DNA diagnostics gets digitised

Quantitative real-time PCR (qPCR) has during the last two decades emerged as the preferred technology for nucleic acid analysis in routine as well as in research. qPCR has the sensitivity to detect a single molecule, the specificity to differentiate targets by a single nucleotide, and, because of its exponential nature, virtually unlimited dynamic range[1].

By Dr Mikael Kubista and Dr Anders Ståhlberg

With the recent appearance of high throughput platforms represented by the OpenArray from Life Technologies[2], the BioMark from Fluidigm[3], the LC1536 from Roche[4] and the SmartChip from Wafergen[5], can it get better? Well, classical qPCR is excellent platform technology and will dominate nucleic acid analyses years to come. But the exponential nature of qPCR, which is the key to its many advantages, is also limiting in some aspects. In particular, for example, multiplexing is challenging. True multiplexing based on one tube amplification and separate detection of targets is limited by competition for reagents and is in practice limited to two to four parallel reactions. In laboratories today most analyses performed are singleplex. When multiple targets are assayed the sample is divided into aliquots that are analysed separately. If sample amount is limited the nucleic acid may be pre-amplified. There are several methods to pre-amplify nucleic acids; some are linear while others are exponential, some are sequence specific while others are global, some target RNA while others target DNA. The methods are in general good, but they all require additional steps that add variation into the analysis, and they may also introduce bias. Another limitation of qPCR is precision. Replicate qPCR amplification curves for high quality assays and with some 25 or more initial template molecules appear usually very similar, and the standard deviation of the number of cycles required to reach threshold (Cq) is generally below 0.25 cycles and with some of the better instruments even below 0.1 cycle. From a technical perspective this reflects excellent instrument and assay performance (at a Cq of 25 cycles, SD of less than 0.25 cycles corresponds to a coefficient of variation below 1%). However, variation in Cq reflects imprecision in the logarithm of the concentration; imprecision in concentration (regular linear scale) scale is substantially larger. This has impact on, for example, copy number determinations using qPCR. Precision in copy number determinations depends on the quality of the standard curve (random error among replicates, number of standards, dynamic range, and position on the standard curve), but is rarely within 50% (strictly, the confidence interval is symmetric in logarithmic scale). Hence, under advantageous conditions the precision of qPCR is sufficient to detect a trisomy (distinguishing between two and three copies) with acceptable false negative and positive rates, but not sufficient to measure smaller differences in copy number. Third limitation of qPCR is limited specificity. True, a well-designed PCR assay amplifies with high preference a target with perfect sequence match for primers binding than sequences with even a single mismatch in a primer region, and

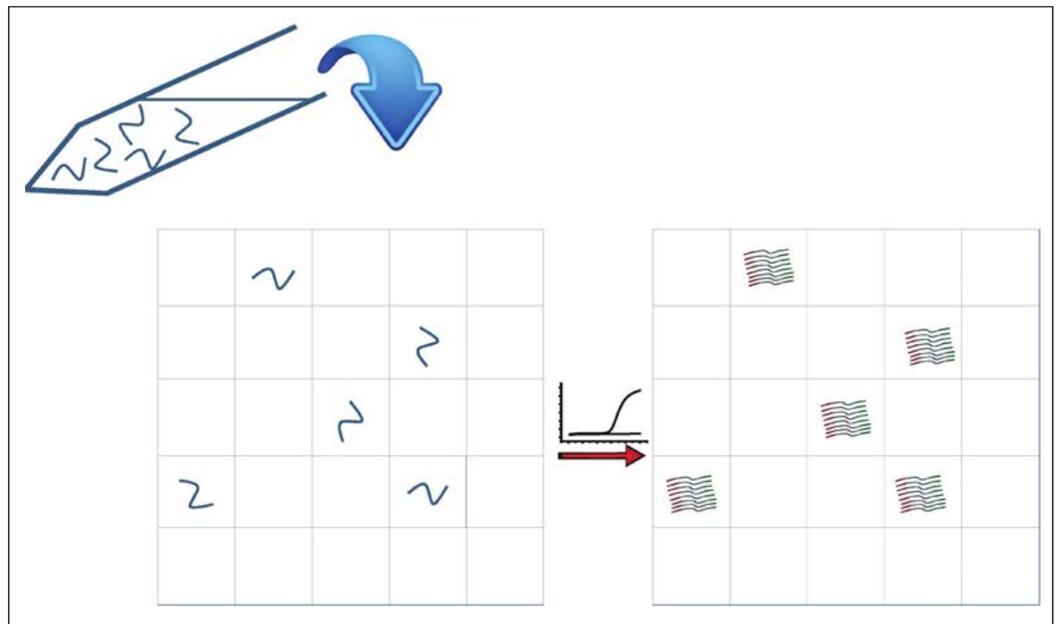


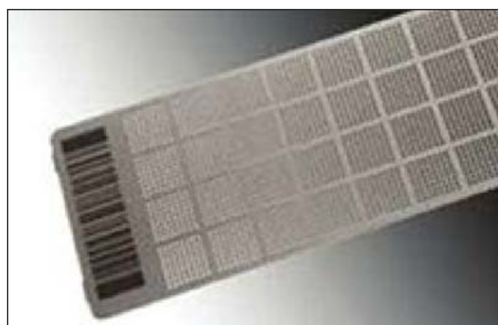
Figure 1
The principle of digital PCR. An extensively diluted sample is partitioned into a large number of reaction chambers, such that most chambers are empty or contain a single template molecule only. PCR is performed in the chambers amplifying the template molecules that are present. At the end of the PCR the number of positive reactions are scored, which reflects the initial number of template molecules that were present in the sample

qPCR is extensively used for mutation and SNP analysis. But there is finite possibility an assay with primers targeting an SNP will also amplify wild-type sequence. This gives rise to false positive signal and limits assay specificity. A SNP present in only a fraction of the template molecules may be missed. An alternative approach is to use generic primers and sense sequence variation using probes. The probe, however, binds with finite probability also to wild-type sequence, which limits assay specificity of this design. The specificity of a qPCR assay limits the background of wild-type sequence that is tolerated in a sample. Using regular Taqman assays already, 10-fold background of wild-type sequence is often challenging. qPCR assay specificity can be marginally increased by using modified primers/probes with elevated thermal stability[6]. Other strategies to enhance qPCR specificity is by sequestering, as used in CastPCR recently made available from Life Technologies[7], or by using modified primers such as the myT primers developed by Swift Biosciences[8]. Final limitation of qPCR is its sensitivity to inhibition. Analysing field sample substances from the sample matrix that have not been removed or carryover of reagents from upstream steps may interfere with the PCR influencing the measured Cqs and thus the estimate of target DNA concentration[9]. Wild-type Taq as well as several engineered variants are exceedingly sensitive to common inhibitors such as human blood, and major attempts are made to find more resistant variants using either rational design or selection strategies[10].

Already in 1992, even a year before Russ Higuchi described qPCR[11], Sykes et al[12] had the idea of quantifying target numbers by PCR using limiting dilution. Diluting a sample to such an extent that it contains a very small number of target molecules such that when aliquoted into reaction containers most will be empty, while some will contain a single template molecule only (**Figure 1**). Performing PCR the number of positive reactions will correspond to the number of template molecules in the original sample[13]. In 1999 Bert Vogelstein used the technique to quantify K-ras mutations in stool DNA from colorectal cancer patients and named it digital PCR (dPCR)[14]. It took time for dPCR to gain popularity as its precision was limited by the rather small number of reaction chambers in conventional 96 and 384-well plate instruments that often had to be filled manually. This changed when the high throughput platforms with integrated loading systems became available. Most convenient is the OpenArray from Life Technologies (**Figure 2**)[2] that we use in our dPCR services in Europe[15]. The measurement platform is a small metal plate the size of a microscope slide with 3072 reaction chambers in the form of small through holes that each hold 33nl of sample. They are arranged in 48 subarrays with 64 (= 8×8) chambers in each. This offers flexibility to tune the loading to the requested resolution and is also convenient for the serial dilutions initially performed to find optimum loading concentration. Other excellent platforms are the EP1 and the BioMark from Fluidigm[2]. These use Fluidigm's ingenious

integrated fluidic circuits that are able to run 12 samples, each partitioned into 765 reaction chambers (12.765) and 48 samples each partitioned into 770 reaction chambers (48.770). Reaction volumes are 6nl and 0.85nl, respectively.

Many of the limitations of qPCR are alleviated in dPCR and the technology has important advantages in several applications. Most popular dPCR application is copy number determination. In theory it should be possible to determine absolute copy numbers with dPCR, ie, loading a sample and measuring the number of target DNA copies it contains directly without having to use a reference. In practise, this is not so easy because of dead volume, losses during processing and handling of the sample, and ambiguity in separating positive and negative reads[16]. In the future some of these problems may be solved, perhaps by calibration, which would bring dPCR a step closer to becoming an absolute standard and as such exceedingly important as a reference method in nucleic acid testing. For now, copy number determinations, even with dPCR are performed as relative measurements. One may measure the relative abundance of one target in two samples or the relative amounts of two targets in a single sample. The latter is a typical comparison with an endogenous control, such as a conserved sequence present in exactly one copy per haploid genome. One such assay is ValidPrime[17], which targets a non-transcribed region in the genome and is therefore not amplifying any cDNA. One of the obvious applications is to determine allelic imbalance such as gene duplication/amplification, genetic aberrations, allelic loss and similar variations in gene copy numbers. The precision in the dPCR determination depends on the number of reaction chambers used, and has been detailed mathematically[18] and confirmed experimentally[19]. For example, four from five copies can be distinguished using some 1,200



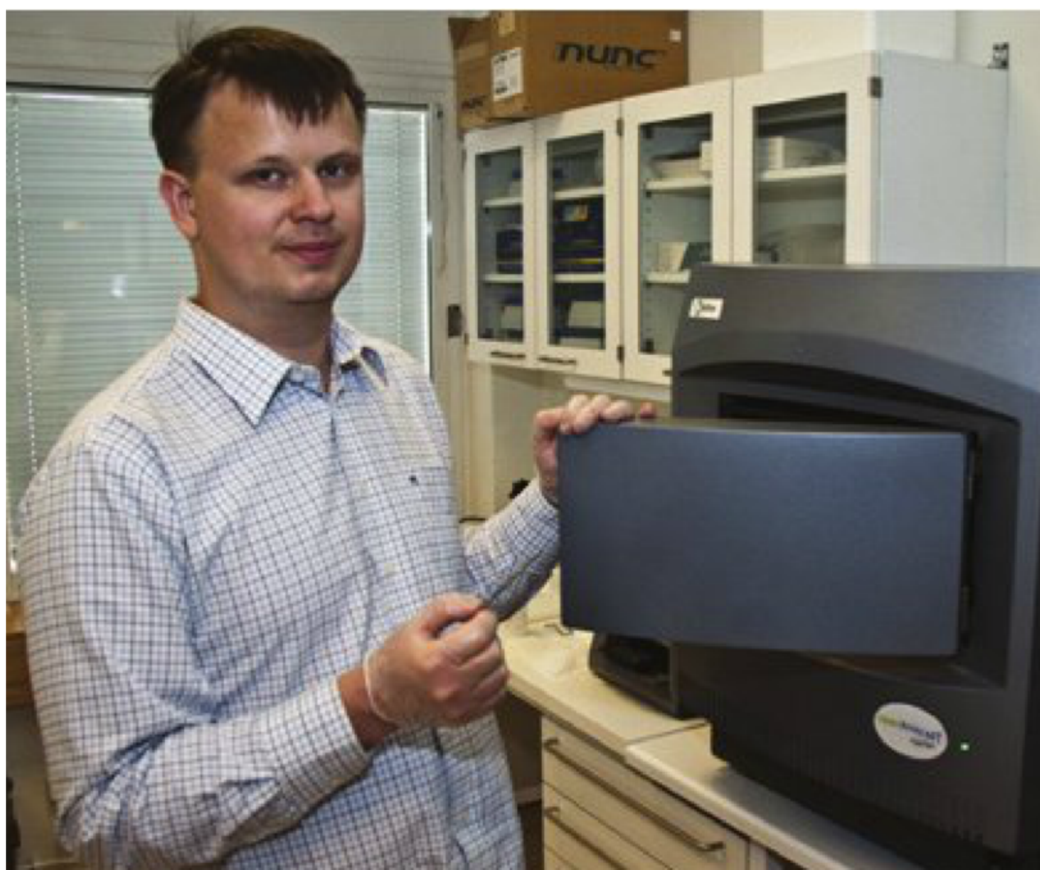


Figure 2
Digital PCR service in Europe. Left: Dr Anders Ståhlberg loading the OpenArray at the TATAA Biocenter digital PCR service facility in Gothenburg, Sweden. Above: The OpenArray through hole plate with 3072 reaction chambers

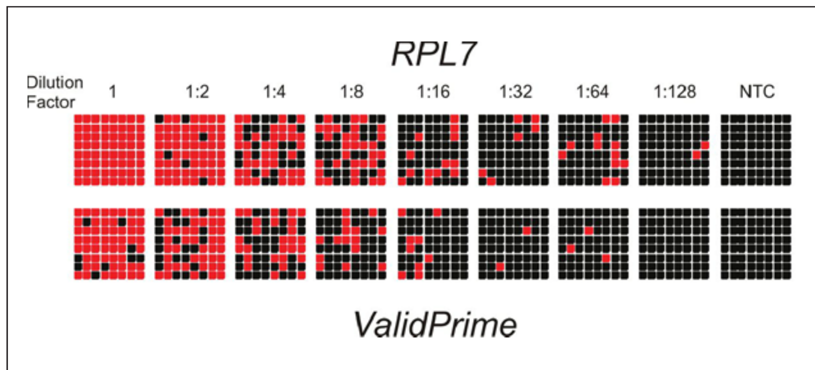


Figure 1
Probing sensitivity to genomic DNA of RT-PCR assays using ValidPrime. The intron spanning RPL7 assay and ValidPrime, which targets a non-transcribed sequence present in exactly one copy per haploid human genome, are used to amplify purified human genomic DNA in the dPCR OpenArray platform. A two-fold dilution series is performed across the 64-well subarrays. Comparison reveals that the RPL7 assay, despite being designed to span an intron, generates more PCR product from the genomic DNA than the ValidPrime

chambers, while with 8,000 chambers 10 from 11 copies can be separated. With qPCR it is challenging to separate two from three copies[20]. Recently Henrik Laurell pointed to a problem in RT-qPCR that can be approached with dPCR[21]. Analysing gene expression mRNA is reverse transcribed into cDNA and quantified using qPCR. Since transcripts are copies of the genome the RT-qPCR assay may also amplify genomic DNA that has not been removed during the processing of the sample. Current recommendation to deal with the problem is to design qPCR assays spanning introns. The intron-spanning primers will be close to each other in the transcript and readily amplified, while they are far apart on the genomic sequence and will not give rise to PCR product. True, at least for genes that have introns. But, Henrik Laurell says, many genes have intron-less pseudo genes that are amplified even when intron spanning primers are used. In fact, the number of intron-less pseudo genes of a gene in the genome may be large and highly variable among species. For example, the commonly used reference gene glyceraldehyde-3-phosphate dehydrogenase (GAPDH) has 62 pseudogenes in human and 331 in mouse[22]. These pseudogenes may contribute with substantial background. The sensitivity of RT-qPCR assays optimised for the quantification of cDNA to genomic DNA can be tested with dPCR. Dilution series of genomic DNA with an intron-spanning RFPL assay and ValidPrime (**Figure 3**) reveals the RFPL assay is picking up more DNA signal than ValidPrime, which amplifies one sequence copy per haploid genome. Obviously, the intron-spanning RFPL assay is picking up signal from intronless pseudogenes in the genome. Indeed, presence of pseudogenes is a major complication in RT-qPCR and for serious quantitative studies it is good practice to experimentally test the importance of genomic contribution to the RT-qPCR signal. It is also advisable to consult the excellent database of known human pseudogenes made available by Gerstein's lab at Yale University[23].

The sequence discrimination is enhanced in dPCR by the partitioning of the sample into the large number of reaction containers, which effective reduces the background level of related sequences. Consider a somatic mutation present in one out of 1,000 copies. With conventional probe-based qPCR we would not be able to detect the mutation because of insufficient discrimination. However, in dPCR the reaction chamber that happens to contain the mutated sequence will have substantially lower background of the wild-type sequence and the mutation is readily detected. This advantage was employed already in the pioneer paper by Vogelstein for early detection of a mutant ras oncogene in the stool of patients with colorectal cancer[10], and later to quantify EGFR alterations in lung cancer patients. The same approach has been to detect non-cultivatable pathogens against excessive backgrounds[25], and to probe individual environmental bacteria for viruses[26]. The improved discriminative ability of dPCR has also been employed for analysis of foetal DNA in plasma[27], and for sensitive quantification of minimal residual disease in chronic myeloid leukemia[28]. An esoteric dPCR/qPCR application is tomography to study the spatial distribution of transcripts in samples including gradients of transcripts within individual cells[29].

More extensive dilution, such that reaction chambers rarely contain more than a single template molecule, confers additional advantages to dPCR, including more extensive multiplexing. In conventional qPCR multiplexing is in practise limited to some four targets because of competition for reagents of the parallel reactions that compromises sensitivity and accuracy, and by crosstalk between detection channels when simultaneously quantifying multiple signals[1]. Clonal amplification obliterates these complications since there is no competition between reactions and there is no crosstalk. In fact, clonal amplification allows for an even higher degree of multiplexing using combinatorial detection based on multiple probes binding per target[30,31]. dPCR is potentially also less sensitive to inhibition. While inhibitors may delay amplification affecting the Cq values in qPCR leading to erroneous concentration estimates, they do not obliterate product formation and therefore do not influence the count in dPCR.

Clearly, digital PCR is a forthcoming technique that is rapidly becoming established in the field of nucleic acids analysis. Still, the wider spread of the

technique to routine applications is hampered by the rather high cost per sample analysed. This is about to change. Since digital PCR is an end-point technique that does not require continuous monitoring of product formation, the arena is open for other platforms. A spinning disc of inexpensive plastic with microfluidic architecture has been designed that passively compartmentalises a sample into 1,000 nanolitre-sized wells by centrifugation. A rapid air thermocycler is used for PCR and a CCD camera to acquire a fluorescent image to count positive reads[32]. The SlipChip has been described to perform digital PCR in a very simple and inexpensive format. Elongated wells in two plates are designed to overlap during sample loading. The fluidic path is broken up by slipping of the two plates that removes the overlap among wells and brings each well in contact with a reservoir preloaded with oil to generate 1,280 nanolitre reaction compartments. After thermal cycling end-point fluorescence intensity is measured[33]. A microfluidic 'megapixel' digital PCR device was recently presented that uses surface tension-based sample partitioning and dehydration control to enable single molecule amplification in 1 million reactors of picolitre volume. The device has a dynamic range of 10^7, can detect a single-nucleotide-variant in one copy per 100,000 wild-type sequences and discriminates 1% difference in chromosome copy number[34]. Out of the forthcoming next generation dPCR platforms, closest to market are droplet techniques based on water-oil emulsions. The aqueous droplets act as micro-reactors in the bulk oil phase. Biochemical reactions in emulsion systems date back to 1998[35], and the first report on single molecule PCR in droplets came almost 10 years ago[36]. The early emulsion-based systems either had poor amplification efficiency or required a complicated workflow. But major improvements during the last decade have led to the commercialisation of these platforms[37]. QuantaLife was first to launch with its Droplet Digital™ PCR system available at the end of last year[38]. The system is based on a droplet generator that partitions samples and master mix into nanolitre-sized droplets. DNA in the droplets is amplified by conventional thermal cycling whereafter a droplet reader (which is essentially a FACS) detects and measures the fluorescence of the droplets at a rate of 1,000 droplets per second. 32 samples can be analysed per hour. Raindance is expected to be next to launch[39], tentatively early in 2012[40]. Also Stokes Bio, which is part of Life Technologies, develops microfluidic solutions for quantitative PCR[41].

Future view

Digital PCR is here to stay and as awareness of the technology is increasing it will gain popularity. Commercialisation may be hampered by a complex IP landscape that may take some time to cross licence. In parallel, in-house development may be spurred by recent publication of 'Idiot-proof emulsion PCR' protocol[42]. Routine diagnostic applications that require IVD approval are still far ahead and will require the technology to mature. DDW

Dr Mikael Kubista *is head of the department of gene expression at the institute of Biotechnology of the Czech Academy of Sciences in Prague[43], and CEO and founder of the TATAA Biocenters (www.tataa.com). TATAA Biocenters are leading providers of qPCR services in Europe and have an OpenArray from Life Technologies as well as a BioMark from Fluidigm. In collaboration with Life Technologies, TATAA offers digital PCR services in Europe. TATAA has an intensive R&D programme related to qPCR and has developed several important products such as the dyes Chromofy and Visiblue, the 1-step extraction, RT, qPCR CelluLyser reagent, ValidPrime for RT-qPCR quality control, and proprietary panels for the identification of optimum reference genes, for profiling of embryonic stem cells, and for profiling of circulating tumor cells. TATAA also offers hands-on training courses in qPCR and molecular diagnostics world-wide (www.tataa.com/Courses/Courses.html) and arranges an annual qPCR symposium (www.qpcrsymposium.eu).*

Dr Anders Ståhlberg *is working as Researcher at the Department of Pathology, Sahlgrenska Cancer Center, University of Gothenburg, Sweden. His primary research interest is to understand molecular mechanisms in human sarcoma development and stem cell differentiation. He has developed several strategies for gene expression profiling, especially at single cell level.*

References

1 Kubista, M, Andrade, JM, Bengtsson, M, Forootan, A, Jonák, J, Lind, K, Sindelka, R, Sjöback, R, Sjögreen, B, Strömbom, L, Ståhlberg, A, Zoric, N. The real-time polymerase chain reaction. Mol Aspects Med. 2006 Apr-Jun;27(2-3):95-125. Epub 2006 Feb 3.
2 http://www.appliedbiosystems.com/absite/us/en/home/applications-technologies/real-time-pcr/real-time-pcr-instruments/openarray-real-time-pcr-system.html.
3 http://www.fluidigm.com/.
4 http://www.roche.com/products/product-details.htm?type=product&id=134.
5 http://www.wafergen.com/.
6 Letertre, C, Perelle, S, Dilasser, F, Arar, K, Fach, P. Evaluation of the performance of LNA and MGB probes in 5'-nuclease PCR assays. Mol Cell Probes. 2003 Dec;17(6):307-11.
7 http://www.lifetechnologies.com/news-gallery/press-releases/2011/oe-i-1-millio-o-loger-a-log-shot.html.
8 http://www.swiftbiosci.com/technology/my-t-primers.
9 Rådström, P, Knutsson, R, Wolffs, P, Lövenklev, M, Löfström, C. Pre-PCR processing: strategies to generate PCR-compatible samples. Mol Biotechnol. 2004 Feb;26(2):133-46.
10 Kermekchiev, MB, Kirilova, LI, Vail, EE, Barnes, WM. Mutants of Taq DNA polymerase resistant to PCR inhibitors allow DNA amplification from whole blood and crude soil samples. Nucleic Acids Res. 2009 Apr;37(5):e40. Epub 2009 Feb 10.
11 Higuchi, Russell, Fockler, Carita, Dollinger, Gavin and Watson, Robert. Kinetic PCR Analysis: Real-time Monitoring of DNA Amplification Reactions. Nature Biotechnology 11, 1026 – 1030 (1993). doi:10.1038/nbt0993-1026.
12 Sykes, PJ, Neoh, SH, Brisco, MJ, Hughes, E, Condon, J, Morley, AA. Quantitation of targets for PCR by use of limiting dilution. Biotechniques. 1992 Sep;13(3):444-9.
13 Kalinina, O, Lebedeva, I, Brown, J, Silver, J. Nanoliter scale PCR with TaqMan detection. Nucleic Acids Res. 1997 May 15;25(10):1999-2004.
14 Vogelstein, B, Kinzler, KW. Digital PCR. Proc Natl Acad Sci U S A. 1999 Aug 3;96(16):9236-41.
15 http://www.tataa.com/Services/Projects.html#Digital.
16 Sanders, R, Huggett, JF, Bushell, CA, Cowen, S, Scott, DJ, Foy, CA. Evaluation of digital PCR for absolute DNA quantification. Anal Chem. 2011 Sep 1;83(17):6474-84. Epub 2011 Aug 1.
17 http://www.tataa.com/products/ValidPrime-Control-of-Genomic-Background/View-all-products.html.
18 Dube, Simant, Qin, Jian, Ramakrishnan, Ramesh. Mathematical Analysis of Copy Number Variation in a DNA Sample Using Digital PCR on a Nanofluidic Device. PLoS ONE 3 (2008), pp. e2876–e2883.
19 Weaver, Suzanne, Dube, Simant, Mir, Alain, Qin, Jian Gang, Sun, Ramakrishnan, Ramesh, Jones, Robert C and Livak, Kenneth J. Taking qPCR to a higher level: Analysis of CNV reveals the power of high throughput qPCR to enhance quantitative resolution. Methods Volume 50, Issue 4, April 2010, Pages 271-276. The ongoing Evolution of qPCR.
20 http://www.tataa.com/Courses/Course-Modules.html#2daysbiostatistics.
21 Laurell, Henrik. ValidPrime for accurate real time qPCR quantification of RNA without DNase treatment. Presentation at the 2011 years qPCR Symposium in Prague (www.qpcrsymposium.eu).
22 Liu, YJ, Zheng, D, Balasubramanian, S, Carriero, N, Khurana, E, Robilotto, R, Gerstein, MB. Comprehensive analysis of the pseudogenes of glycolytic enzymes in vertebrates: the anomalously high number of GAPDH pseudogenes highlights a recent burst of retrotrans-positional activity. BMC Genomics. 2009 Oct 16;10:480.
23 http://www.pseudogene.org/.
24 Wang, J, Ramakrishnan, R, Tang, Z, Fan, W, Kluge, A, Dowlati, A, Jones, RC, Ma, PC. Quantifying EGFR alterations in the lung cancer genome with nanofluidic digital PCR arrays. Clin Chem. 2010 Apr;56(4):623-32. Epub 2010 Mar 5.
25 Ottesen, EA, Hong, JW, Quake, SR, Leadbetter, JR. Microfluidic digital PCR enables multigene analysis of individual environmental bacteria. Science. 2006 Dec 1;314(5804):1464-7.
26 Tadmor, AD, Ottesen, EA, Leadbetter, JR, Phillips, R. Probing individual environmental bacteria for viruses by using microfluidic digital PCR. Science. 2011 Jul 1;333(6038):58-62.
27 Lun, Fiona MF, Chiu, Rossa, WK. Chan, KC Allen, Leung, Tak Yeung, Lau, Tze Kin and Lo, YM Dennis. Microfluidics Digital PCR Reveals a Higher than Expected Fraction of Fetal DNA in Maternal Plasma. Clinical Chemistry. 2008;54:1664-1672.
28 Goh, HG, Lin, M, Fukushima, T, Saglio, G, Kim, D, Choi, SY, Kim, SH, Lee, J, Lee, YS, Oh, SM, Kim, DW. Sensitive quantitation of minimal residual disease in chronic myeloid leukemia using nanofluidic digital polymerase chain reaction assay. Leuk Lymphoma. 2011 May;52(5):896-904. Epub 2011 Feb 21.
29 Sindelka, R, Sidova, M, Svec, D, Kubista, M. Spatial expression profiles in the Xenopus laevis oocytes measured with qPCR tomography. Methods. 2010 Jan 4. doi:10.1016/j.ymeth.2009.12.011.
30 Huang, Q, Zheng, L, Zhu, Y, Zhang, J, Wen, H, Huang, J, Niu, J, Zhao, X, Li, Q. Multicolor combinatorial probe coding for real-time PCR. PLoS One. 2011 Jan 14;6(1):e16033.
31 Zhong, Q, Bhattacharya, S, Kotsopoulos, S, Olson, J, Taly, V, Griffiths, AD, Link, DR, Larson, JW. Multiplex digital PCR: breaking the one target per color barrier of quantitative PCR. Lab Chip. 2011 Jul 7;11(13):2167-74. Epub 2011 May 17.
32 Sundberg, SO, Wittwer, CT, Gao, C, Gale, BK. Spinning disk platform for microfluidic digital polymerase chain reaction. Anal Chem. 2010 Feb 15;82(4):1546-50.
33 Shen, F, Du, W, Kreutz, JE, Fok, A, Ismagilov, RF. Digital PCR on a SlipChip. Lab Chip. 2010 Oct 21;10(20):2666-72. Epub 2010 Jul 1.
34 Heyries, Kevin A, Tropini, Carolina, VanInsberghe, Michael, Doolin, Callum, Petriv, Oleh I, Singhal, Anupam, Leung, Kaston, Hughesman, Curtis B and Hansen, Carl L. Megapixel digital PCR. Nature methods VOL.8 NO.8, AUGUST 2011, 649.
35 Tawfik, DS, Griffiths, AD. Man-made cell-like compartments for molecular evolution. Nat Biotechnol. 1998 Jul;16(7):652-6.
36 Nakano, Michihiko, Komatsu, Jun, Matsuura, Shun-ichi, Takashima, Kazunori, Katsura, Shinji, Mizuno, Akira. Single-molecule PCR using water-in-oil emulsion. Journal of Biotechnology, Volume 102, Issue 2, 24 April 2003, Pages 117-124.
37 Kiss, MM, Ortoleva-Donnelly, L, Beer, NR, Warner, J, Bailey, CG, Colston, BW, Rothberg, JM, Link, DR, Leamon, JH. High-throughput quantitative polymerase chain reaction in picoliter droplets. Anal Chem. 2008 Dec 1;80(23):8975-81.
38 http://www.quantalife.com.
39 http://www.raindancetech.com.
40 http://www.genomeweb.com/pcrsample-prep/raindance-eyes-early-2012-digital-pcr-launch-preps-new-sequence-enrichment-platf.
41 http://www.stokesbio.com/.
42 Schütze, Tatjana and Glökler, Jörn. Idiot-proof emulsion PCR. LabTimes. 2010; 2010(1):50.
43 http://www.img.cas.cz/ge/.