Quantitative analysis of gene expression changes in response to genotoxic compounds.

Ceri A. Morris\textsuperscript{a,b,*}, Gamal A. El-Hiti\textsuperscript{c}, Ian Weeks\textsuperscript{b}, Stuart Woodhead\textsuperscript{d}, Keith Smith\textsuperscript{e}, Peter Kille\textsuperscript{a}

\textsuperscript{a} School of Biosciences, Cardiff University, Cardiff CF10 3AT, UK
\textsuperscript{b} School of Medicine, Cardiff University, Tenovus Building, Heath Park, Cardiff CF14 4XN, UK
\textsuperscript{c} Cornea Research Chair, Department of Optometry, College of Applied Medical Sciences, King Saud University, P.O. BOX 10219, Riyadh 11433, Saudi Arabia
\textsuperscript{d} Invitron, Wyastone Leys, Monmouth NP25 3SR, UK
\textsuperscript{e} School of Chemistry, Cardiff University, Main Building, Park Place, Cardiff CF10 3AT, UK

* Corresponding author at: School of Medicine, Cardiff University, Tenovus, Heath Park, Cardiff CF14 4XN, UK.

E-mail address: morrisc10@cardiff.ac.uk; Tel. +44 (0)2920 874507 (C. A. Morris).
ABSTRACT

Techniques that quantify molecular endpoints sufficiently sensitive to identify and classify potentially toxic compounds have wide potential for high-throughput in vitro screening. Expression of three genes, RAD51C, TP53 and cystatin A (CSTA), in HEPG2 cells was measured by Q-PCR amplification. In parallel, we developed alternative assays for the same 3 gene signature based on an acridinium-ester chemiluminescent reporter molecule.

HEPG2 cells were challenged with eighteen different compounds (n=18) chosen to represent compounds that are genotoxic (n=8), non-genotoxic non-carcinogenic (n=2) or have a less well defined mechanism of action with respect to genotoxicity (n=8). At least one of the three genes displayed dysregulated expression in the majority of compounds tested by Q-PCR and ten compounds changed the CSTA expression significantly.

Acridinium-ester labelled probes for the three genes were synthesised and tested. Analytical sensitivity was characterised and suggested a limit of detection generally better than 0.1 fmol but often 10-50 attomol. A linear amplification step was optimised and this quantitative method detected statistically significant increases in RAD51C and CSTA expression in agreement with the Q-PCR results, demonstrating the potential of this technology. The broad agreement of the amplified chemiluminescent method and Q-PCR in measuring gene expression suggests wider potential application for this chemiluminescent technology.

Keywords: acridinium-ester, chemiluminescence, genotoxic, cystatin A (CSTA), RAD51C, tumour suppressor p53 (TP53)
1. Introduction

Assays based on the use of acridinium-ester (AE) as a chemiluminescent reporter molecule are well described in the literature (Arnold et al., 1989; Nelson et al., 1996 a,b) and are applied to the measurement of highly abundant ribosomal RNA of infectious agents. However, this technology has wider application as it is capable of measuring less abundant messenger RNA transcripts in a sensitive and quantitative fashion (Thomas et al., 2003a,b) with exquisite specificity (Morris et al., 2014). Here we aimed to extend the technology further by developing more sensitive chemiluminescent methods capable of measuring gene expression changes for toxicological application in conjunction with in vitro models.

To this end, the human hepatoma HEPG2 cell line was chosen as the test system. Although no cell model can perfectly mimic the in vivo situation, HEPG2 cells are often used in in vitro studies of toxicity as they express many of the phase I and II detoxification enzymes and a functioning gene (TP53) associated with the TP53 protein (Westerink and Schoonen, 2007a,b). Three genes were identified as potential targets of interest: RAD51C, TP53 and cystatin A (CSTA) primarily based on a publication by Westerink et al., (2010) which identified these genes from microarray analysis and demonstrated their potential with genotoxic compounds in HEPG2 cells in the form of reporter gene assays. The RAD family including RAD51C was also found to have altered gene expression in response to metal carcinogenesis (Kawata et al., 2009). Using genotoxic compounds identified from the publication by Kirkland et al. (2008), the gene expression in response to eighteen compounds with different mechanisms of action was elucidated using quantitative polymerase chain reaction (Q-PCR) for comparative purposes.

There is clear rationale for the involvement of RAD51C and TP53 protein in the downstream events of a chemical exposure as both are involved in DNA damage and repair pathways but the involvement of CSTA is less apparent. Westerink et al. (2010) noted that the induction of CSTA is quite specific to genotoxicity although the LXR agonist Tularik 0191317 (Westerink et al., 2010) and bile acids (Jones et al., 1998) are also known to induce this anti-apoptotic protein.
RAD51C protein is well documented to be involved in homologous recombination after DNA double strand breaks which are either a primary or secondary effect of exposure to many carcinogenic compounds. Initially, RAD51 accumulates at the sites of DNA damage in order to direct the other members of the DNA repair pathway and is believed to promote DNA strand exchange. RAD51C has been detected in a number of protein complexes with other members of the family including a complex with RAD51B, RAD51D and XRCC2 and a second with XRCC3 (Liu et al., 2004). There is evidence to suggest that RAD51C functions to resolve Holliday junctions that form to replicate DNA from the sister chromatid (Liu et al., 2004; Sharan et al., 2007). Defects in this gene have been found in families with a history of both breast cancer and ovarian cancer (Meindl et al., 2010). RAD51C was also identified in a study which looked at the expression profiles of HEPG2 cells after exposure to cadmium, nickel and arsenic (Kawata et al., 2009).

The ‘guardian of the genome’, TP53, is a tumour suppressor gene that encodes a pivotal protein in the response to DNA damage, mitotic spindle disruption and activation of oncogenes. Under non-stress conditions, the TP53 protein is a short-lived transcription factor. When cells are exposed to various forms of stress, it is activated by phosphorylation by checkpoint kinase Chk2 and then in turn binds to a TP53-responsive element in the promoter of target genes activating downstream target genes. Targets of the activated TP53 protein are found in cell-cycle arrest, DNA repair, senescence and apoptosis pathways (Powell et al., 2014). Due to its critical function in maintaining the genomic stability of the cell, levels of TP53 protein are tightly regulated and the main mechanism of TP53 activation, in response to cellular stresses such as DNA damage, is through post-transcriptional activation (Lakin and Jackson, 1999). There is a transient increase in TP53 mRNA levels which occurs as the cell transitions from G0/G1 phase of the cell cycle and into early S-phase in preparation for TP53 protein synthesis prior to DNA synthesis during the cell cycle (Boggs and Reisman, 2006; Takahasi et al., 2011) and this has been hypothesised to be to ensure a rapid response to DNA damage before exiting S-phase (Takahasi et al., 2011).

Cystatin A, also known as stefin A, is a cysteine protease inhibitor from family 1 of the cystatin superfamily (Keppler, 2006). It is found in the cytoplasm because, unlike other cystatins, it does not contain a signal peptide for secretion. High levels of expression have been reported in epithelium and lymph nodes.
as well as the liver, spleen, keratinocytes and poly-morphonuclear granulocytes (Keppler, 2006; Barrett 1985, 1986). The cysteine protease family are involved in many biological functions including apoptosis, MHC class II immune responses, prohormone processing and remodelling of the extracellular matrix (Moffitt et al., 2010). As such, cysteine proteases are important in maintaining the normal function of the extracellular matrix and are normally in balance with their inhibitors, the cystatins/stefins. There is growing evidence to suggest that in patients with malignant tumours the expression and activity of both the proteases and their inhibitors are altered (Li et al., 2011).

CSTA has been suggested as an indicator of aggressive sub-types of prostate and breast cancer (Kuopio et al., 1998; Sinha et al., 1999) though generally the expression is lost during progression of lung, prostrate, skin and breast cancer with the loss of the cells expressing CSTA (Keppler, 2006). It has been shown to be upregulated in the saliva of breast cancer patients and suggested as a possible specific biomarker (Zhang et al., 2010). Protein expression of CSTA was predictive of laryngeal carcinoma and negatively linked to lymphatic metastasis, reoccurrence and survival (Li et al., 2011). Increased immunostaining for CSTA was related to a favourable outcome in squamous cell carcinoma patients (Anicin et al., 2013). A study by Jais et al. (2008) identified a 16-gene signature which included CSTA that predicted the survival rate of elderly patients with diffuse large B-cell lymphoma being treated with doxorubicin and a mixture of other drugs. Despite the lack of functional information on CSTA, the growing number of studies that have identified this gene as a potential biomarker for genotoxicity or the identification of various types of cancer suggest it may have future application in such fields.

In this study, the gene expression of RAD51C, TP53 and CSTA in HEPG2 cells exposed to a range of genotoxic compounds was measured by Q-PCR amplification. In parallel, we developed an alternative assay for the same 3 gene signature based on chemiluminescence (Arnold et al., 1989; Nelson et al., 1996a,b) which we have used for gene expression analysis of abundant genes (Thomas-Jones et al., 2003a, 2003b) and the identification of single nucleotide differences in RNA transcripts (Morris et al., 2014). The chemiluminescence method allows the direct quantitation of RNA transcripts by the use of acridinium-ester (AE)-labelled complementary probes and is capable of analytical sensitivity of 10-50 attomol of target with a
routine coefficient of variance of less than 15% (Morris et al., 2012). It consists of three sequential steps: hybridisation, where the probe forms a double stranded duplex with the target and the AE intercalates; hydrolysis, where any AE unprotected by intercalation is hydrolysed rendering it incapable of producing light; and detection, where the light output is triggered. Due to the chemical hydrolysis step, these assays have the ability to distinguish between targets differing by a single base pair allowing high specificity (Nelson et al., 1996a, 1996b).

We have used this technology in past studies to directly measure highly abundance messenger RNA transcripts (eg. vitellogenin mRNA). Here we aimed to determine if this technology was capable of measuring medium abundance genes such as those that would be useful for a genotoxicity screen in vitro and also if it could, due to its reported low CVs, be adapted to detect trends of gene expression changes at lower concentrations of compound than Q-PCR amplification.

2. Materials and Methods

2.1 Human Gene/ Protein nomenclature

The gene and protein nomenclature used in this manuscript is that recommended by the HUGO gene nomenclature committee (Wain et al., 2002).

2.2 Materials

Standard chemicals were obtained from Sigma-Aldrich (St. Louis, MO, USA). Those used in HEPG2 exposures are listed in Tables 3 and 4. Specialised Q-PCR materials and molecular grade glycogen were supplied by Life Technologies (Carlsbad, CA, USA) except the Q-PCR enzyme buffer mix which was supplied by PCR-BIO (PCR BIO, London, UK).

Eurofins MWG Operon (Ebersberg, Germany) supplied all the oligonucleotides with the exception of the modified oligonucleotides. These were synthesised with an internal linker containing a free primary
amine attached to the oligonucleotide via a six carbon aliphatic spacer arm and were supplied by Integrated DNA Technologies (Coralville, IA, USA).

2.3 Cell Culture and exposures

All cell culture growth and exposures on the HEPG2 cell line were conducted by AvantiCell Science Ltd, Ayr, Scotland, according to conditions described below, including initial testing on compound concentration range and solubility. A test on the compound’s solubility in DMSO was performed. For a number of compounds, the top concentration resulted in precipitation when diluted with culture medium for addition to the well. If any precipitation was observed, a range of solvents was tested to determine the optimal solvent.

2.4 Cytotoxicity prescreen

Once an appropriate solvent had been determined, a MTT cytotoxicity test was performed (data not shown). Human HEPG2 cells were cultured under standard growth conditions in a culture medium containing Minimum Essential Medium, heat inactivated foetal calf serum, glutamine, non-essential amino acids, fungizone, and antibiotics. Cells were seeded at an assay-optimised density in a 96 well plate format. Cells were at passage 16 at the time of assay.

Following 48 hours’ incubation at 37°C and 5% CO₂, in relative humidity >95%, the cells were challenged for a 24 hour exposure period with 18 test compounds at 3 concentrations (n=3). The cytotoxicity assay is based upon measuring the leakage of cytosolic enzyme lactate dehydrogenase from dead or dying cells that have lost membrane integrity. Detergent-lysed cells were measured to provide the maximum value, which was then used to calculate % cytotoxicity induced by the test material. In addition to measuring the % cytotoxicity in response to each of the treatments, all cells were examined microscopically, and a visual observation was recorded (data not shown). Staurosporine (10 μM), a cell permeable inhibitor of protein kinases, was included as a positive control to demonstrate assay performance. Appropriate test compound vehicle controls were also included.
In summary, cytotoxicity was detected in response to 4 of the chemicals tested. Benzo(a)pyrene caused cytotoxicity at all three of the concentrations tested (100 μM = 35%, 50 μM = 30% and 25 μM = 12%). Cadmium chloride caused cytotoxicity at all three of the concentrations tested (50 μM = 53%, 25 μM = 51% and 10 μM = 30%). Arsenic (III) oxide was estimated to cause 19% cytotoxicity at 50 μM but none was detected at 25 and 10 μM. Cisplatin caused 21% cytotoxicity at 100 μM. The concentration range was adjusted accordingly in the case of observed cytotoxicity, solubility issues (benzopyrene) or precipitation.

2.5 Genotoxicity testing

Human HEPG2 cells were cultured as described above. Cells were seeded at a density of 3.0 x 10⁶ cells per culture well (6-well plates). Cells were at passage 17 at the time of sample harvest. Following 24 hours’ incubation at 37°C and 5% CO₂, in relative humidity >95%, the cells were challenged with the 18 test compounds for a further 24 hour exposure period. For each test chemical, triplicate wells of HEPG2 cells were exposed to 4 concentrations. Triplicate measurements of a solvent control were included for comparison.

Following the 24 hour exposure period, the cell lysates were prepared by the following method. Media/test compound was removed and the cells were washed gently with 2 ml of sterile phosphate buffered saline (PBS). Care was taken to remove as much PBS as possible from the wells. The cells were lysed in RLT buffer (600 μL) supplied by Qiagen (Venlo, The Netherlands), which was added to each well, then scraped vigorously with a sterile inverted pipette tip. The plate was tipped at an angle to collect as much sample as possible. Samples were collected into nuclease-free non-stick eppendorf tubes (Alpha Laboratories, Eastleigh, UK), containing 40 U of RNase inhibitor (RNase OUT, Life Technologies). The tubes were vortex mixed, placed on ice and stored at -80°C.

Two further exposures were done at a later date: the 4-chloroaniline exposure was repeated with a higher concentration (1000 μM) and doxorubicin exposure was repeated (1 μM) to gain more material for direct measurement experiments.
2.6 RNA extraction

The total RNA was extracted using the Qiagen RNeasy kit according to manufacturers’ instructions and eluted in 50 µl nuclease-free H₂O and 40 U RNaseOUT (Life Technologies) was added before storage. Purity and concentration were assessed by measuring the absorbance at 260 nm on a Nanodrop spectrophotometer (Thermo Scientific, Waltham, MA, USA).

A Bioanalyser 2100 (Agilent, Santa Clara, CA, USA) was used to confirm the integrity of the RNA from the benzopyrene exposure and an average RNA integrity number (RIN) of 7.8 was determined. To further validate the process, two hydroxymethylbilane synthase (HMBS) Q-PCR reactions were performed on the same cDNA and regression analysis was conducted and also HMBS and GAPDH expression were compared (data not shown). Both experiments showed good correlation ($R^2 > 0.9$) suggesting the RNA quality was acceptable and the Q-PCR methodology reproducible.

2.7 cDNA synthesis and Quantitative Polymerase Chain reaction (Q-PCR)

Total RNA (800ng) was reverse transcribed into single-stranded cDNA using the QuantiTect reverse transcription method according to manufacturers’ instructions (Qiagen). A control which included all reaction components and RNA but no reverse transcription enzyme was included for each exposure to ensure there was no genomic DNA contamination present. The cDNA was diluted 1:10 in nuclease-free water before using 4 µL in a Q-PCR reaction with a total volume of 20 µL, containing 10 µL of 2 x qPCRBIO PROBE MIX LO-ROX (PCR BIO, London, UK), 1 µL of TaqMan probe primer mix (Life Technologies) and 5 µL of RNase free water (Ambion, Austin, Texas). The final concentration of each primer was 900 nM and of the 6-FAM™ dye-labelled TaqMan MGB probe was 250 nM. Thermocycling conditions were 95°C for 2 minutes, then 40 cycles of 95°C for 10 s, 60°C for 30 s. All PCR measurements were made in triplicate. TaqMan probes and primer sets were obtained from Life Technologies for the following genes (manufacturer’s part numbers given in brackets): RAD51C (Hs00365220_m1), CSTA (Hs 00193257_m1), TP53 (Hs 01034254_g1), CYP1A1 (Hs01054797_g1). HMBS (Hs 00609293_g1) was measured as a control gene for
all samples. Where HMBS gene expression did not conform to the stringent criteria for use (see below), GAPDH (Hs02758991_g1) and 18S (Hs99999901_s1) were measured.

Q-PCR amplifications were performed on a Stratagene Mx3000P instrument (Agilent, Santa Clara, CA, USA) using a 96 well plate format. MxPro software (version 4.10) was used to adjust raw fluorescence readings by subtracting baseline background, normalizing against the ROX reference dye (dRn) and estimating the threshold to give C(t) values. Several acceptance criteria were applied to each run which must be met for the run to be considered valid: no amplification after 35 cycles in the no-template (NTC) or the no-amplification (NAC) controls and the average standard deviation (SD) of technical replicates within the experiment to be equal or less than 0.3 SD. If these criteria were not met, the experiment was repeated.

2.8 Acceptance criteria applied to control genes

The control gene data was examined for each exposure and a Student’s t-test (two-tailed) applied to each condition against the solvent control. For a gene to be accepted as a control gene and used for data normalisation, it was essential two criteria were met: there must not be a two-fold difference (or greater) in expression over the range of concentrations used (Schmittigan and Livak, 2008) nor a statistically significant difference in gene expression.

For the majority of the exposed samples, hydroxymethylbilane synthase (HMBS: accession number NM_000190), the gene for a key enzyme in the haeme biosynthetic pathway, proved to be an appropriate control gene and was used to normalise the data. In the case of doxorubicin, nickel chloride, sodium dichromate, arsenic oxide and copper sulfate, two other control genes were investigated: 18S and GAPDH. The gene found to have the least variable expression for each individual compound according to the criteria detailed above was chosen for data normalisation.

2.9 Statistical treatment

The Q-PCR data were analysed by relative expression analysis as described in the paper by Schmittgen and Livak (2008) to generate \( 2^{-\Delta\Delta C(t)} \) values after normalisation against the most stable reference
gene and then the mean was expressed as a ratio against the solvent control (0 µM). A two-tailed Student’s t-test, was applied to each condition to identify statistically significant differences between the solvent control and the exposure samples.

2.10 Heat map and hierarchical clustering of genotoxic compounds based on their gene profiles

The gene profile at all concentrations was used to generate a heat map, from which hierarchical clustering analysis was performed. The data treatment was as follows: $2^{-\Delta Ct}$ values were calculated, and then the mean was expressed as a ratio against the solvent control value before log transformation. The hierarchical cluster analysis was performed using a one minus Pearson Correlation Coefficient with average linkage. Data analysis was performed with GENE-E software (Broad Institute, MA, USA). Software can be downloaded from http://www.broadinstitute.org/cancer/software/GENE-E/download.html.

2.11 Chemiluminescent AE-probe preparation

Probe sequences were designed by alignment of all isoforms found in Genbank using CLC Main workbench (version 6.9) from CLC Bio (Aarhus, Denmark), and selecting regions where all isoforms were conserved. The probe sequences are listed below (Table 1) and # designates the position of the linker. In all cases, the reverse complement sequence was used to make a dilution series for the standard curves.

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Sequences of modified probe oligonucleotides (# designates the position of the linker).</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>AE- labelled probes</strong></td>
<td></td>
</tr>
<tr>
<td>Hs RAD51C</td>
<td>$5'$ GAA GCC CTG GG # TAT GCT CCT GCT C $3'$</td>
</tr>
<tr>
<td>Hs TP53</td>
<td>$5'$ CCC AAC TGT AGA AAC T # A CCA ACC CAC C $3'$</td>
</tr>
<tr>
<td>Hs CSTA</td>
<td>$5'$ GAT TTC TGG AGT # GGC GGG TTT GGC CTC AG $3'$</td>
</tr>
<tr>
<td>Hs HMBS</td>
<td>$5'$ CCC AGG CAT CT # GTG CCC CAC AAA $3'$</td>
</tr>
</tbody>
</table>

2.12 Probe labelling and purification

Oligonucleotides were labelled with acridinium-ester (NHS-AE) through an internal amine linker and purified using high-performance liquid chromatography (HPLC) based on the procedures described in Nelson et al., (1992) and Morris et al. (2014). HPLC purification was conducted using an instrument from Thermo
Separation Products (Waltham, MA, USA) with automatic gradient elution and equipped with a detector capable of measurement at 260 and 325 nm. The crude labelled oligonucleotide was subjected to purification by gradient elution IE-HPLC using a 4.0 x 125 mm Nucleogen-DEAE 60-7 column (Macherey-Nagel, Düren, Germany).

2.13 Direct method of measuring chemiluminescence

AE-labelled probe (1 pmol per reaction) was added to either ssDNA target or total RNA after dilution in 100 mM lithium succinate, 2 mM EDTA, 2 mM EGTA, 10% lithium lauryl sulfate pH 4.8 followed by hybridisation at 60°C for 30 minutes. A second reagent was added (150 mM sodium tetraborate, 5% Triton X102 pH 8.5) followed by a further 15-20 minutes at 60°C. The temperature was reduced to 4°C and remaining chemiluminescence was measured over 2 s using a Leader 50i luminometer (tube based) (Gen-Probe, San Diego, CA, USA) with sequential injections of 32 mM hydrogen peroxide in 1 mM nitric acid followed by 1.5 M NaOH. For quantitation, a standard curve of reverse complementary ssDNA target was also measured.

The plate format was essentially as described above except the probe concentration was 100 fmol per reaction, the hybridisation and hydrolysis incubations were performed at 65°C on a Techne PCR instrument (Bio-Techne, Minneapolis, MN, USA) with a heated lid at 110°C and a Centro Luminometer LB960 (96-well plate format, Berthold Technologies, Bad Wildbad, Germany) with MikroWin software (Version 4.41) was used to measure the light output.

2.14 Linear amplification step prior to chemiluminescence measurement

For the first strand cDNA synthesis, the following reaction was set up for each RNA sample to be analysed: RNA (1 µg) and nuclease-free H2O to a total volume of 5 µL, then 2 µL dNTPs (10 mM, Clontech UltraPure), 1 µL each of MPXForward primers (RAD51C-MPXF, CSTA-MPXF, TP53-MPXF, HMBS-MPXF) and of MPXReverse primers each (RAD51C-MPXR, CSTA-MPXR, TP53-MPXR, HMBS-MPXR). Primers were diluted to
a concentration of 100 pmol/µL prior to use and sequences are given in Table 2. An additional sample was included as a background control (cycle 1 (C1) control).

Thermocycling conditions were one cycle of 65°C, 5 min, 60°C 3 min, 55°C 3 min, 50°C 60 min then 70°C 15 min in a Techne PCR machine (heated lid at 110°C). During the 50°C step, 8 µl of a mix consisting of 4 µL 5x First strand buffer (250 mM Tris-HCl (pH 8.3), 375 mM KCl, 15 mM MgCl₂), 2 µl dithiothreitol (0.1 M stock, 8 mM final concentration), 40 U RNase OUT and 200 U of Superscript III was added to each reaction and mixed gently by pipetting.

To the ssDNA, 50 µL 2 x AmpliTaq Gold 360 mastermix, 49 µL nuclease-free H₂O and 2 U of RNase H were added, mixed then the following thermocycling conditions were used to synthesise double stranded cDNA: 37°C 20 min, 95°C 10 min, 57°C 5 min, 72°C 30 min. One sample was removed after the first complete cycle (C1) and placed on ice, then cDNA was linearly amplified with a further 14 cycles of 95°C 30 s, 57°C 30 s, 72°C 2 min. All components were sourced from Life Technologies except where indicated.

The dsDNA (8 µL) was used in a chemiluminescent detection assay as described above for the plate format except it was heated to 95°C for 5 minutes before adding AE-labelled probe. The RLU value of the C1 control was subtracted from the data.

<table>
<thead>
<tr>
<th>Multiplex primers</th>
<th>Forward (-MPXF)</th>
<th>Reverse (-MPXR)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CSTA</td>
<td>5′ CCA CAC TTC CCT GTT CAC TT 3′</td>
<td>5′ CAG CTC GTC ATC CTT GTT TTT 3′</td>
</tr>
<tr>
<td>RAD51C</td>
<td>5′ AGC GGG ATT TGG TGA GTT T 3′</td>
<td>5′ ACT CCA CCC CCA AGA ATA TCA 3′</td>
</tr>
<tr>
<td>TPS3</td>
<td>5′ ACA TCT GGC CTT GAA ACC 3′</td>
<td>5′ AAA GAA ATT GAC CCT GAG CA 3′</td>
</tr>
<tr>
<td>HMBS</td>
<td>5′ ACG GCT CAG ATA GCA TAC AA 3′</td>
<td>5′ ATA GCA GTG AGA ATG GGG CA 3′</td>
</tr>
</tbody>
</table>

3. Results

We measured the changes induced in a 3 gene signature in HEPG2 cells in response to 16 genotoxic and 2 non-genotoxic non-carcinogenic chemicals (n=18) with different mechanisms of action. The chemicals
were chosen to give 7 compounds with well-characterised genotoxic mechanisms from Group 1 of the recommended list of genotoxic chemicals (Kirkland et al., 2008). A further 8 compounds were chosen that were not present in the Kirkland recommended list nor in the Westerink study (Westerink et al., 2010), which tested over 200 compounds in total. The 8 compounds were chosen to have a wide range of mechanisms and were very diverse, including metal salts to endocrine disrupting compounds, and so had a less well defined mode of action with regard to genotoxicity. Two chemicals, (+)-menthol and phthalic anhydride, were chosen to act as negative controls from Group 3 of non-carcinogens; these are negative or equivocal for genotoxicity in vivo and show no carcinogenicity in rat and mice studies (Kirkland et al., 2008) and doxorubicin was included as a positive control (Westerink et al., 2010).

The starting point for determining the concentration range for the exposures was taken from the literature (Westerink et al., 2010; McElwee et al., 2009; Song et al., 2009; Verma et al., 2011; Osterburg et al., 2010; Wang et al., 2009; Wang et al., 2013; Kuzbari et al., 2013; Konac et al., 2005; Permenter et al., 2011; Kawata et al., 2009; Chin et al., 1994; Hodges et al., 2004; Hu et al., 2004; Muller et al., 2005). Where possible, several concentrations below the lowest observed effect concentration were examined to provide a stringent test of sensitivity for the different assays tested. A test of solubility and cytotoxicity were performed and Table 3 shows the actual concentrations and solvents used in the exposure experiments.
### Table 3
List of compounds with their associated solvents.

<table>
<thead>
<tr>
<th>Compound</th>
<th>CAS no.</th>
<th>vehicle control</th>
<th>Dose 1 (µM)</th>
<th>Dose 2 (µM)</th>
<th>Dose 3 (µM)</th>
<th>Dose 4 (µM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Doxorubicin</td>
<td>25316-40-9</td>
<td>0.01% DMSO</td>
<td>0.05</td>
<td>0.1</td>
<td>0.5</td>
<td>1</td>
</tr>
<tr>
<td>(±)-Menthol</td>
<td>15356-70-4</td>
<td>0.5% DMSO</td>
<td>1</td>
<td>10</td>
<td>100</td>
<td>1000</td>
</tr>
<tr>
<td>Phthalic anhydride</td>
<td>85-44-9</td>
<td>0.5% DMSO</td>
<td>1</td>
<td>10</td>
<td>100</td>
<td>1000</td>
</tr>
<tr>
<td>Benzopyrene</td>
<td>50-32-8</td>
<td>0.625% Acetone</td>
<td>3.125</td>
<td>6.25</td>
<td>12.5</td>
<td>25</td>
</tr>
<tr>
<td>7,12-Dimethylbenzanthracene (DMBA)</td>
<td>57-97-6</td>
<td>0.25% DMSO</td>
<td>1</td>
<td>5</td>
<td>10</td>
<td>50</td>
</tr>
<tr>
<td>2-(Acetylamino)fluorene (2-AAF)</td>
<td>53-96-3</td>
<td>0.5% Acetone</td>
<td>50</td>
<td>100</td>
<td>250</td>
<td>500</td>
</tr>
<tr>
<td>Cadmium chloride</td>
<td>10108-64-2</td>
<td>0.05% DMSO</td>
<td>0.1</td>
<td>0.5</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>Cisplatin</td>
<td>15663-27-1</td>
<td>0.05% DMSO</td>
<td>1</td>
<td>5</td>
<td>10</td>
<td>50</td>
</tr>
<tr>
<td>4-Chloroaniline (first expt)</td>
<td>106-47-8</td>
<td>0.5% DMSO</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>1000</td>
</tr>
<tr>
<td>4-Chloroaniline (second expt)</td>
<td>106-47-8</td>
<td>0.5% DMSO</td>
<td>10</td>
<td>50</td>
<td>100</td>
<td>500</td>
</tr>
<tr>
<td>2,4-Diaminotoluene</td>
<td>95-80-7</td>
<td>0.5% DMSO</td>
<td>10</td>
<td>50</td>
<td>100</td>
<td>500</td>
</tr>
<tr>
<td>Nickel (II) chloride</td>
<td>7791-20-0</td>
<td>H₂O</td>
<td>10</td>
<td>50</td>
<td>100</td>
<td>500</td>
</tr>
<tr>
<td>Sodium dichromate</td>
<td>7789-12-0</td>
<td>H₂O</td>
<td>5</td>
<td>10</td>
<td>50</td>
<td>100</td>
</tr>
<tr>
<td>Arsenic (III) oxide</td>
<td>1327-53-3</td>
<td>0.1% NaOH</td>
<td>1</td>
<td>5</td>
<td>10</td>
<td>50</td>
</tr>
<tr>
<td>Lead (II) acetate</td>
<td>6080-56-4</td>
<td>H₂O</td>
<td>5</td>
<td>10</td>
<td>50</td>
<td>100</td>
</tr>
<tr>
<td>Sodium tungstate</td>
<td>10213-10-2</td>
<td>H₂O</td>
<td>50</td>
<td>100</td>
<td>500</td>
<td>1000</td>
</tr>
<tr>
<td>Copper sulfate</td>
<td>7758-98-7</td>
<td>H₂O</td>
<td>10</td>
<td>50</td>
<td>100</td>
<td>500</td>
</tr>
<tr>
<td>Hydroxyurea</td>
<td>127-07-1</td>
<td>H₂O</td>
<td>10</td>
<td>50</td>
<td>100</td>
<td>500</td>
</tr>
<tr>
<td>Diethylstilbestrol</td>
<td>56-53-1</td>
<td>0.05% MeOH</td>
<td>1</td>
<td>5</td>
<td>10</td>
<td>50</td>
</tr>
</tbody>
</table>

#### 3.1 RNA concentration

There was no statistically significant difference in the concentration of total RNA in the solvent control (0 µM) compared to RNA extracted from exposed HEPG2 samples for most chemicals in this study, suggesting that, in general, the chemicals chosen had no effect on global transcription based on a Student’s t-test, two tailed between the solvent control and each concentration of compound used in the experiment.

There were three exceptions to this (Fig. S1). As sodium dichromate concentration increased to 50 µM and above, the concentration of RNA decreased (one-way ANOVA, p = 0.0008) suggesting decreased transcription. Doxorubicin had a complex trend over the concentrations in this study with an initial increase and then a decrease (one-way ANOVA, p = 0.0039) which only reached significance at 0.1 µM (Student’s t-
test against the solvent control). However, copper sulfate showed a striking decrease in RNA concentration over all four concentrations compared to the solvent control suggesting a large impact on global transcription (one-way ANOVA, p = 4.58E-06). To normalise for any global effect on RNA transcription, a standard amount of RNA (800ng) was used in cDNA synthesis.

3.2 Gene expression analysis using Q-PCR

The gene expression of RAD51C, TP53 and CSTA was measured, using Q-PCR based on TaqMan primers and FAM-probes in an in vitro system based on the human liver carcinoma cell line, HEPG2. Overall, the Q-PCR results showed excellent technical replication, with an average standard deviation of ≤ 0.1 (0.3% coefficient of variance (CV)) and an average inter-sample CV of 23%.

The trends in gene expression for each compound are shown in Fig. 1, 2 and 3. Two chemicals were included in the study as good examples of non-genotoxic non-carcinogenic chemicals based on the recommendations of Kirkland et al. (2008): phthalic anhydride and menthol (Fig. 1A & B). Our study showed no trend of dysregulation of gene expression for the four genes (Table 4) over a wide range of concentrations (0, 1, 10, 100 and 1000 µM) for either of these chemicals.

Doxorubicin (DOX) was included in the study as previous studies have reported a strong response to genotoxic insult (Westerink et al., 2010) (Fig. 1C). This was confirmed for DOX exposure in this study with a strong change in gene expression for all three genes with increasing concentration. This was despite the use of lower concentrations of DOX than any other compound in the study.

Position for Figures 1, 2 and 3 (provided separately)

Figure 1, 2, 3. Q-PCR analysis of 3-gene signature. Relative gene expression (log2 transformation of the genes after normalisation) plotted against the concentration of the chemicals (µM). The symbol —— desig-nates RAD51C expression, ▲ TP53, □— CSTA, ○— HMBS, ×— GAPDH / 18S. Figure 1 shows three compounds expected from previous studies to give a clear genotoxic
(doxorubucin) or non-genotoxic result (menthol and phthalic anhydride). **Figure 2** shows seven compounds included in the Kirkland list of genotoxic compounds. Results from two separate experiments are shown for chloroaniline. **Figure 3** shows a further eight compounds not included in the Kirkland list.

### Table 4
Overview of the gene expression changes detected using Q-PCR for the chemicals included in the study.

<table>
<thead>
<tr>
<th>Compound Designation</th>
<th>Compound Name</th>
<th>Q-PCR</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td><strong>RAD51C</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>TP53B</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>CSTA</strong></td>
</tr>
<tr>
<td><strong>Non-genotoxic non-carcinogenic</strong> (Kirkland et al., 2008))</td>
<td>(1) Menthol</td>
<td>Dark grey</td>
</tr>
<tr>
<td></td>
<td>Phthalic anhydride</td>
<td>Light grey</td>
</tr>
<tr>
<td><strong>Positive control</strong> (Westerink et al., 2010))</td>
<td>Doxorubicin</td>
<td>Light grey</td>
</tr>
<tr>
<td><strong>Genotoxic chemicals</strong> (Kirkland et al., 2008))</td>
<td>Benzopyrene</td>
<td>Light grey</td>
</tr>
<tr>
<td></td>
<td>7,12-Dimethylbenzanthracene</td>
<td>Light grey</td>
</tr>
<tr>
<td></td>
<td>2-(Acetylamino)fluorene</td>
<td>Light grey</td>
</tr>
<tr>
<td></td>
<td>Cadmium chloride</td>
<td>Light grey</td>
</tr>
<tr>
<td></td>
<td>Cisplatin</td>
<td>Light grey</td>
</tr>
<tr>
<td></td>
<td>4-Chloroaniline (first expt)</td>
<td>Light grey</td>
</tr>
<tr>
<td></td>
<td>4-Chloroaniline (second expt)</td>
<td>Light grey</td>
</tr>
<tr>
<td></td>
<td>2,4-Diaminotoluene</td>
<td>Light grey</td>
</tr>
<tr>
<td><strong>Less well defined mode of action with regards to genotoxicity</strong></td>
<td>Nickel chloride</td>
<td>Light grey</td>
</tr>
<tr>
<td></td>
<td>Sodium dichromate</td>
<td>Light grey</td>
</tr>
<tr>
<td></td>
<td>Arsenic oxide</td>
<td>Light grey</td>
</tr>
<tr>
<td></td>
<td>Lead acetate</td>
<td>Light grey</td>
</tr>
<tr>
<td></td>
<td>Sodium tungstate</td>
<td>Light grey</td>
</tr>
<tr>
<td></td>
<td>Copper sulfate</td>
<td>Light grey</td>
</tr>
<tr>
<td></td>
<td>Hydroxyurea</td>
<td>Light grey</td>
</tr>
<tr>
<td></td>
<td>Diethylstilbestrol</td>
<td>Light grey</td>
</tr>
</tbody>
</table>

**Legend:** Table 4 shows the difference in the gene expression between the solvent control (0 µM) and the highest concentration tested (Student’s t-test, $p \leq 0.05$). Dark grey designates a statistically significant difference and light grey designates no statistically significant difference.
Across the study, most compounds with a positive response affected the expression of more than one gene (Table 4). Ten compounds affected the expression of CSTA whereas seven affected TP53 or RAD51C expression. Only four compounds showed changes in one target gene – BaP and chloroaniline (TP53) and nickel chloride and 2-(acetylamino)fluorene (CSTA), suggesting a weaker response; this was also suggested by the higher concentration required of two of these compounds to stimulate a response (chloroaniline: 1000 µM and 2-(acetylamino)fluorene: 500 µM). The fact that the majority of compounds showed a change in at least one of these three genes, despite the stringency of the exposure test, suggests that this 3-gene signature has potential as an indicator for genotoxicity for a variety of mechanisms of action.

A number of compounds show a statistically significant RAD51C increase suggesting direct DNA damage and the subsequent induction of the DNA damage repair pathways (Table 4). These include doxorubicin (Fig 1C), cisplatin (Fig 2D) and hydroxyurea (Fig 3G) and to a lesser degree cadmium chloride (Fig 2C). Initial increases in RAD51C expression were replaced by decreases in higher concentrations for sodium dichromate (Fig 3B), arsenic oxide (Fig 3C) and copper sulfate (Fig 3H), suggesting inhibition of DNA repair.

The mechanism of TP53 activation is well established as being post-translational (Lakin and Jackson, 1999). However, this study detected some statistically significant changes in TP53 RNA expression: the following p values were derived from a Student’s t-test between the solvent control and the highest concentration of compound used in the experiment. There were very modest changes in expression with BaP (1.69 fold, p = 0.03), DOX (1.28 fold, p = 0.03) and chloroaniline (1.21 fold, p = 0.04). More strikingly, we also found evidence of decreased transcription in the presence of diethylstilbestrol (0.73 fold, p = 0.001), sodium dichromate (0.27 fold, p = 0.001), arsenic oxide (0.08 fold, p = 0.001) and at the three highest concentrations of copper sulfate (50, 100 and 500 µM), reaching a maximum decrease at 0.16 fold (p = 0.0008). All of these exposures were checked for cytotoxicity before use, and of the compounds that changed TP53 expression, only benzopyrene and arsenic showed some evidence of cytotoxicity at the highest dose. Benzopyrene was estimated to cause 12% cytotoxicity at 25 µM and arsenic (III) oxide was estimated to cause 19% cytotoxicity at 50 µM though none was detected at 25 and 10 µM. There was no cytotoxicity detected with any of the other compounds that changed TP53 expression. The TP53 protein
requires zinc for normal protein folding into a conformation capable of specific DNA binding (Meplan et al., 2000). The presence of 30 μM copper is sufficient to displace zinc and cause an abnormal conformation and altered transcriptional properties (Hainaut et al., 1995, Tassabehji et al., 2005). There have been previous reports of altered TP53 mRNA expression in response to copper (VanLandingham et al., 2002; Narayanan et al., 2001; Obata et al., 1996; Tassabehji et al, 2005.; for a review see Phatak and Muller, 2015).

CSTA dysregulation was the single most effective indicator in this study, changing in the presence of 10 out of 16 compounds studied (4 out of 8 known genotoxic agents). This is in contrast to the findings of Westerink et al. (2010) who found that the HEPG2 p53_luc reporter assay had the highest sensitivity both across 62 compounds tested (85%, ECVAM (European Centre for the Validation of Alternative Methods) compound set) and also across the same compounds used in this study (8 out of 8 compounds, including doxorubicin) gave a positive result compared with 6 and 5 for the HEPG2 cystatin A_luc and RAD51C_luc reporter assays respectively, though this may reflect the differences in methodology, e.g. four TP53-responsive elements were used to construct the TP53-luciferase reporter constructs. In this study, CSTA expression reached 294-fold for doxorubicin and 12.7-fold for cisplatin relative to the expression levels in the absence of compound.

Some of the strongest responses seen for the different compounds were found with the set that contained metal salts, which changed the expression of all genes examined. Sodium dichromate and arsenic oxide showed initial increases in gene expression at the two lower concentrations followed by a marked decrease. Copper sulfate however, changed CSTA expression by 3.43 fold (p = 0.004) and expression of RAD51C by 0.35 fold (p = 0.0003) and TP53 by 0.16 fold (p = 0.0008).

However, two compounds (7,12-dimethylbenzanthracene and 2,4-diaminotoluene) did not show a significant dysregulation of the 3-gene signature and two (benzopyrene and 2-(acetylamino)fluorene) did not show as strong a response as expected despite the use of concentrations comparable to those used in the Westerink study (2010). A common feature of these four compounds is that they all require metabolic activation before causing toxicity. A number of PAH compounds are metabolised by CYP1A1 including
benzopyrene and 2-(acetylamino)fluorine (Muñoz and Albores, 2011). The expression of several genes required for detoxification and metabolic activation, GST2A, GADD45α and CYP1A1, was confirmed in the HEPG2 cell line (data not shown). The expression of CYP1A1 in response to benzopyrene was measured by Q-PCR and found to increase with increasing concentration as expected (168 fold at 25µM, p = 2.8E−05, data not shown). This suggests to us that the reason for the underlying difference between the two studies is not the ability of the cells to respond to the compound nor the batch of chemical used but is due to methodological difference, for example measurement of native gene expression compared to the reporter gene approach utilised by the Westerink study.

3.3 Hierarchical cluster analysis of the Q-PCR results

To organise the compounds into categories with similar gene profiles, a heat map was generated with GENE-E, a matrix visualisation and analysis platform, followed by hierarchical clustering. The hierarchical cluster analysis was performed using a one minus Pearson Correlation Coefficient with average linkage (Fig. 4). The clustering suggested two major groups. The first group contains the non-carcinogenic chemicals menthol and phthalic anhydride and a cluster of compounds that caused changes in 2 or 3 genes. Group 2 contained compounds with more complex profiles where both increases and decreases within an expression profile were observed.
Figure 4: Heat map and hierarchical clustering of genotoxic compounds based on Q-PCR gene profiles. The gene profile at all concentrations was used to generate a heat map, from which hierarchical clustering analysis was performed in order to categorise the compounds with similar characteristics. The hierarchical cluster analysis was performed using a one minus Pearson Correlation Coefficient with average linkage. Data analysis was performed with GENE-E data visualisation software. The horizontal axis of the dendrogram (tree diagram) represents the distance or dissimilarity between clusters. The intensity of the red and blue colour indicate high and low expression levels, respectively, relative to the mean (see scale bar) up to a maximum of +3 or -3 fold. Details of the concentrations of each compound used are shown in Table 1. CTR designates the solvent control data (0 µM) for each experiment. * designates the cluster discussed in the results section.
The hierarchical clustering suggested a close relationship between the gene profiles of doxorubicin and cisplatin, both compounds that directly interact with and damage DNA by either intercalating within it or forming DNA adducts (Tacar et al., 2012; Chu, 1994). These belonged to a larger group of clustered compounds (cluster designated by * on Fig 4), which also included a cluster of cadmium chloride, hydroxyurea and lead acetate. With this group of compounds, expression of both RAD51C and CSTA increased with increasing concentration but cadmium chloride, hydroxyurea and lead acetate had a general trend of decreasing TP53 expression rather than the increase measured for doxorubicin and cisplatin (though few of the trends in TP53 expression reached statistical significance). Underlying this is a different mechanism of action; the literature suggests predominantly an indirect mechanism of DNA damage for cadmium chloride (via oxidative stress and apoptosis (Waisberg et al., 2003)), hydroxyurea (oxidative stress via base oxidation (Sakano et al., 2001, Kovacic, 2011)) and lead acetate (associated with increased reactive oxygen species (ROS) production (Hernandez et al., 2009)), as opposed to the direct DNA interaction of doxorubicin and cisplatin.

This analysis suggested that there might be many trends of perturbed gene expression at lower concentrations, which only reached statistical significance at the highest concentrations of compound, i.e. only relatively large changes of expression were identified as statistically significant. Assay sensitivity is therefore a limiting factor and could potentially be improved by a technology with less associated measurement error. We investigated a technology based on a chemiluminescent acridinium-ester, which is capable of measuring RNA transcripts directly and has high reported accuracy (low coefficient of variation (CV)). The chemiluminescent method is capable of detecting and quantitating RNA transcript concentration directly from RNA with no need to convert to cDNA nor is an amplification step required, making it quicker and easier to use. A specialist dual injection luminometer is required but all other reagents and equipment needed are available in the average molecular biology laboratory. An added advantage is that as this technology uses no enzymic steps, it has potential to provide a quicker, cheaper and more robust method for measurement of gene expression over those that do.
3.4 Chemiluminescent detection assay analysis:

Probes were designed for CSTA, TP53, RAD51C and HMBS, synthesised with a non-nucleotide amino linker and labelled with acridinium ester. Initially, the assay was performed in a tube based format using a Leader 50i luminometer as described in the Materials and Methods. Briefly, the AE probe is hybridised to the target for 30 minutes at 60°C followed by chemical hydrolysis of any unbound probe for 15-20 minutes. The performance of each probe was tested by measuring a dilution series of an oligonucleotide of the reverse complement sequence in triplicate (1000 – 0.1 fmol). The performance of a representative assay designed to measure CSTA is shown in Figure 5A and B and the performance characteristics are summarised in Table 5. The inter-assay variation for the CSTA assay was ~10% coefficient of variance (CV) (Table 5). A similar average inter-assay variation was estimated for the other assays (10.4% CV). The limit of detection (LOD), which was defined as the lowest quantity of the target oligonucleotide that can be distinguished from the absence of that oligonucleotide using a Student's t-test (p ≤ 0.05), was 0.5 fmol, in the presence of 1 pmol of AE-probe.
Figure 5: Optimisation of the CSTA chemiluminescent detection assay.
Figure 5. Optimisation of the chemiluminescent detection assay. Panels A, B C and D: The performance of all AE-probes was characterised by measuring a dilution series of reverse complementary DNA oligonucleotide (target) and the results of a representative CSTA AE-probe, are shown. Panels A and B show the tube format and Panels C and D the 96-well format. Panels A and C show the linear regression analyses of the data over a range of 0.01 to 10 fmol target whereas Panels B and D show the data after log transformation (log RLU vs log DNA oligonucleotide concentration). Panel E shows optimisation of CSTA AE-probe concentration (0.5, 0.1 and 0.05 pmol AE-probe) using a serial dilution of a reverse complement oligonucleotide (0.1, 0.5, 1, 5, 10, 100 and 1000 fmol). Panel F To determine if the 96-well format assay could measure CSTA transcript in total RNA, HEPG2 cells were exposed to doxorubicin (0 and 1µM) and total RNA extracted. The CSTA AE-probe (100 fmol per reaction) was used to measure CSTA expression directly in 5µg of HEPG2 RNA.

ND designates ‘not detected’.

Table 5
Performance of the chemiluminescent assay for CSTA (average values from three experiments).

<table>
<thead>
<tr>
<th>Format</th>
<th>EXPT 1</th>
<th>EXPT 2</th>
<th>EXPT 3</th>
<th>Average Inter-assay CV (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tube</td>
<td>9.3</td>
<td>9.6</td>
<td>10.5</td>
<td>9.8</td>
</tr>
<tr>
<td>Plate</td>
<td>17.0</td>
<td>15.4</td>
<td>11.4</td>
<td>14.6</td>
</tr>
</tbody>
</table>

With the tube format, there was appreciable evaporation at 60°C, which limited the range of temperatures that could be used for the hybridisation and hydrolysis steps. The format was then further optimised to a 96-well plate format using the Centro Luminometer LB960 which allowed for higher throughput as well as for higher temperatures to be used (the assay was performed in a PCR instrument with a heated lid at 110°C to prevent condensation). The temperature of both hybridisation and hydrolysis steps was increased to 65°C.
The performance of the plate format was characterised and shown to be sensitive and capable of quantitative measurement over 4 orders of magnitude (Fig. 5C and D). Each probe was tested with different probe concentrations to determine the optimal sensitivity and the results for the CSTA assay are shown in Figure 5E. The sensitivity of the assay improved with the lower concentrations of probe but the dynamic range of the assay also decreased. As expected, as the target concentration passed that of the probe, the light output (Relative Light Units) reached a plateau and further increases of target concentration no longer resulted in an increase in RLU. These changes led to an improvement in the assay sensitivity with assays having a limit of detection of 0.1 fmol and often as low as 10-50 attomol. The intra-assay variation from three representative CSTA assays was ~15% CV (Table 5).

3.5 Direct measurement of RNA transcripts with chemiluminescent probes

The Q-PCR data indicated that exposure of HEPG2 cells with 1 µM doxorubicin (DOX) caused a ~300 fold change in the expression levels of cystatin A transcript. A second DOX exposure (0 and 1 µM) was performed to determine if direct chemiluminescent measurement, i.e. directly from the RNA with no further manipulation, was capable of measuring the CSTA transcript. The limit of detection for this assay was determined to be 50 attomol of DNA oligonucleotide. As can be seen from Fig. 5F, CSTA transcript was below the limit of detection of the assay in total RNA (5 µg) from HEPG2 cells. However, in total RNA from HEPG2 cells exposed to 1 µM DOX, there was a detectable signal indicating an increased amount of measureable CSTA RNA transcript equating to ~0.2 fmol target. The expression of the other genes could not be detected with this format (data not shown).

To address the sensitivity of the assay, a linear multiplex amplification step of the RNA transcripts (RADS1C, TP53, CSTA, HMBS) was optimised. In this procedure, the RNA was converted to double stranded cDNA before undergoing a defined number of cycles of amplification then the amount of target present was measured using the chemiluminescent assay as described previously with the addition of an initial 95°C heat step to denature the dscDNA before hybridisation to the AE-probe. The minimal number of cycles of
amplification required to give an amount of cDNA measurable by the chemiluminescent detection assay was determined to be 15 (data not shown).

3.6 Amplified measurement of gene expression using chemiluminescent probes

To measure the gene expression, total RNA (1µg) from the initial DOX exposure was reverse transcribed into dsDNA and then amplified for 15 cycles before being hybridised with AE-labelled probes targeted to the 4 genes of interest (Fig. 6). The results confirmed that doxorubicin at very low concentrations caused clear changes in gene expression of this 3-gene signature (Fig. 4). RAD51C and CSTA showed a statistically significant increase in expression (Student’s t-test, \( p \leq 0.05 \)) showing general agreement between the two technologies. The CSTA measurements showed significant increase for all 4 concentrations, as demonstrated by Q-PCR. However, an increase of RAD51C was detected only for the 2 highest concentrations contrasting to an increase detection for all 4 concentrations with Q-PCR amplification. A modest change of TP53 expression reaching a maximum of 1.28 fold (\( p = 0.03 \)) was detected in the two highest concentrations using Q-PCR; however this was not detected by the chemiluminescent detection assay (CLA, 0.95 fold, \( p = 0.3 \)). The comparative analysis of the three technologies is shown in Table 6 and shows the performance of the two amplification techniques to be similar (15 and 16% CV, respectively, Table 6). However, the maximum CV observed in the experiment was 49% for the amplified samples when using the CLA technology, higher than the 29.6% seen with Q-PCR. When the RNA was measured directly, i.e. no amplification step, a lower average CV of 10% was estimated (Table 6).

To our knowledge, this is the first time the technology has been shown to be capable of quantitatively measuring messenger RNA transcripts that are of medium abundance equivalent to the majority of mRNA transcripts present in the cell. This technology has been used previously with a transcription mediated amplification (TMA) step to measure RNA targets in a qualitative fashion or directly to measure highly abundant transcripts (Thomas-Jones et al., 2003a, 2003b; Morris et al., 2014). Teaming this technology with a linear amplification step has a number of consequences; it adds enzymic steps increasing the cost, time to result and observed CVs but it lowers the limit of detection and by doing so, increases the
Figure 6: Measurement of the 3-gene signature after doxorubicin exposure of HEPG2 cells using the amplified chemiluminescent detection method.

RNA (1ug) for doxorubicin exposed HEPG2 cells was reverse transcribed to ds cDNA then amplified for 15 cycles followed by measurement of the 3-gene signature with the chemiluminescent assay. Relative gene expression ($2^{- \Delta \Delta C(t)}$) values calculated after normalisation expressed as a ratio against the solvent control (0 µM) is shown against doxorubicin concentration (µM) ± standard error of mean (SEM). The * symbol designates a statistically significant difference (Student’s t-test, $p \leq 0.05$).
number of genes that can be measured to an extent that it is now a viable option for research with \textit{in vitro} cell lines.

In general, the chemiluminescent assay results reflect the changes determined using Q-PCR though the latter method detected statistically significant gene expression changes in lower doses suggesting it is a more sensitive technique. However, the broad agreement of the amplified chemiluminescent method with the gold standard method, Q-PCR, suggests that this method may have utility in the future and broaden the use for chemiluminescent methods. Further optimisation of the new technique may improve both the sensitivity and the precision of the assay.

<table>
<thead>
<tr>
<th>Table 6</th>
<th>Inter-assay variation for the doxorubicin exposure.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Average CV (%)</td>
</tr>
<tr>
<td>Q-PCR</td>
<td>15</td>
</tr>
<tr>
<td>CLA$^a$ (amplified)</td>
<td>16.1</td>
</tr>
<tr>
<td>CLA$^a$ (direct)</td>
<td>10</td>
</tr>
</tbody>
</table>

$^a$ CLA designates chemiluminescent assay

4. Discussion

Our study confirmed that the 3-gene signature could be measured directly in exposure studies of a human hepatoma cell line, HEPG2. This was despite using an exposure range that was designed to be challenging. We found that $\text{CST}A$ expression was dysregulated by the largest number of compounds, demonstrating a significant change in 10 out of 16 compounds included in this study. For a few chemicals, the increase in $\text{CST}A$ expression was surprising, reaching 12.7 fold for the highest concentration of cisplatin and 294 fold for doxorubicin. The implications of a massive alteration in cytoplasmic $\text{CST}A$ protein levels are hard to decipher but may be an attempt to form a protective environment against the cysteine proteases involved in the apoptotic pathways (Jones \textit{et al}., 1998). Doxorubicin (DOX) is an anthracycline drug used in cancer chemotherapy that works by intercalating DNA, disrupting DNA replication, disrupting the cell cycle and causing DNA damage which ultimately leads to the triggering of the apoptosis pathway in rapidly dividing cells (Tacar \textit{et al}., 2012). DOX causes an apoptotic effect by up-regulating the apoptotic marker, Bcl-
2-associated X protein (Bax) expression (Tacar et al., 2012; El-Moselhy and El-Sheikh, 2014). This leads to downstream activation of the caspases and ultimately apoptosis. The main cytotoxicity mechanism for cisplatin is to form DNA inter- and intra-strand crosslinks though it does also interact directly with RNA and protein (Chu, 1994). The damage triggers apoptosis via the intrinsic pathway at least in part by TP53 activation and by inducing the BAX gene and the downstream caspases. This could explain the strong induction of RAD51C (as part of the DNA repair pathway) and CSTA (as an inhibitor of cysteine proteases) caused by exposure to these two compounds seen in this study.

CSTA is known to inhibit UVB-induced apoptosis caused by caspase 3 in keratinocytes (Takahashi et al., 2007). A study on rat hepatoma cell line provides evidence that CSTA reduces apoptosis induced by bile salts through the interaction with cysteine proteases such as cathepsin B but not caspase 3 and the authors suggest that mechanically cathepsin B is downstream from the caspases (Jones et al., 1998). This raises the possibility that the increase of a specialised member of the family which lacks the secretion peptide is an attempt by the cell to protect against damage by proteases.

We have previously used the chemiluminescent detection assay, which is based on specific hybridisation of oligonucleotides labelled with acridinium-ester, to measure directly high abundance genes such as vitellogenin, zona radiata proteins and beta actin (Thomas-Jones et al., 2003 a,b). Here we demonstrate that the chemiluminescent assay can be adapted to measure genes with much lower levels of expression. In the Q-PCR workflow, cDNA synthesis is considered to be the least efficient step of the process and gene expression studies might be greatly enhanced if this step could be removed. Here we show that, in some circumstances, the changes in expression of an averagely abundant gene, CSTA, can be detected without the need for cDNA synthesis or amplification. The ability to measure RNA transcripts directly and accurately without the need for expensive and time-consuming steps may find application. Addition of a linear amplification step in order to quantify amplified products may also increase the utility of this technology. However, the disadvantage of introducing any form of target amplification is the probability of increasing the error associated with measurement and as such, the precision of the assay decreases.
This modified format may increase the types of application suitable for this technology, hence increasing its utility. The confirmation that the addition of a linear amplification step into the chemiluminescent assay is still capable of quantitative measurements is an important extension of the capability. Our study demonstrates that this novel method of measuring gene expression is capable of generating similar data to the gold standard technology, Q-PCR, and suggests that this adapted method may broaden the use for chemiluminescent methods.

Acknowledgments

This study was supported by funding from the Welsh Government under the CIRP A4B scheme [A4B2122TF]. In addition, PK was a co-inventor of a patent ‘Genetic Biosensors using Chemiluminescence’ (European patent 00971618.4), and PK and CM were co-inventors on a patent ‘Detection of short non-coding RNA using chemiluminescence labelled nucleic acid probes’ (GB1320560.4). A director of Invitron, Professor Stuart Woodhead and, in their roles as the Directors of CatCelt, Professor Keith Smith and Gamal El-Hiti advised on the technology and experimental design as part of the oversight committee. Invitron loaned equipment for part of the experimental work. We would especially like to thank Drs. Colin J. Wilde, Tracey Duff and Austin Logan at AvantiCell Science Ltd, Ayr, Scotland, for conducting the cell line exposures and for their helpful advice and feedback. We would also like to acknowledge Mike O’Reilly for his technical advice on HPLC purification and Ms Joan Hubbard for her technical assistance through the work. G. A. El-Hiti extends his appreciation to the Deanship of Scientific Research at King Saud University for its funding for his research through the research group project RGP-239.
References


Wang, Y.T., Tzeng, D.W., Wang, C.Y., Hong, J.Y., Yang, J.L., 2013. APE1/Ref-1 prevents oxidative inactivation of ERK for G(1)-to-S progression following lead acetate exposure. Toxicology 305, 120–129.


Supplementary Figure 1: Changes in global gene expression.

- **Sodium dichromate (µM)**
  - RNA concentration (ng/µL)
  - Concentrations at 0, 5, 10, 50, and 100 µM with error bars.

- **Doxorubicin (µM)**
  - RNA concentration (ng/µL)
  - Concentrations at 0, 0.05, 0.1, 0.5, and 1 µM with error bars.

- **Copper sulfate (µM)**
  - RNA concentration (ng/µL)
  - Concentrations at 0, 10, 50, 100, and 500 µM with error bars.