNP sites.

**Acknowledgement:**
Cardioscreening, Plan Nacional Subdirección General Proyectos de Investigación, BIO2011-25369. This project has also received funding from the European Union’s Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 659211.
References