

emerging real-time PCR applications

Since its introduction on the commercial market little more than 10 years ago, real-time PCR has become the main technical platform for nucleic acid detection in research and development, as well as in routine diagnostics. In 2007 the real-time PCR market revenue in the US was estimated at \$740 million with annual growth of more than 10%. In this article latest developments and future expectations are presented.

In 1983 Kary Mullis at Cetus conceptualised the most important of biotechnological reactions – the polymerase chain reaction (PCR)¹. The idea, for which he was awarded the 1993 year's Nobel Prize in chemistry, was as simple as brilliant. Based on the natural ability of polymerases to copy nucleic acids in the presence of short complementary oligonucleotides, Kary Mullis reasoned that using a heat stable polymerase, the reaction could be automated to perform multiple copying by cycling the temperature. Heating to 95° would separate the strands of the double stranded nucleic acid template, a temperature between 50-60° will allow primer oligonucleotides to hybridise to the complementary template strands, and raising the temperature to 72° will activate the heat stable polymerase from *Thermus aquaticus* to extend the primers generating a copy of the template molecule. With PCR virtually any DNA molecule could be amplified to large amounts from a single copy. The technique became rapidly important in biological research, for species identification, forensic applications, and qualitative infectious disease testing. When Bob Griffith and Russell Higuchi accidentally ran a PCR in the presence of ethidium bromide, intended for post PCR gel staining, and found that the reaction was not inhibited they realised it should

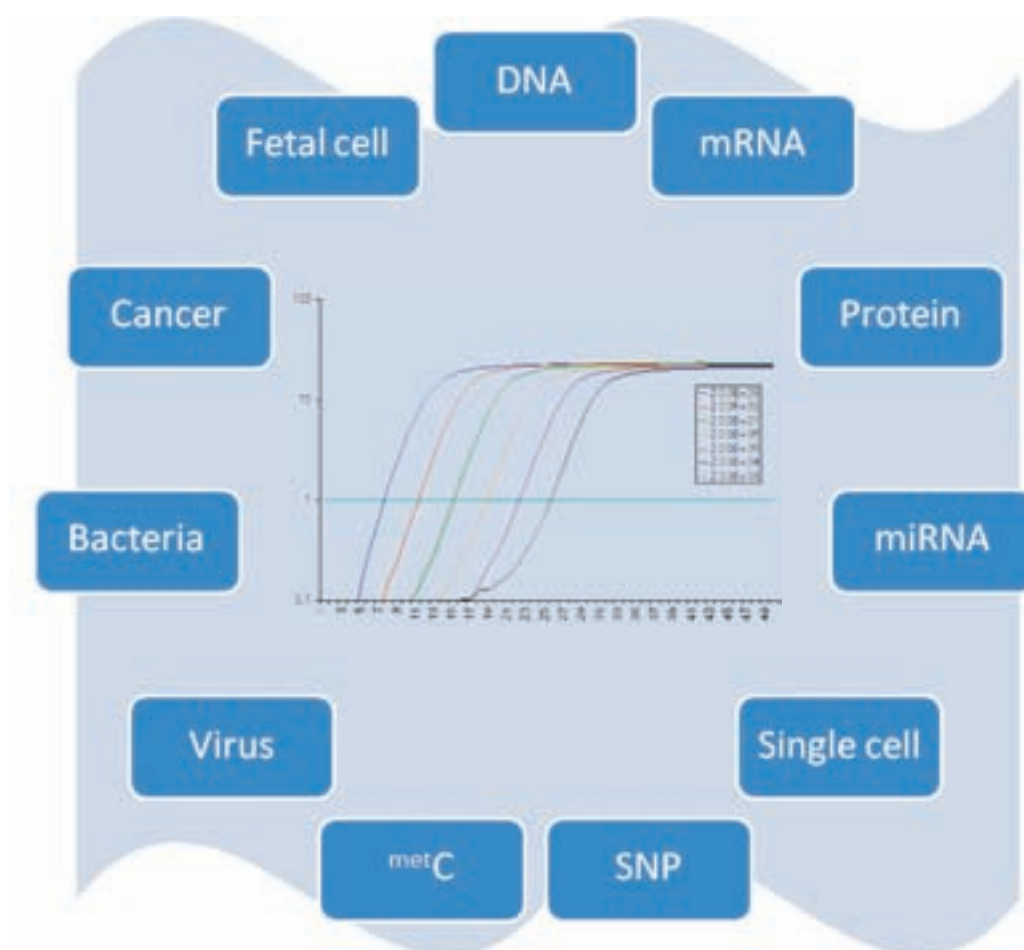
be possible to monitor the amplification in real-time through the dye's fluorescence². By monitoring the number of amplification cycles needed to produce the amount of product that generates certain threshold fluorescence signal the PCR became quantitative. In practice the fluorescence is measured during the on-going reaction giving rise to the technique's popular name today: real-time PCR (Figure 1).

The intercalative dye ethidium bromide, which was used to demonstrate the concept of real-time PCR, was soon replaced with asymmetric cyanine dyes. These bind in the groove of DNA and are much less inhibitory to the PCR. In particular SYBR Green I from Invitrogen has become popular. Dyes are sequence non-specific reporters. They do not distinguish between PCR products and bind also to any aberrant products that may form from interacting primers only (so called primer-dimer products). These false positive signals can be distinguished from that of the target amplicon by melt curve analysis. A melt curve is recorded by slowly heating the PCR product while monitoring the fluorescence. Intensity decreases slowly with temperature due to increased dynamics of the bound dye, but when the temperature is reached that denatures the DNA double helix the separation of strands releases the dye and the fluorescence drops. The

Dr Mikael Kubista

Molecular Biology

Figure 1
Targets for emerging real-time
PCR applications



temperature required to denature a PCR product is called the melting temperature (T_m) and depends on the DNA length, base composition and sequence. Aberrant PCR products are usually shorter than the target amplicon and have a lower T_m .

Dyes are cost-efficient reporters but they are suitable only when a single target is amplified per tube. Multiple targets can be quantified using sequence specific reporters. The hydrolysis probe, also known as Taqman, was the first probe invented³. It is an oligonucleotide labelled with a fluorophore and a quencher that are close to each other in the intact probe, such that the fluorescence is quenched. Upon extension of primers by the Taq polymerase the enzyme's nuclease activity degrades the probe, the labels are released and the free fluorophore fluoresces. Other probe designs include the Molecular Beacons⁴, adjacent probes⁵, and LightUp probes⁶. These are also oligonucleotides, but designed to fluoresce upon hybridisation to target already. Third category reporters are labelled primers that generate fluorescence when incorporated into amplicons. These include

Scorpion primers⁷, Amplifluor primers⁸, and QZyme. Labelled primers are suitable for multiplexing, but any aberrant products they form are fluorescent which limits assay sensitivity. A recent innovation is the Universal ProbeLibrary from Roche. It is based on nuclease probes containing Locked Nucleic Acid (LNA) nucleotides. The LNA bases confer stability to the DNA duplex and 8-9 nucleotides are sufficient for hybridisation. Because of their short length each probe is complementary to several genes, and the library of selected 90 probes is good enough to cover virtually the entire transcriptome.

High resolution melting

Recently instruments that offer high temperature resolution, such as the MR-1 and Lightscanner from Idaho Technologies, the RG-6000 from Corbett Research, the LC480 from Roche, and the 7500 Fast Real-time PCR System from Applied Biosystems, have become available. Using dyes whose fluorescence sensitively reflect melting, such as BEBO from TATAA Biocenter, LCGreen from Idaho Technology,

SYTO9 from Invitrogen, and EVAGreen from Biotium, it is possible to distinguish even single base variants with High Resolution Melting (HRM)⁹. The A:T to G:C interchange, which is the most common single base variation in the human genome (about 84% of cases) results in a difference of about 1 °C in T_m , which is readily detected. Other base changes give rise to smaller differences in T_m , resulting in lower specificity that depends on exact base change and also on the surrounding sequence. These cases can still be addressed by mixing in a known sequence. If the added sequence is the same as the one in the test sample it will have no effect on T_m , while a biphasic melt curve will be obtained if the sequences are different because of heteroduplex formation. HRM can be further improved by including a short unlabelled oligonucleotide probe that hybridises to the region containing the presumptive SNP. Uneven PCR primer concentrations are used to produce excess of this strand, which then hybridises to the unlabelled probe that is designed to generate large differences of the T_m s of the interesting sequence variants.

Since HRM relies on uncomplicated dye reporters it is today the most cost-efficient method to test for known sequence variations. Applications include SNP genotyping for allelic prevalence in a population for the identification of candidate predisposition genes, DNA fingerprinting, screening for loss of heterozygosity, association (case/control) studies, characterisation of haplotype blocks – ‘hap maps’, mutation discovery (gene scanning), somatic acquired mutation ratios, microsatellite analysis (tandem repeats), species identification/taxonomy, and HLA compatibility typing.

A recent application of HRM with great promise is DNA methylation analysis¹⁰. 2-5% of cytosines in eukaryotic genomes are methylated in the 5-position. Methylation occurs in CpG islands upstream of promoters, where it has a regulatory function. The methylation pattern is preserved throughout cell divisions by maintenance mechanisms, but may change in disease, for example, silencing tumour suppressor genes. Methylated sites can be identified using HRM by first treating the sample with bisulfite. This converts unmethylated cytosine to uracil, which in the PCR is transcribed to adenine, which in turn is transcribed to thymine. Methylated cytosine is resistant to bisulfite treatment. Hence, normal C:G base-pairs in the template end up as A:T base-pairs in the PCR product, while ^{met}C:G base-pairs end up as regular C:G base-pairs. This sequence difference is readily detected by HRM, and it is even possible to estimate the degree of promoter methylation from the T_m (Figure 2).

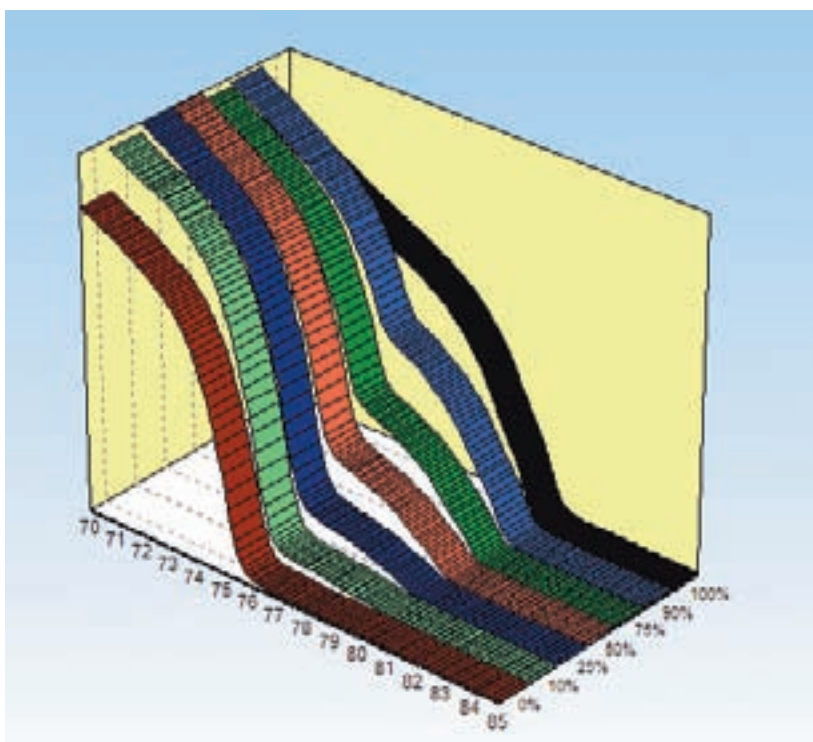
Pathogen detection

Real-time PCR was early introduced for infectious disease testing and kits are today offered from Roche, Abbott, Qiagen, Biotools, Roboscreen and LightUp Technologies. All assays are probe-based to minimise the rate of false positives¹¹. Recent developments include multiplex assays, fully automated systems such as the COBAS® AmpliPrep/COBAS® TaqMan® from Roche, and freeze dried reagents and portable instrument for field use such as the R.A.P.I.D. cyclor from Idaho Technologies. Assays for targets with high mutation rates, such as HIV and HCV are optimised for subtypes and clades prevalent in developed countries, and may have poor sensitivity for the variants common in the parts of the world where these tests are most needed.

miRNA detection

Recently, small RNA molecules called microRNA (miRNA), were discovered that regulate cellular processes by affecting the translation or stability of target mRNA. More than one-third of human genes are thought to be regulated by miRNAs and several studies report that expression levels of miRNAs are altered in several human cancers, suggesting miRNAs are potential disease and prognostic markers¹². The short length of miRNAs (22-24 bases), which is essentially the length of a normal PCR primer, makes it a challenge to detect them

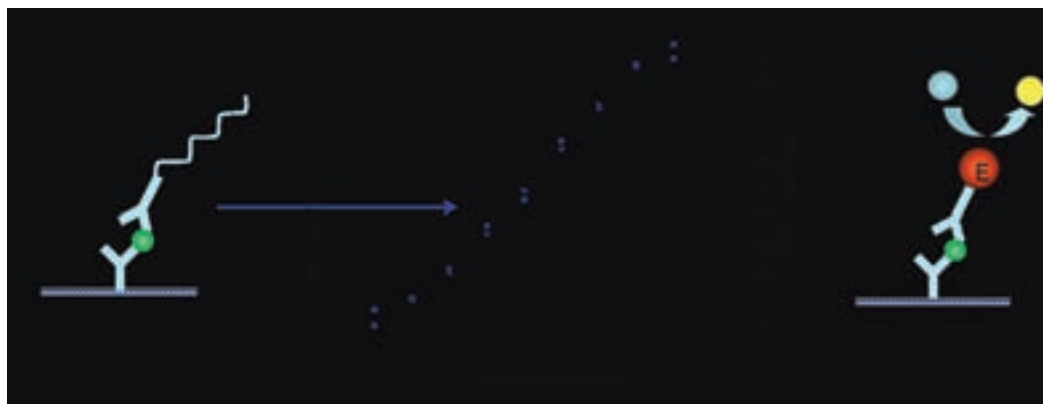
Figure 2
Methylation analysis with High Resolution Melting. The seven samples contained 0%, 10%, 25%, 50%, 75%, 90% and 100% thylated material. Data presented in GenEx



Molecular Biology

Figure 3

Comparison of ELISA and immuno QPCR for the detection of prostate specific antigen (PSA). Both assays are based on antibodies provided from Fujirebio Diagnostics. The immune QPCR assay shows much wider dynamic range and greater sensitivity than the corresponding ELISA assay



with high specificity. Applied Biosystems has developed Taqman MicroRNA assays based on stem-loop primers to achieve specificity, and Exiqon has developed miRCURY LNA microRNA PCR System based on LNA, which makes it possible to use shorter primers.

Figure 4

3-way expression profiling.

Left: Temporal expression profiles measured after replacing ethanol for glucose as nutrient in four strains of yeast.

Right: Principal Component Analysis (PCA) of matrix augmented data. Strains are distinguished by symbol and gene groups by color. Data presented in GenEx

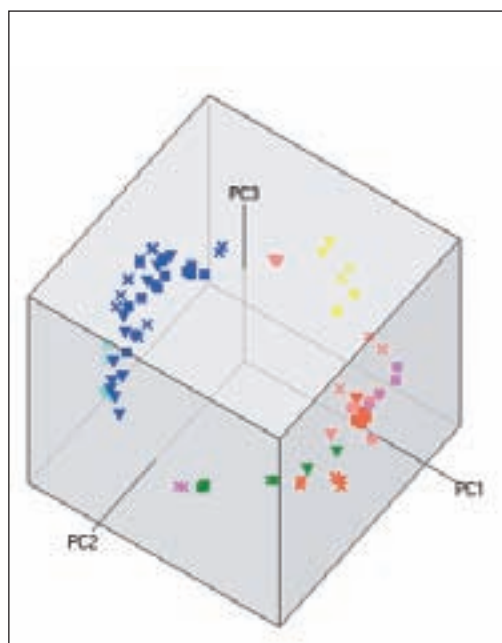
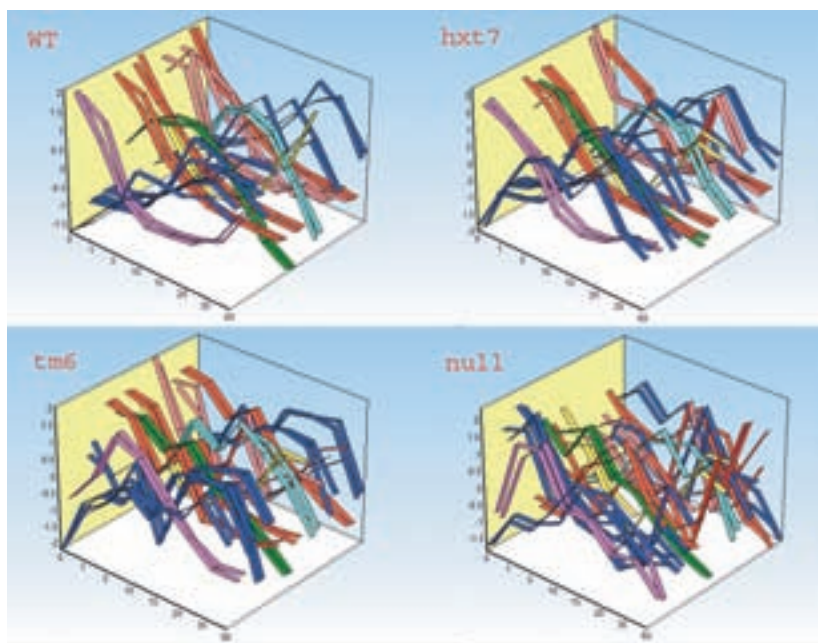
Real-time immune PCR

That RNA can be quantified with PCR by converting it first to cDNA in a reverse transcription reaction is familiar to most researchers, but less well known is that also proteins can be quantified. The technique is called immune qPCR and is offered by Chimera Biotech¹³. The design is similar to sandwich ELISA. A capture antibody is immobilised on the surface of the real-time PCR container. Test solution is added and the antigen is bound by the antibody. Detection antibody tagged with a DNA

molecule is then added. Careful washing removes unwanted material, leaving only the sandwiched antigen, which is quantified by PCR amplification of the DNA tag. As usual the signal is measured in real-time and the number of amplification cycles required to reach a threshold fluorescence level is registered. The main advantage of real-time PCR is that the signal increases exponentially with time, while conventional ELISA gives rise to linear signal increase (Figure 3). This gives much wider dynamic range, and for several assays higher sensitivity.

Expression profiling and multimarker diagnostics

While infectious and monogenic diseases usually can be assessed by measuring the expression of a single dominant marker, the majority of diseases have more complex pictures. A challenging goal of



Need help in understanding the market for new screening technologies?



HTStec is an independent market research consultancy, focused on providing informed opinion and market research on the technologies that underpin drug screening today. HTStec offers companies that are developing novel liquid handling, detection instruments, laboratory automation, assay reagents and platform technologies a range of consulting services and published market reports.

To find out how HTStec can help you maximize the market potential of your developments visit...

www.htstec.com



modern molecular methods is to find suitable markers for diagnosing, staging and monitoring complex diseases. It is not meaningful to base diagnostics on global expression, since most genes are not sensitive to the disease condition and will only contribute with confounding variance to any analysis. The challenge is to select reliable panels of markers that are sensitive to the disease state. Microarray screening can be a first step in such process, but the number of markers should be reduced and the test preferably converted to real-time PCR to improve dynamic range, enhance sensitivity and reduce costs. One such test is the Oncotype Dx from Genomic Health for the staging of breast cancer patients. Expression of 16 cancer related genes and five reference genes in formalin fixed paraffin embedded (FFPE) tissue is measured by real-time PCR and, using an empirically derived prospective mathematical algorithm, a recurrence score is calculated based on which a patient is assigned to one of three groups with estimated risk of low, intermediate and high distant recurrence¹⁴. A number of predictive panels for other types of cancers and other complex diseases are being developed, many of which are based on bodily fluids, such as sputum, urine, ejaculate, serum and plasma making the testing non/minimal invasive, which is a requisite for primary diagnostics. Epigenomics is developing tests for several cancers based on methylation markers and several promising studies on miRNA expression profiling have been published. Future tests will possibly be based on panels of both mRNA and miRNA markers, combined with tests for genetic and epigenetic factors.

Multiway expression profiling

Normal expression profiling is performed on a fairly large number of samples where each is characterised by measuring the expression of a fairly large number of genes, and the goal is to identify samples, or possibly genes, that have common expression patterns based on which they can be classified. The data make up a matrix with genes as columns and samples as rows that is analysed with multivariate methods such as Principal Component Analysis (PCA), Hierarchical clustering, the Self-Organized Map (SOM), Artificial Neural Networks (ANN) and Support Vector Machines (SVM). Recently it was shown that a third way can be introduced in the study to create a cube of data¹⁵. The model system was yeast, and the authors studied changes in the expression of metabolic genes over time after replacing ethanol as nutrient for glucose in four strains. The study

yielded a cube of data arranged along the three ways: genes, genotypes, and time. The three-way data were analysed using matrix augmentation followed by PCA, SOM and hierarchical clustering. Because of the very large amount of information in multiway studies, new correlations could be found. The authors classified four new genes based on functionality, and revealed several distinct relations between specific genotypes and gene function (Figure 4). This study has paved the ground for multiway designs, and we expect multiway approaches to be common in the future.

Laser assisted microdissection and QPCR tomography

Analysing tissue samples heterogeneity is often a problem. Tissues contain many different cell types that can have very different expression profiles, and tumours are made up of clones that can differ substantially. This makes it difficult to identify gene expression signatures characteristic for disease conditions. The problem can be approached with laser assisted microdissection as offered by Palm, Leica, Zeiss and MMI¹⁶. With these techniques homogeneous portions of the sample can be excised under a microscope and collected in tubes for analysis. The precision of the cut is 1µm and individual cells are readily excised for downstream real-time PCR analysis.

In qPCR tomography the studied sample is sliced along a direction of interest and expression is measured in the slices resulting in spatially resolved expression profiles. Recently the technique was applied to a single oocyte revealing intracellular expression profiles¹⁷ (Figure 5).

Digital PCR

Even though PCR is sensitive enough to detect a single molecular copy, its specificity is not always sufficient to discriminate between like products. We learnt above that with HRM it is possible to distinguish even molecules that differ in a single position. But this was in homozygous sample or heterozygous samples containing the two variants in equal amounts. Specificity can be somewhat

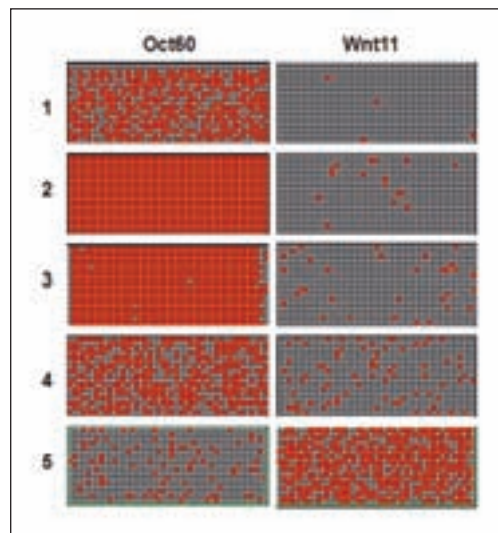


Figure 5: Intracellular expression profiles measured with digital PCR on the BIOMARK digital array. A single oocyte was sliced into five segments across the animal – vegetal, and the number of Oct60 (left) and Wnt11 (right) transcripts was measured. While Oct60 is most abundant in the 2nd and 3rd segments Wnt11 predominates in the segment that is closest to the vegetal pole

improved using probes that target one sequence variant and clamps that prevent hybridisation to the other. But sensitivity is still not sufficient to detect rare mutations, such as some somatic mutations, or mutations in shredded tumour cells in lymph or peripheral blood of cancer patients, or mutations in fetal cells present in peripheral blood of pregnant women. These targets are present in about one per 1,000-10,000 wild type sequences. An approach to detect rare mutations is digital PCR. In digital PCR the sample is diluted extensively and analysed in aliquots that each contains one or at most a few molecules. Most aliquots will not contain the mutated DNA, but in anyone that does, the background is tremendously reduced and the mutation is readily detected by conventional strategies. Another application of digital PCR is to count the number of target molecules in a sample. Diluting the sample even more extensively, such

References

- 1 Mullis, K et al. Specific enzymatic amplification of DNA in vitro: The polymerase chain reaction. *Cold Spring Harb. Symp. Quant. Biol.* 51: 263–273 (1986).
- 2 Higushi, R. Kinetic PCR Analysis: Real-time Monitoring of DNA Amplification Reactions. *Bio/Technology* 11, 1026 – 1030 (1993).
- 3 Livak, K et al. Oligonucleotides with fluorescent dyes at opposite ends provide a quenched probe system useful for detecting PCR product and nucleic acid hybridization. *PCR Methods Appl.* 4:357–362 (1995).
- 4 Tyagi, S and Kramer, FR. Molecular beacons: probes that fluoresce upon hybridization. *Nat Biotechnol* 14, 303 (1996).
- 5 Cardullo, RA et al. Detection of nucleic acid hybridization by nonradiative fluorescence resonance energy transfer. *Proc Natl Acad Sci U S A* 85, 8790 (1988).
- 6 Svanvik, N et al. Light-up probes: thiazole orange-conjugated peptide nucleic acid for detection of target nucleic acid in homogeneous solution. *Anal Biochem* 281, 26 (2000).
- 7 Whitcombe, D et al. Detection of PCR products using self-probing amplicons and fluorescence. *Nat Biotechnol* 17, 804 (1999).
- 8 Nazarenko, IA et al. A closed tube format for amplification and detection of DNA based on energy transfer. *Nucleic Acids Res* 25, 2516 (1997).
- 9 Wittwer, C et al. High-Resolution Genotyping by Amplicon Melting Analysis Using LCGreen. *Clinical Chemistry* 49, 853 (2003).

Continued on page 66

Figure 6: Single cell expression correlation. Table shows correlation between expressions of five genes in β cells from the pancreas. A number close to 1 reflects high positive correlation (both genes are either on or off), a number around zero reflects no correlation (the genes are expressed in different cells) and a number close to -1 is negative correlation (when one gene is on the other is off)

	ActB	Ins1	Ins2	Abcc8
Ins1	0.15			
Ins2	0.12	0.9		
Abcc8	-0.02	-0.01	0.06	
Kcnj11	0.11	-0.02	0.24	-0.15

Molecular Biology

Continued from page 65

10 Wojdacz, T and Dobrovic, A. Methylation-sensitive high resolution melting (MS-HRM): a new approach for sensitive and high-throughput assessment of methylation. *Nucleic Acids Research*, 2007, Vol. 35, No. 6 e41.

11 Espy, MJ et al. Real-Time PCR in Clinical Microbiology: Applications for Routine Laboratory Testing. *CLINICAL MICROBIOLOGY REVIEWS*, 19, 165 (2006).

12 Calin, G and Croce, C. MicroRNA signatures in human cancers. *Nature Reviews Cancer* 6, 857-866 (2006).

13 Lind, K and Kubista, M. Development and evaluation of three real-time immuno-PCR assemblages for quantification of PSA. *J. Immunological Methods* 304, 107-116 (2005).

14 Cronin, M et al. Measurement of gene expression in archival paraffin embedded tissues: Development and performance of a 92-gene reverse transcriptase-polymerase chain reaction assay. *Am J Pathol* 164, 35 (2004).

15 Ståhlberg, A et al. Multiway real-time PCR gene expression profiling in yeast *Saccharomyces cerevisiae* reveals altered transcriptional response of ADH-genes to glucose stimuli. *BMC Genomics* 9, 170 (2008).

16 Espina, V et al. Laser capture microdissection technology. *Expert Review of Molecular Diagnostics* 7, 647 (2007).

17 Sindelka, R et al. Intracellular expression profiles measured by real-time PCR tomography in the *Xenopus laevis* oocyte. *Nucleic Acids Research*. 1-6 (2007).

18 Bengtsson, M et al. Gene expression profiling in single cells from the pancreatic islets of Langerhans reveals lognormal distribution of mRNA levels. *Genome Research* 15, 1388 (2005).

19 Which mean do you mean? *Research Highlights, Nature Reviews Genetics* 6, 1 (2005).

that most aliquots contain none or only a single target, the total amount of target molecules in the sample can be calculated from the number of positive PCRs (Figure 5). The BIOMARK microfluidic digital array from Fluidigm automatically distributes loaded samples into 760 chambers for qPCR analysis and is a very convenient tool for digital PCR applications.

Single cell expression correlation

Above we concluded that tissues are heterogeneous because they contain several cell types. But there is substantial heterogeneity also among cells of the same type¹⁸. The variation in genes' expression among like cells is normally distributed when expressed in logarithmic scale, which is so called log normal distribution. The expression of the typical cell in a population is therefore given by the geometric mean of the cells' expression levels and not by the arithmetic mean that is measured on a sample containing many cells¹⁹. Measuring the expression of several genes per cell it was found that the levels of some mRNAs were highly correlated, while there was essentially no correlation between the levels of other mRNAs (Figure 6). This correlation suggests it is vital for the cell to produce the corresponding proteins at the same time and at a certain ratio, clearly indicating they have related functions. With the recent development of exponential preamplification methods such as the TaqMan PreAmp from Applied Biosystems and the Transplex Whole Transcriptome Amplification from Sigma, the transcripts can be amplified prior to qPCR analysis allowing for many more genes to be assayed per cell. At TATAA we are measuring expression of 50 genes per cell in a high throughput format using the BIOMARK system from Fluidigm. A signature of 50 selected genes per cell should be enough to characterise different cell types, identify new subtypes of cells, and study moderately complex expression pathways. This new expression correlation approach on single cell level is expected to become a very important tool in both basic research and target discovery.

Future challenges

Real-time PCR is becoming a mature technique. Instruments are very robust and reliable, and optimised assays can be ordered from several vendors. Even panels of optimised markers are available from companies such as SuperArray and TATAA Biocenter. Although the technique will be further refined and more niche products will be developed, main future improvements are not expected within the real-time PCR technology itself. Instead, future

challenges are to improve the down and upstream steps of molecular diagnostics. These are sample management, where efforts are needed to improve and standardise the sampling process, storage and transportation of biological samples, and to devise proper controls to normalise for variations in mRNA levels that occur post sampling. These variations include gene activation as well as repression, and RNA degradation. In Europe this is addressed through the European Union Framework Seven programme, which in autumn is launching the project 'Standardisation and improvement of generic pre-analytical tools and procedures for *in vitro* diagnostics' (SPIDIA). Upstream the challenge is to handle, annotate and analyse the very large amounts of data that will be generated. Roche is launching a 1536 well plate version of its LC480 instrument, the OpenArray platform from Biotrove runs 3072 reactions in parallel, and Fluidigm is launching a 96x96 microfluidic array, which in a single run measures 9216 reactions. Also Wafergen has announced it will launch a high throughput real-time PCR platform in 2009. In multimarker diagnostics a further challenge will be to combine and weigh information from different kinds of markers and optimally other information available about the patients. This challenge is addressed by MultiD Analyses which is developing dedicated software GenEx for real-time PCR data management and analysis and LabOnNet, which develops software for quality control of real-time PCR data.

DDW

Dr Mikael Kubista is founder of TATAA Biocenters (www.tataa.com), which is the leading provider of real-time PCR services in Europe and arranges basic as well as advanced hands-on training in real-time PCR and its applications worldwide. November 10-13 TATAA organised the main US Real-time PCR Symposium in San Francisco (www.qpcrsymposium.com). Dr Kubista also founded LightUp Technologies (www.lightup.se) as Europe's first company focusing on real-time PCR-based molecular diagnostic, and MultiD Analysis (www.multid.se), focusing on real-time PCR data analysis. Presently, Dr Kubista is heading the laboratory of gene expression at the recently founded Biotechnology Institute in Prague, Czech Republic (www.ibt.cas.cz), where he is setting up high throughput single cell expression profiling.