

Design and Validation of a qPCR Panel for Stem Cell Differentiation

Bachelor of Science Thesis
Project performed at TATAA Biocenter, Gothenburg

LINN LINDBLÖM

Design and Validation of a qPCR Panel for Stem Cell Differentiation

LINN LINDBLOM

SUPERVISOR: ROBERT SJÖBACK

EXAMINER: JOAKIM NORBECK



Department of Biology and Biological Engineering
Division of Systems and Synthetic Biology
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, 2016

Design and Validation of a qPCR Panel for Stem Cell Differentiation

Linn Lindblom

© LINN LINDBLÖM, 2016

Department of Biology and Biological Engineering
Division of Systems and Synthetic Biology
Chalmers University of Technology
SE-412 96 Gothenburg
Sweden

Cover: qPCR analysis with SYBR Green for assay CD44

Tryckeri: Biologi och Bioteknik, Chalmers
Gothenburg, Sweden 2016

ACKNOWLEDGEMENTS

This Bachelor of Science Thesis is the final part of the chemical engineering program at Chalmers University of Technology and corresponds to 15 credits. The project was performed at the company TATAA Biocenter in Gothenburg.

I want to thank TATAA Biocenter and its staff for giving me the opportunity to accomplish my thesis at the company. A special thanks to my supervisor Robert Sjöback for his involvement and support during the project. Furthermore, I want to thank Jan Nygren for all help with the laboratory work and Julie Elling for the help with the references. I received guidance from Emelie Lott and Caroline Olsson with design and the Fragment Analyzer and for that I am very grateful. I also want to thank Biosearch Technologies for the sponsorship of the probes.

Gothenburg, May 2016

Linn Lindblom

Design and Validation of a qPCR Panel for Stem Cell Differentiation

LINN LINDBLÖM

Department of Biology and Biological Engineering
Division of Systems and Synthetic Biology
Chalmers University of Technology

ABSTRACT

The purpose of this project was to develop and validate a qPCR panel for stem cells differentiation for human cells. The panel consists of 62 selected genes, which are marker for cell differentiation for neurons, cardiac cells, kidney cells, embryonic cells and detection of pluripotency. The validation of the panel implied design of assays, qPCR analysis with SYBR Green and probe, control of PCR product length with a capillary electrophoresis instrument and a preamplification. The assays were designed between an exon-exon junction or separated with large introns as far as possible. Some assays are designed over an exon, in order to detect as many protein coding transcript variants as possible. The selected genes have been analyzed successfully with SYBR Green and probe.

A validation of a preamplification is performed to enable analysis of limited sample volumes. The preamplification implied a mixed pool consisted of all primers and reference genes. The solution was analyzed with a regular PCR analysis with SYBR Green. Products from the regular PCR was used as template in the final qPCR analysis. The results from the preamplification were good. Efficiency for the assays were $\geq 90\%$ for almost all assays and they were reproducible. So far, the panel is only validated for preamplification with SYBR Green. In the near future, the panel also will be validated with probe.

The result of this project is a developed qPCR panel for stem cells differentiation for human cells. The qPCR based panel can now be used of scientists to detect the differentiation state of added stem cells.

Key words: qPCR, stem cells, stem cells differentiation

Design och validering av en qPCR baserad panel för stamcelldifferentiering

LINN LINDBLÖM

Institutionen för Biologi och Bioteknik
Avdelningen för Systembiologi
Chalmers Tekniska Högskola

SAMMANFATTNING

Syftet med detta projekt var att utveckla och validera en qPCR panel för stamcelldifferentiering för humana celler. Panelen består av 62 utvalda gener som är markörer för celldifferentiering till bland annat nervceller, hjärtceller, njurceller, embryonala celler och detektion av pluripotens. Valideringen av panelen har innefattat design av assayer, qPCR analys med SYBR Green och probe, kontroll av PCR-produktens längd med hjälp av ett kapillärelektroforetiskt instrument och avslutningsvis en preamplifiering. Assayerna designades i möjligaste mån mellan en exon-exon korsning eller separerade med stora intron. Detta kunde inte uppfyllas för alla assayer, därför designades en del assayer över ett exon för att detektera så många proteinkodade transkriptvarianter som möjligt. De utvalda generna har analyserats med gott resultat med SYBR Green och probe.

En validering av en preamplifiering är gjord för att möjliggöra analys av begränsade provvolymmer. Preamplifieringen innebar en pool bestående av alla primers från panel och referensgener. Lösningen analyserades med PCR med SYBR Green. Produkter från PCR analysen användes sedan som templat i den avslutande qPCR analysen. Resultatet från preamplifieringen var bra. Effektiviteten för assayerna var $\geq 90\%$ för nästan alla assayer och de var även reproducerbara. Än så länge är panelen endast validerad för preamplifiering med SYBR Green. Inom en snar framtid kommer panelen även att valideras för användning av probe.

Resultatet av projektet är en utvecklad qPCR panel för stamcelldifferentiering. Panelen kan nu användas av forskare för att detektera differentieringstillstånd av tillsatta stamceller.

Sökord: qPCR, stamceller, stamcelldifferentiering

Table of Content

ABBREVIATIONS	1
1. INTRODUCTION	2
1.1 Background	2
1.2 Purpose	2
2. THEORY	3
2.1 Real-Time PCR, qPCR	3
2.2 Stem Cells and Stem Cells Differentiation	5
2.3 Recent Research	6
3. METHOD	7
3.1 Primer and Probe Design	7
3.2 Real-Time PCR Method	7
3.3 Preamplification	8
4. EXPERIMENTAL	9
4.1 Materials	9
4.1.1 Reagents and Chemicals	9
4.1.2 Apparatus	9
4.1.3 Other Equipment	9
4.2 Procedures	9
4.2.1 Preparations	9
4.2.2 qPCR Analysis with SYBR Green	10
4.2.3 Verification of Products with Fragment Analyzer	11
4.2.4 qPCR Analysis with Probe	11
4.2.5 Preamplification with SYBR Green and Probe	12
5. RESULTS AND DISCUSSION	14
5.1 Design of Selected Genes	14
5.2 qPCR Analysis with SYBR Green and probe	14
5.3 Verification of the Product Length	19
5.4 Preamplification	20
6. CONCLUSION	23
REFERENCES	24
APPEENDIX	I

Abbreviations

PCR – Polymerase chain reaction

qPCR – Real-time polymerase chain reaction

DNA – Deoxyribonucleic acid

NTC – No template control

SNP – Single-nucleotide polymorphism

1. INTRODUCTION

1.1 Background

Polymerase chain reaction, PCR, is a common technique used to make replicates of a specific DNA sequence. (Britannica Online) DNA can be amplified from one single molecule to a large number of molecules. (Kubista 2010) Therefore, the technique gives opportunities for researchers to receive sufficient amount of DNA for microbiology experiments. Common applications for PCR are forensic analysis, evolutionary biology and medical diagnostics. (Britannica Online) The PCR-technique is also used as an alternative to gene cloning, for further analysis by DNA sequencing. (Dictionary of Science)

PCR was developed in 1983 by Kary Mullis and since then the method has been an important qualitative analytical tool. Russell Higuchi made an important discovery in the beginning of the 90s. He discovered that PCR could occur in presence of nucleic acid, which becomes fluorescent when it binds into DNA. This observation provided an opportunity to measure the fluorescent during the reaction. It implicated that the amplification of DNA could be measured in real time. (Kubista 2010) The new method was called Real-time Quantitative PCR, qPCR.

The qPCR method can be used for various application since the method is very flexible. (Noaksson et al.) qPCR combines the objectivity of fluorescence detection with the original PCR reaction and have become a very important application for detection of bacterial pathogens, SNP analyses and solid tumor diagnostic. There is a variety of uses of qPCR and the technique had evolved from general detection of cells to identifications of specific genes. (Bustin 2004)

TATAA Biocenter is a company, which provides nucleic acid services by qPCR. The company offers the entire field of qPCR services, which includes commissioned services and research projects, hands-on training and product development. (TATAA Biocenter) The developed products consist of reagent for qPCR and molecular tests which in turn include panels of qPCR assays. TATAA Biocenter have a patent for a panel of qPCR assays for stem cells differentiation together with another company. This panel only consists of a few genes for stem cells differentiation. The research is constantly evolving and scientists have discovered more genes for detection of stem cells differentiation. Therefore, a new qPCR panel will be developed with more interesting genes for stem cells differentiation.

1.2 Purpose

The purpose of these experiments is to develop a qPCR panel for stem cells differentiation for human cells. The developed panel will be used to analyze the differentiation state of stem cells. Furthermore, this will be a good method for scientist to control result from addition of stem cells. A validation for preamplification will be performed to enable analysis at a High Throughput instrument with small sample volumes. The panel will consist of 48 genes selected for stem cells differentiation excluding 12 reference genes and ValidPrime.

2. THEORY

2.1 Real-Time PCR, qPCR

Real time PCR is a scientific technology to make replicates of a specific sequence. The method is based on a quantitative relationship between the amount of target nucleic acid present in the start of the assay and the amount of product amplified during the exponential phase. PCR takes place in a number of cycles and the first cycle is called threshold cycle (C_t). This cycle is the point where the instrument first detects fluorescence, see *Figure 1*. (Bustin 2004) The threshold cycle (C_t) is indirectly proportional to the number of copies of the original template. The relation between the original copy number and threshold cycle (C_t) can be determined with a known standard. Through this method the number of copies in the original sample can be determined. (Tymoczko et al. 2013)

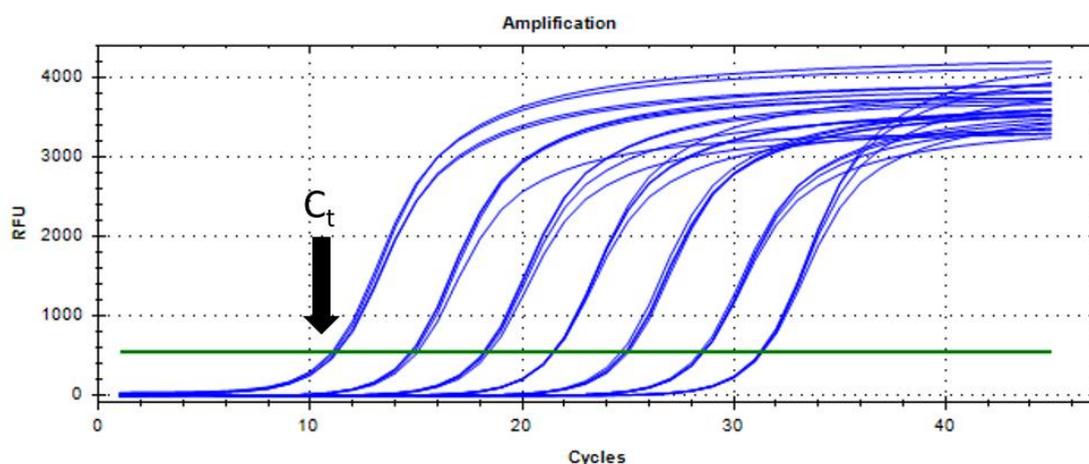


Figure 1. The first cycle in a PCR reaction, which calls threshold cycle (C_t)

The threshold cycle depends on initial copy number, efficiency of PCR amplification, efficiency of cleavage or hybridization of the probe and sensitivity of detection. It requires fewer cycles to reach detection of fluorescence with greater initial copy numbers. (Bustin 2004) A PCR cycle consists of three steps; denaturation, annealing and extension. During the first step, denaturation, an excess of primer pairs is added and the solution is heated to 95°C in 5 seconds to enable separation of the DNA strand. (Tymoczko et al. 2013) There are two different types of primers, forward primer and reverse primer. The primer pair binds to the template at their complementary strands. (Britannica Online) The length of the primers varies between 20 – 30 nucleotides. The next step is annealing, which intends to cool the solution to 60°C in 30 seconds to allow each primer to hybridize to a DNA strand. Last step in a cycle is extension. The solution is heated again to 72°C in 10 seconds. The temperature is the optimal temperature for Taq polymerase. The function of the Taq polymerase is to elongate the primers in the direction of target sequence. Direction for DNA synthesis is from the 5' end - to-3' end. After the first cycle there are no more reagents added. (Tymoczko et al. 2013) The reagents for a PCR reaction consists of a master mix and a template. The master mix includes primer pairs, master mix with Taq polymerase and water.

PCR reactions can be performed with either SYBR Green or probes. SYBR Green fluoresce when it binds to double-stranded DNA. The SYBR Green is released when the DNA is denatured, which results in reduced fluorescence. All new copy of double-stranded DNA binds to SYBR Green and this results in an increase in fluorescence intensity proportioned to

the amount of PCR product. Since the SYBR Green binds to all double-stranded DNA, there will be a risk to bind to nonspecific sequences.

There are multiple probe systems, which are suited for various applications. One of these systems is Taqman Probes. A Taqman probe consists of a fluorescent reporter dye and a quencher dye, in each end of the probe. When the probe is intact the quencher avoids fluorescence but during the extension cycle. The DNA polymerase cleaves the reporter dye from the probe. This generates the reporter dye to fluoresce. (ThermoFisher) Taqman probes are linear probes and can be used for both sensitivity and specificity analyzes. The linear probes are often simpler to design and manufacture compared to the probes with secondary structure. Molecular Beacons are an example of a probe with secondary structure. (Bustin 2004)

A melt curve analysis can be used to identify nonspecific amplification products. One example of a nonspecific products is primer dimer. (Nolan & Bustin 2013) A melt curve is only executed with SYBR Green. To perform a melt curve analysis, the PCR products must be heated to 95°C to ensure a complete separation of the products. Afterwards, the products must be cooled to 55°C to ensure complete hybridization. Then the temperature increases gradually according to a fixed interval to 95°C. A distinct result of the melting profile is obtained by longer hold with smaller temperature steps. A melt curve with a single melting profile indicates a single product, see *Figure 2*. Multiple melting profiles indicates multiple products and the primers must be redesigned. (Bustin 2004)

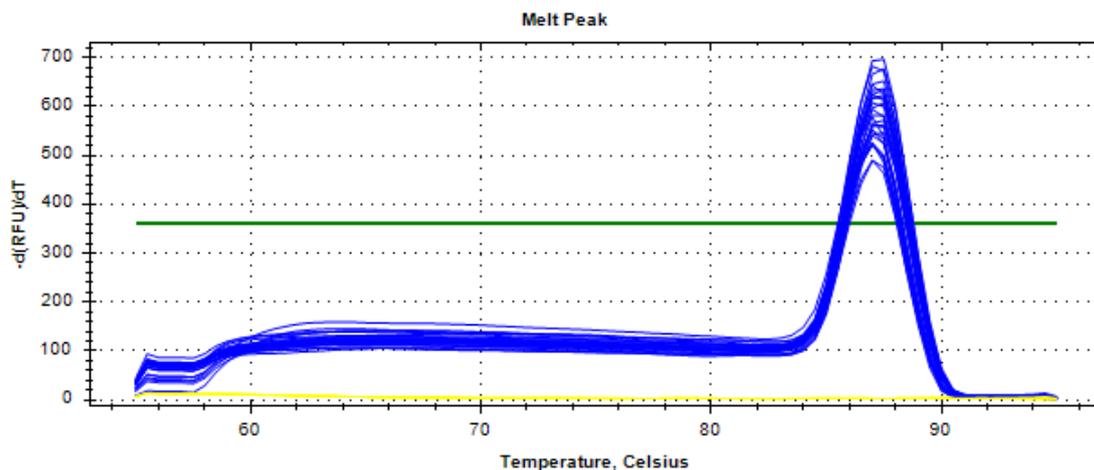


Figure 2. Melt curve with a single melting profile

When real samples will be analyzed, there can be limited sample volumes. This problem can be avoided by a preamplification. Samples are usually preamplified in order to be compatible with Real-time PCR constraints. (Andersson et al. 2015) The purpose of preamplification is to multiply the number of copies to enable analysis with singleplex PCR. (Ståhlberg & Kubista 2014) Preamplification can be global, targeting all molecules or a specific sequence. The PCR method intends a pool of specific primer pairs, which is performed in a limited number of cycles to avoid introducing bias. Concentration of the primers is 10 – 20 times lower in preamplification PCR compared to a standard PCR. The concentration is lower to reduce the risk of formation of nonspecific products. The annealing is extended to compensate for the low concentration of primers. (Andersson et al. 2015) A successful preamplification has efficiency over 90 % and the assays have high reproducibility. A low standard deviation is also important for a successful preamplification. (Ståhlberg & Kubista 2014)

There are different types of qPCR instruments, digital PCR and High-throughput instruments are two examples. Digital PCR generates multiple PCR reactions, some of these reactions consist of the target molecule and others are blank. After the PCR amplification, the number of target molecules in the initial sample corresponds to positive reactions. The blank samples are negative. Another type of qPCR instrument is a High-Throughput instrument. These instruments are used for applications that are not practical and cost-efficient on the original qPCR instruments. The system is based on chip, where 96 assays are loaded on one side and 96 samples on the other side. The assays and samples are mixed into $96 \times 96 = 9.216$ reaction chambers for analysis. High-throughput instruments can be used for multi-marker diagnostics and prognostics. (Kubista 2010)

Real-time PCR can be used for many technical applications. The method has become important for genetic diagnostic since DNA mutations can be identified even with small sample aliquots. qPCR can also be used to detect infectious agent in blood and tissues. (Nationalencyklopedin, Genteknik) HIV is an example of an infectious agents in blood, which can be detected by PCR. (Lakshmi et al. 2011) Analysis of the differentiation state of stem cells can also be analyzed with qPCR. (Noaksson et al.)

2.2 Stem Cells and Stem Cells Differentiation

Stem cells are immature cells and by cell division may generate more differentiated cells or more stem cells. There are two different types of stem cells, embryonic stem cells and adult stem cells. The embryonic stem cells are also called totipotent and these cells manage the embryonic development. The further growth of the embryo results in differentiated stem cells. Therefore, no cells will retain the totipotency. The function of the adult stem cells is to synthesize new cells with specific properties. These stem cells are called pluripotent stem cells. The reason to form new stem cells depends on cells with limited lifetime or lost after injury. The pluripotent stem cells can differentiate into all types of cells in living organisms, for example muscle cells, neurons and cardiac cells. (Nationalencyklopedin, Stamcell) Since the pluripotent stem cells can differentiate to all cell types in a human body gives these stem cells a significant potential for tissue engineering, regenerate medicines and diseases modelling. (Tsankov et al. 2015)

2.3 Recent Research

There is more upcoming research in the qPCR field and previously it has been developed qPCR panels for stem cells differentiation. One method to detect the differentiation of human embryonic stem cells has been developed by using qPCR. The genes have been selected through observation of morphological changes with light microscopy and immunohistochemistry. The sensitivity of the method has been increased by measure the expression of up- and downregulated genes, before and after the differentiation. The results of the analysis proved the differentiation state of the cells by calculating the expression based on mRNA levels of four different proteins. (Noaksson et al.)

Furthermore, there is another study of stem cells with Real-time PCR. The purpose of the study was to clarify the molecular status of single- and colonies of human adult germ stem cells. These cells were compared with human embryonic stem cells and human fibroblasts. The pluripotency of these cells were analyzed with Real-time PCR and then compered. The expression of germ cell- and pluripotency-associated genes were compared in the various stem cells. For instance, the selected pluripotency genes were EPCAM, PROM1 and SALL2. The result of the study showed marked difference between the human adult germ stem cells and the human fibroblast in terms of the expression of germ- and pluripotency-associated genes. The human adult and embryonic stem cells have only a few similarities with the human fibroblasts. (Conrad et al. 2016)

3. METHOD

The panel consists of 62 genes for stem cell differentiation. The genes have been selected from literature studies focused on qPCR-panels for stem cell differentiation. (Conrad et al. 2016) Studies of commercial panels have also been made to selected coveted genes. (R&D SYSTEMS) The selected genes also consist of received requests of genes from external partners, which have expertise in stem cell research. The panel will focus on stem cells differentiation for detection of kidney cells, pluripotency, cardiac cells, smooth muscle cells, endothelial cells, mesenchymal stem cells and hepatic tissue.

3.1 Primer and Probe Design

Design of primers and probes is an important part of the validation of a qPCR panel. The primers must be specific for the gene to be as sensitive as possible. The sensitivity should preferably be sensitive enough that it could be analyzed at single-cell level. A commercial design program, Primer-BLAST, has been used to design the primers. The primers have been designed between an exon-exon junction or separated with an intron with a length of 1000-3000 base pairs. All transcript variants which are protein coding, have been including as far as possible. The length of the primers is between 20 – 30 base pairs long and the length of the amplicon has been designed between 80 – 200 base pairs. The settings for the primers melt temperature has been set at 57 – 63°C, with a max difference with +/- 2°C. A synthetic template has been designed by dint of the forward and reverse primer. The synthetic template, also called gBlocks, has been used to make a standard curve for each assay. The program Beacon designer has been used for probe design. The primer design and the amplicon size have great significance for the probe design. A short amplicon can make it difficult to design a good probe since it is few bases to bind in to. The probe was designed with a melt temperature at 67°C.

3.2 Real-Time PCR Method

The validation of the real-time PCR method consists of three steps, where every step must be approved. The first step intends a qPCR analysis with the dye SYBR Green with a following melt curve analysis. The used qPCR instrument was BIO-RAD CFX384. Analysis with SYBR Green included a master mix with the interesting gene, cDNA, gDNA and a dilution series of gBlocks. cDNA was used to verify the assay, cDNA and the assays must have the same melting profile in the melt curve analysis. The purpose of the gDNA analysis was to control if the assay detects genomic DNA. The melting profile for gDNA should preferably not be the same as the assay and cDNA. A dilution series of gBlocks was used to make a standard curve. The results must pass the criteria to proceed to the next step in the validation plan. The criteria were an efficiency between 80 – 110 %, r^2 -values over 0.995, a linear standard curve and a melt curve without nonspecific products. There was also important with a great dilution series, without a wide dispersion.

The length of the PCR products was controlled with the instrument Fragment Analyzer, which is a multiplexed capillary electrophoresis instrument for performing separation and quantification of double stranded DNA. The criteria for this experiment were one lower and one upper marker, the difference between the practical and theoretical length of the product must be ± 10 base pairs and the same length of the assay and cDNA. The gDNA sample could have several peaks.

If these criteria were passed, the assays were analyzed with probe. The used probes were Taqman probes with modification of the 5' and 3' end. In 5' end the dye FAM and in 3' end a Black Hole Quencher. Analysis with probe occurred without a melting analysis and included master mix with the interesting gene and a dilution series of gBlocks. The criteria for the probe analysis were an efficiency between 80 – 110 %, r^2 -values over 0.995 and a linear standard curve. There was also important with a great dilution series, without a wide dispersion.

3.3 Preamplification

Last step in the validation was a preamplification and the procedure occurred in two parts. First, a preamplification with regular PCR analysis and then a qPCR analysis of the amplified products from the preamplification reactions. The preamplification procedure was performed with SYBR Green and probe. The regular PCR analysis consisted of a pool of all primer pairs including reference genes and Validprime. The purpose to have the reference genes in the pool was for normalization. Some of the reference genes are normally expressed in samples and are therefore used to adjust the background. The solution of primers and reference genes was mixed with cDNA and spiked gBlocks or distilled water. For low expressed assays with high Cq-values for cDNA, ≥ 35 cycles, the cDNA was spiked with gBlocks.

The last part of the preamplification was a qPCR analysis of the amplified products, and this part differ between SYBR Green and probe. The difference is the content of the master mix. The content of the master mix for SYBR Green consisted of primers and the probe master mix consisted of both primers and probes. There were also criteria to pass for the preamplification step. The general criterion was the efficiency, which should be ≥ 80 %. Reproducibility and standard deviation were two other important aspects.

4. EXPERIMENTAL

4.1 Materials

4.1.1 Reagents and Chemicals

Primers from IDT, Integrated DNA Technologies
gBlocks from IDT, Integrated DNA Technologies
Probe from BIOSEARCH Technologies
cDNA Human
gDNA Human
TATAA SYBR® Grandmaster® Mix from TATAA Biocenter
TATAA Probe Grandmaster® Mix from TATAA Biocenter
GenElute™-LPA from Sigma Aldrich
Tris-EDTA buffer solution from Sigma Aldrich
Distilled water, DNase/RNase free from Life Technologies
Reagent kit for Fragment Analyzer

4.1.2 Apparatus

Pipetting robot: Eppendorf epMotion 5070
qPCR instrument: BIO-RAD CFX384/96 Real Times System
Centrifuge: Eppendorf Centrifuge 5804R
Capillary electrophoresis instrument: Fragment Analyzer from Applied Biosystems

4.1.3 Other Equipment

Pipettes: Eppendorf Research
Plates: Framestar® 480/384 from 4titude
Strip: Molecular Biology from Thermo Scientific

4.2 Procedures

4.2.1 Preparations of Reagent

The primers were diluted before use and the concentration of stock solution was 100 μM . Consequently, a working solution was prepared with a concentration of 10 μM . The working solution consisted of 10 μL of forward primer and 10 μL of the reverse primer and was diluted with 80 μL of TE(1x). The gBlocks were diluted in two steps. The first step intended to add distilled water to the pellets, the concentration of stock solution was $1 \cdot 10^{10}$ molecules/ μL . The stock solution was heated in a hot cabinet of 50°C in 20 minutes to dissolve the pellets and was then diluted to a working solution with a concentration of $1 \cdot 10^8$ molecules/ μL . The working solution consisted of 2 μL of stock solution and 198 μL of distilled water. cDNA and gDNA were also diluted before use. cDNA was diluted 10x with distilled water, 10 μL of cDNA and 90 μL of distilled water. gDNA was diluted 20x with distilled water, 5 μL of gDNA and 95 μL of distilled water. A TE-LPA buffer, which was used to the dilution series, was prepared. The TE-LPA-buffer contained 3957 μL of distilled water, 40 μL of TE(100x) and 3.2 μL of LPA (25 $\mu\text{g}/\mu\text{L}$).

4.2.2 qPCR Analysis with SYBR Green

The validation started with qPCR analysis using SYBR Green. Master mix for 42 reactions contained 16.8 μL of working solution of the primer mix, 210 μL of TATAA SYBR GrandMaster mix and 109.2 μL of distilled water. The reagents were mixed and spun down in a 1.5 mL tube. This subsection was repeated for all analyzed genes. gBlocks were diluted to a concentration of $1 \cdot 10^7$ molecules/ μL , through 200 μL distilled water minus 20 μL gBlocks for each assay. Therefore, the volume of the water varied between each run. The diluted gBlock was used to execute a dilution series for a standard curve. The dilution series was performed as below:

1. C: $1 \cdot 10^7$ molecules/ μL
2. 20 μL from 1. + 180 μL TE-LPA buffer. C: $1 \cdot 10^6$ molecules/ μL
3. 20 μL from 2. + 180 μL TE-LPA buffer. C: $1 \cdot 10^5$ molecules/ μL
4. 20 μL from 3. + 180 μL TE-LPA buffer. C: $1 \cdot 10^4$ molecules/ μL
5. 20 μL from 4. + 180 μL TE-LPA buffer. C: $1 \cdot 10^3$ molecules/ μL
6. 20 μL from 5. + 180 μL TE-LPA buffer. C: $1 \cdot 10^2$ molecules/ μL
7. 20 μL from 6. + 180 μL TE-LPA buffer. C: $1 \cdot 10^1$ molecules/ μL

The dilutions were carefully mixed between each concentration and spun down in 1.5 ml tubes. The content of all tubes were conveyed to a strip with 8 wells. In the first well, 180 μL TE-LPA buffer was added as a No template control, NTC. The tubes of master mixes, cDNA, gDNA and the strip were added in the right place according to the robot program. The robot executed the pipetting, 8 μL of the master mixes and 2 μL of the dilution series, cDNA and gDNA.

After the pipetting, the plate was sealed with sealing tape and then carefully mixed and centrifuged in 1 minute at 3000 rpm. Afterwards, the plate was analyzed with a following melt curve analysis in the instrument BIO-RAD CFX384 according to the temperature program in *Table 1*.

Table 1. Temperature program for qPCR analysis with SYBR Green with following melt curve

Program	Temp (°C):	Time (s):	Cycles:
<i>Polymerase activation:</i>	95	60	1
Denaturation	95	5	45
Annealing	60	30	
Extension	72	10	
Denaturation	95	15	1
Annealing	55	15	
Melt	95	-	

4.2.3 Verification of Products with Fragment Analyzer

There were some preparations before analysis with Fragment Analyzer. General preparation intended manufacturing of DNA Gel and change of Inlet Buffer. The DNA Gel was manufactured of 12 ml of dsDNA Separation Gel and 1.2 μL of intercalating dye mixture, this volume was enough for 12 samples. The next preparation was to remove old Inlet Buffer from a 96 well plate, which is used as storage for the capillaries. When the wells were empty, the wells were filled with 1 mL of new Inlet Buffer.

Sample preparation was a preparation of a ladder and samples. 24 μL of 100 bp DNA Ladder solution was added into one well 12. During the next runs, the ladder was imported. It was needed two 96 plates, one for marker solution and one for DNA samples. The marker plate consisted of prepared marker solution and was recycled for all runs. The sample plate was filled with 22 μL of 1x TE buffer in each position and 2 μL of PCR sample were added. For every assay there were four analyzed samples, NTC, gBlock, cDNA and gDNA. Samples were collected from the plate from the qPCR analysis with SYBR Green. The sample plate was sealed with sealing tape and carefully mixed and centrifuged in 1 minute at 3000 rpm. After the centrifuge, the sealing tape was taken off. Both sample plate and marker plate were placed in the instrument and the analysis was started.

4.2.4 qPCR Analysis with Probe

The assays which met the criteria for SYBR Green and Fragment Analyzer were analyzed with Probe. For this experiment a new master mix was prepared. The master mix was prepared for 37 reactions and contained 14.8 μL of the working solution of primer mix, 7.4 μL of probe, 185 μL of TATAA Probe GrandMaster mix and 88.8 μL of distilled water. The reagents were mixed and spun down in a 1.5 mL tube. This subsection was repeated for all analyzed genes. gBlocks were diluted to a concentration of $1 \cdot 10^7$ molecules/ μL , through 200 μL distilled water minus 20 μL gBlocks for each assay. Therefore, the volume of the water varied between each run. The diluted gBlock was used to execute a dilution series for a standard curve. The dilution series was performed as below:

1. C: $1 \cdot 10^7$ molecules/ μL
2. 20 μL from 1. + 180 μL TE-LPA buffer. C: $1 \cdot 10^6$ molecules/ μL
3. 20 μL from 2. + 180 μL TE-LPA buffer. C: $1 \cdot 10^5$ molecules/ μL
4. 20 μL from 3. + 180 μL TE-LPA buffer. C: $1 \cdot 10^4$ molecules/ μL
5. 20 μL from 4. + 180 μL TE-LPA buffer. C: $1 \cdot 10^3$ molecules/ μL
6. 20 μL from 5. + 180 μL TE-LPA buffer. C: $1 \cdot 10^2$ molecules/ μL
7. 20 μL from 6. + 180 μL TE-LPA buffer. C: $1 \cdot 10^1$ molecules/ μL

The dilutions were carefully mixed between each concentration and spun down in 1.5 ml tubes. The content of all tubes were conveyed to a strip with 8 wells. In the first well, 180 μL TE-LPA buffer was added as No template control, NTC. The tubes of master mixes and the strip were added in the right place according to the robot program. The robot executed the pipetting, 8 μL of the master mixes and 2 μL of the dilution series. After the pipetting, the plate was sealed with sealing tape and then carefully mixed and centrifuged in 1 minute at 3000 rpm. Afterwards, the plate was analyzed in the instrument BIO-RAD CFX384 according to the temperature program in *Table 2*.

Table 2. Temperature program for qPCR analysis with probe

Program	Temp (°C):	Time (s):	Cycles:
Polymerase activation:	95	60	1
Amplification:			
Denaturation	95	5	50
Annealing	60	30	

4.2.5 Pre-amplification with SYBR Green and Probe

During the first part of the pre-amplification all 62 assays were mixed with 12 reference genes and ValidPrime. Consequently, 2 µL stock solution of all primer pairs except RRN18S were mixed in a 1.5 mL tube, called primer solution. Concentration of the primer solution was 500 nM. A sample, which expressed all genes, needed to be manufactured. The sample consisted of non-diluted cDNA, spiked gBlocks and distilled water. The concentration of the gBlocks were too high and were diluted to a lower concentration of $1 \cdot 10^5$ molecules/µL. 2 µL of the diluted gBlocks for the low expressed assays and 80 µL of non-diluted cDNA were mixed and diluted with distilled water to a volume of 800 µL.

A master mix was prepared for the PCR analysis and it consisted of 30 µL of the primer solution, 150 µL of TATAA GrandMaster mix and 60 µL of distilled water. The master mix was mixed, spun down and conveyed to an 8 well strip. The first 5 wells were filled with 40 µL of the master mix. Three of the wells were then filled with 10 µL of cDNA + gBlocks and the two remaining wells were filled with 10 µL of distilled water. The strip was carefully mixed, placed in the instrument BIO-RAD CFX96 and the PCR analysis was started according to the temperature program in Table 3.

Table 3. Temperature program for the first step in the pre-amplification

Program	Temp (°C):	Time (s):	Cycles:
Polymerase activation:	95	60	1
Denaturation	95	15	17
Annealing	60	120	
Extension	72	60	

The samples in the strip were frozen immediately after the end of the program by using dry ice and then stored in a freezer until further analysis. The described step above is the same for analysis with SYBR Green and Probe. Last step of the pre-amplification meant an analysis of the products with qPCR. A difference between the preparation of the master mixes differed between analysis with SYBR Green and Probe, see Table 4.

Table 4. The difference of the master mix content between SYBR Green and Probe.

SYBR Green			Probe	
Reagents	Volume (μL)		Reagents	Volume (μL)
Water	46.8		Water	43.2
TATAA SYBR GrandMaster mix	90		TATAA SYBR GrandMaster mix	90
Primer mix	7.2		Primer mix	7.2
			Probe mix	3.6

8 μL of the master mixes were pipetted in each well. The frozen strip was thawed by added 350 μL of distilled water. Thence, 2 μL of each solution in the strip were pipetted in replicate to the plate by using the robot. Two of the empty wells in the strip were filled with cDNA which have not passed through a preamplification. The last well was filled with distilled water as NTC. The plates were sealed with sealing tape, mixed and centrifuged in 1 minute at 3000 rpm and then analyzed with the instrument BIO-RAD CFX384 according to the temperature program in Table 5.

Table 5. Temperature program for SYBR Green and Probe for the qPCR analysis of the last step in the preamplification.

Program	SYBR Green			Probe		
	Temp ($^{\circ}\text{C}$):	Time (s):	Cycles:	Temp ($^{\circ}\text{C}$):	Time (s):	Cycles:
Polymerase activation	95	60	1	95	60	
Denaturation	95	5		95	5	
Annealing	60	30	45	60	30	45
Extension	72	10		-	-	
	Melt curve			-		

5. RESULTS AND DISCUSSION

5.1 Design of Selected Genes

The interesting genes for the panel have been selected from literature studies, requests from external partners and studies of commercial products at the market. The panel consists of important genes for stem cell differentiation and their assay name, full gene name, alias and amplicon length can be read in *Table I* in Appendix. The panel focus on stem cells differentiation for detection of kidney cells, nervous system cells, cardiac cells, smooth muscle cells, endothelial cells, mesenchymal stem cells pluripotency and hepatic tissue. The selected genes are sorted by type of differentiation. According to the purpose of the project, the panel will consist of 48 genes but *Table I* consist of 61 selected genes. The reason is it can be difficult to predict which assays meet the criteria for the validation plan. 50 of these genes have been designed and the remaining genes were already designed of staff at TATAA Biocenter.

The design of the genes has been difficult for some assays. The biggest difficulty has been to get all protein coding transcript variants. To include all variants, there must be common sequences for all transcripts. When it has been difficult to design a gene, a website called Clustal Omega have been used. This website will find common sequences, which is useful for the design of the primers. Sometimes this has not been achieved and the products miss one or more transcripts, where WT1, CD44 and GATA2 are example of these genes. One of the settings in the design tool Primer-BLAST has been a separation of the exons with an intron with ≥ 1000 base pairs or an exon-exon junction. These settings have sometimes caused problems since some of the genes have no introns with that size. In these cases, the primers have been designed over an exon. This may give rise to the same melting profile for gDNA as the assay, which generate the assay to detect genomic DNA. This problem can be solved by using deoxyribonuclease, DNase, specific for single stranded DNA. DNase is an enzyme that catalyzes the hydrolytic cleavage of phosphodiester in the DNA.

The settings for amplicon length were set between 80 – 200 base pairs, but the requirement were changed for the assays WT1, SIX1, EYA1, EPHB2, DMT3B and LIN28A. The changes in the settings depended on difficulties to design an assay with all protein coding transcripts variants.

5.2 qPCR Analysis with SYBR Green and probe

The designed genes were analyzed with the qPCR technique, with both SYBR Green and Taqman probe. The used instrument for the analysis was BIO-RAD CFX384. Analysis with SYBR Green implied a succeeding melt curve analysis. Following figures, *Figure 3* and *Figure 4*, show the results for CD44 from analysis with SYBR Green.

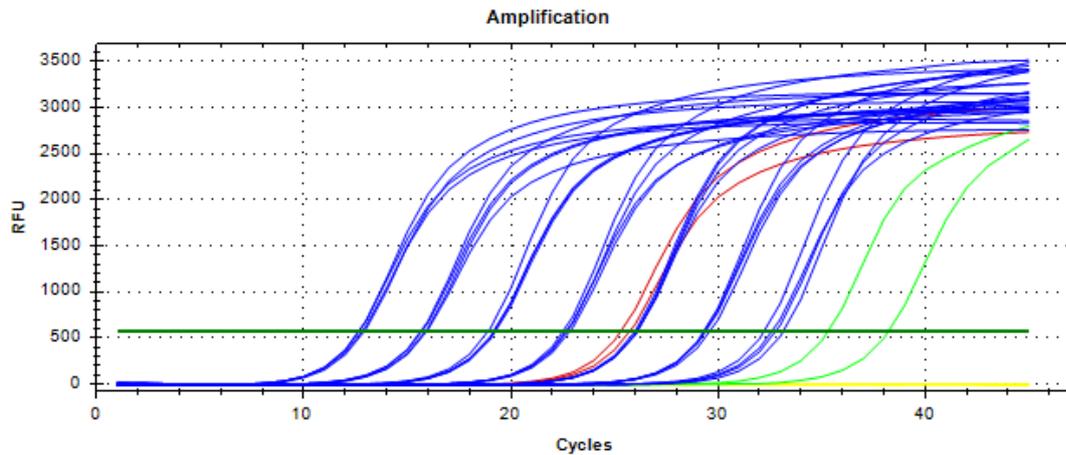


Figure 3. Quantification of assay CD44 with SYBR Green
 Blue: assay, Yellow: NTC, Red: cDNA, Green: gDNA

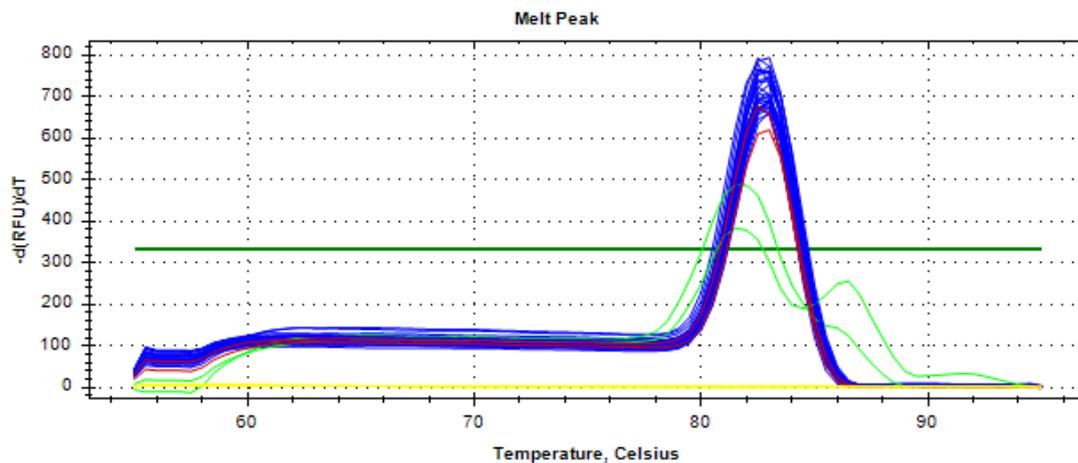


Figure 4. Melt curve for assay CD44 with SYBR Green
 Blue: assay, Yellow: NTC, Red: cDNA, Green: gDNA

Results *Figure 3* demonstrating that the assay has a great dilution series and it is a small dispersion between the replicates. There is a greater dispersion in the dilution 7, but it depends on the low concentrated standard. The replicates with NTC gave no products. *Figure 4* represent a melt curve for the assay. The melt profile for the assay is good and cDNA has the same melt profile as the assay. gDNA have another melt profile compared to the assay and cDNA, which are good. In *Figure 5*, the standard curve for CD44 can be seen.

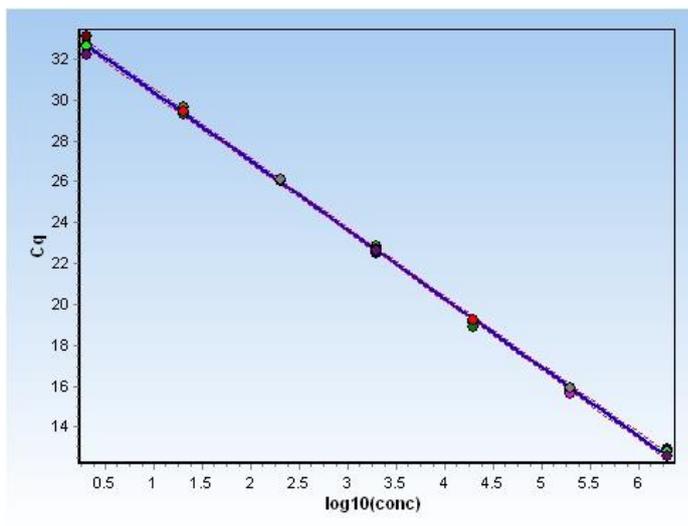


Figure 5. Standard curve for assay CD44 with SYBR Green

The standard curve in Figure 5 is linear and the replicates are located on the custom line. This is an evidence that the dilution was successful. It can be supported by results in Table 6.

Table 6. Results from the qPCR analysis with SYBR for CD44

Assay	Efficiency	R ²	Slope	Intercept	Non-linearity	
CD44	0.98	0.9991	-3.36	33.7	0.16	SYBR
CD44	0.96	0.9995	-3.415	36.12	0.06	Probe
ETV2	0.76	0.9732	-4.39	49.1	linear	SYBR

Results for CD44 with SYBR in Table 6 are good according to the criteria. The efficiency is between; 80 – 110 % and the R²-value is over 0.995. The non-linearity is 0.16 and the limit of linearity is ≤ 0.20. Therefore, the assay is linear and meet the criteria. These values and figures prove that CD44 is an approved assay and will be included in the panel. To illustrate a less good assay, an example with ETV2 will be shown.

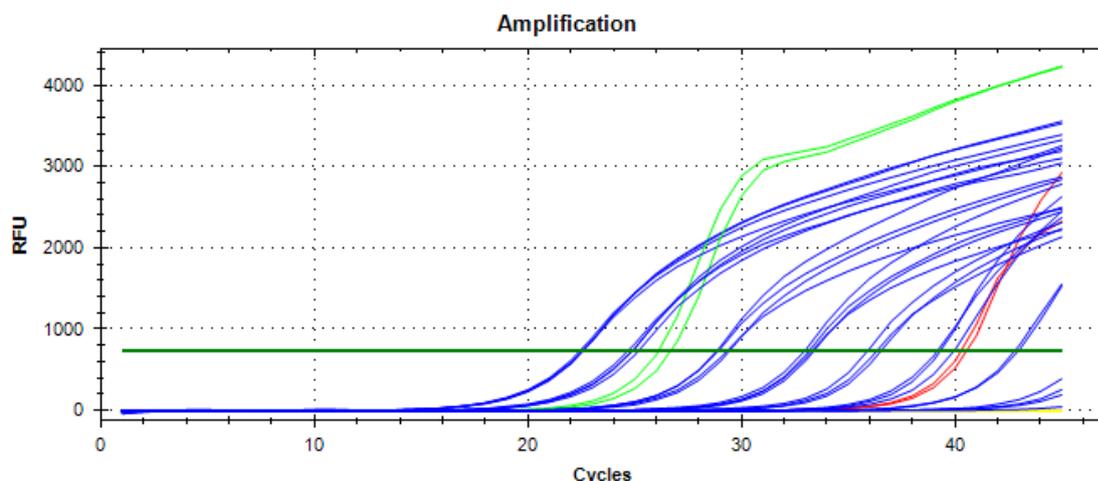


Figure 6. Quantification of assay ETV2 with SYBR Green
Blue: assay, Yellow: NTC, Red: cDNA, Green: gDNA

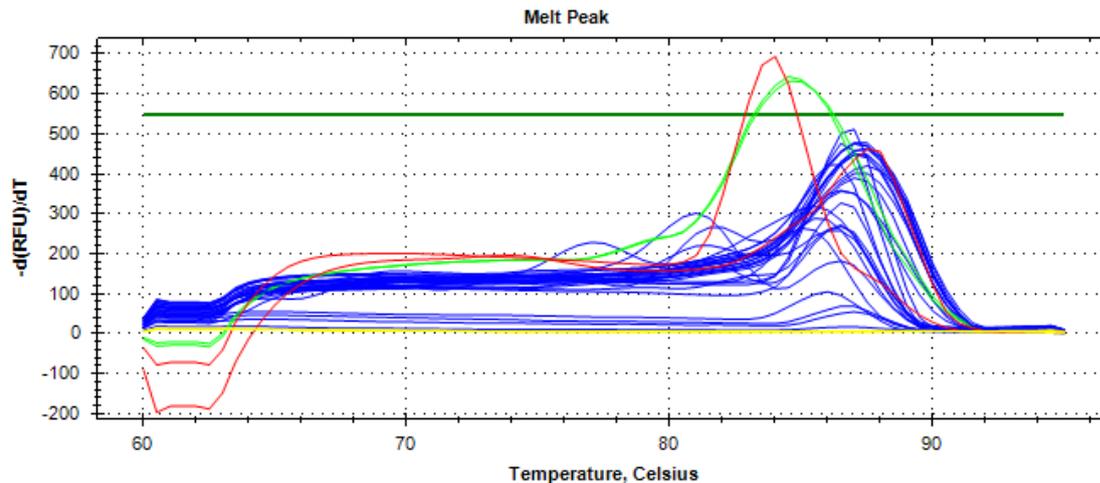


Figure 7. Melt curve for assay ETV2 with SYBR Green
Blue: assay, Yellow: NTC, Red: cDNA, Green: gDNA

Results for the assay ETV2 in Figure 6 and Figure 7 are not good compared to the result for the assay CD44. The blue curves in the melt curve are primer dimers. This is not desirable for an assay and is a sign of redesign. This is confirmed with the standard curve in Figure 8. The values for efficiency and R^2 in Table 6 are not good either.

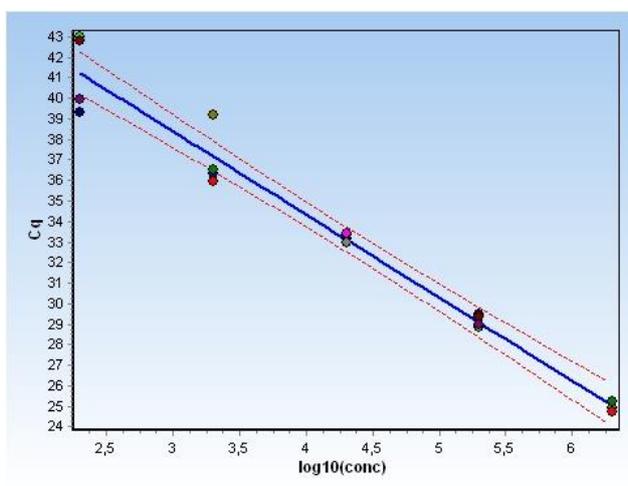


Figure 8. Standard curve for assay ETV2 with SYBR Green

ETV2 has been redesigned since the efficiency was too low. Except ETV2, these assays have been redesigned: TAT, GATA4, CDC42, CDK1, DPPA2, DPPA3, DPPA4, DPPA5, HSPA9, MYBL2, NR6A1, SIX2, SALL1, EPHA7 and DLK1. These have been redesigned on account of pseudogenes, primer dimers and low efficiency. Pseudogenes are genes that have lost their gene expression or ability to code protein and these genes are not desirable.

Analysis with probe exclude cDNA, gDNA and a melt curve analysis. The purpose of the probe analysis is to increase the efficiency of the assay. The use of probe avoids primer dimers since the master mix does not take consideration of nonspecific products. Master mix for SYBRE Green is trying to minimize the formation of primer dimer, but usually not enough. Therefore, there is an advantage to use probe since there are no formation of primer dimers and an increased efficiency. Figure 9 shows results from qPCR analysis for the assay CD44 with probe.

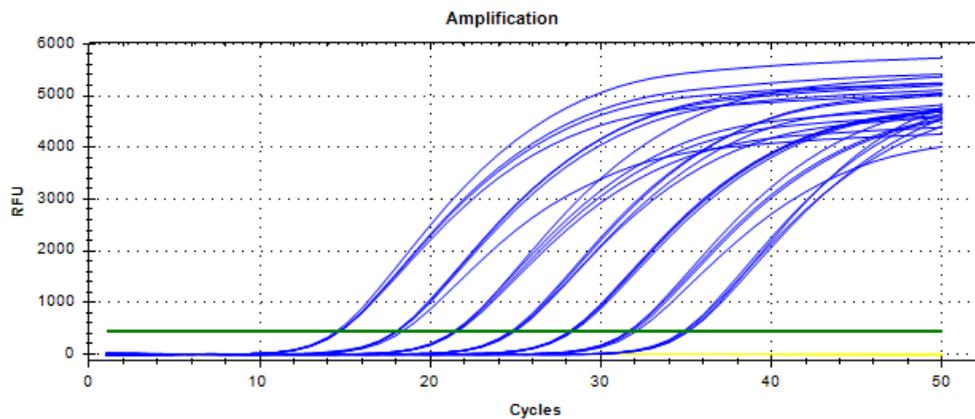


Figure 9. Quantification of the assay CD44 with probe
Blue: assay, Yellow: NTC

Usually, the efficiency increase when probe is used. This does not apply on the assay CD44. See Table 6 for a comparison of the results from the SYBR Green analysis and the probe analysis of CD44. The efficiency has decreased for assay CD44, but for other assays the efficiency has increased. Results from all assays can be found in Table II in Appendix. It is only the designs which are approved. Further, the results in the Appendix are only the designed assays, others assays were already designed and have only been analyzed during the preamplification. Examples of plate setups for both SYBR Green and probe can also be seen in Appedendix. Figure 10 shows an example of a less good probe.

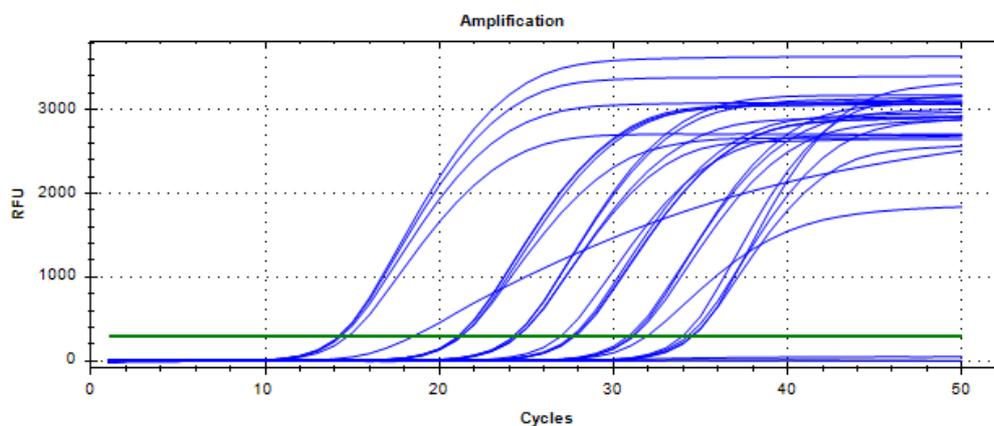


Figure 10. Quantification of the assay CD44 with probe
Blue: assay, Yellow: NTC

Amplification curve in Figure 10 is not good. Probably something has happed with the dilution of the standards since Figure 9 is the same probe with a new dilution series. An assay like Figure 10 has probably not good values for R^2 and is not linear, which are two criterions for an assay to be approved.

5.3 Verification of the Product Length

The length of the product was controlled with the instrument Fragment Analyzer. NTC, gBlocks, cDNA and gDNA for all assays were checked. To facilitate, the example of the assay CD44 will be used. *Figure 11 – 14* shows the results from the analysis with Fragment Analyzer.

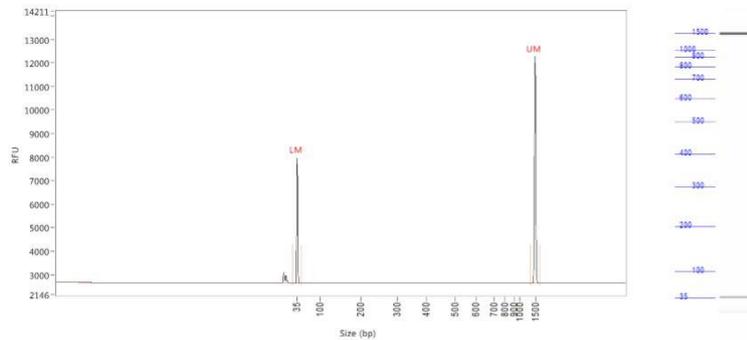


Figure 11. Analysis with Fragment Analyzer for CD44, NTC

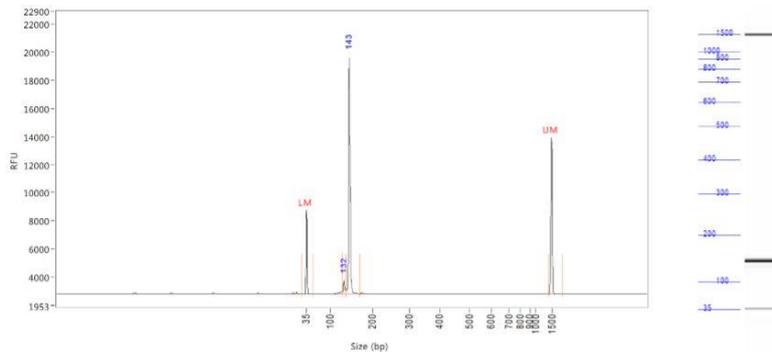


Figure 12. Analysis with Fragment Analyzer for CD44, gBlocks

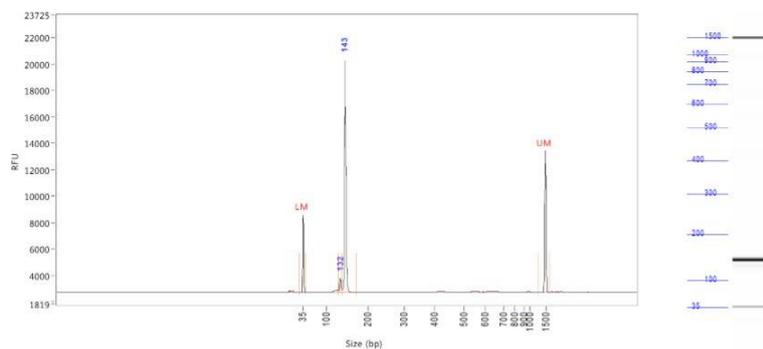


Figure 13. Analysis with Fragment Analyzer for CD44, cDNA

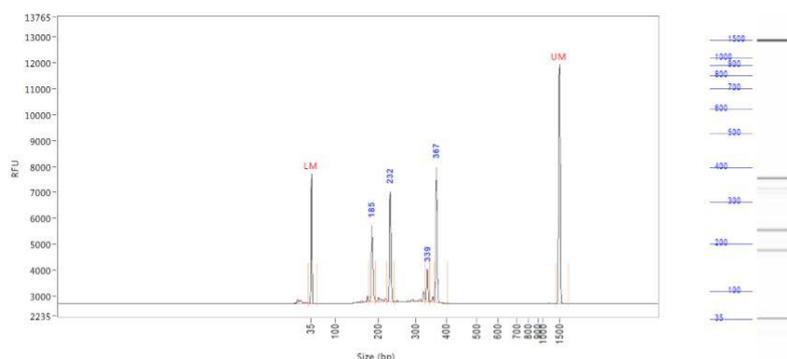


Figure 14. Analysis with Fragment Analyzer for CD44, gDNA

The results in the figures above shows a good assay with both lower and upper markers for NTC, gBlock, cDNA and gDNA. The length of the product is the same for the gBlock and cDNA, which is good. In the gDNA sample, there are many products, which is good. Accordingly, the assay does not detect genomic DNA. A comparison between the theoretical and the practical length will be ± 10 base pairs. The theoretical length is 137 base pairs and the practical length is 143 base pairs, which is within the restriction. Assays with the same length of the assay, cDNA and gDNA is passing, but it should be avoided. It also applies to unspecific products in NTC replicates. The length of the unspecific products should be ≤ 50 base pairs to pass. The results for all assays can be found in *Table II* in Appendix. One example of plate setup can be seen in Appendix.

5.4 Preamplification

The last step of the validation plan was a preamplification. All primers were mixed with Validprime and all the reference genes expect RRN18S. Validprime is added to avoid the performance of no reverse transcriptase controls to test for the presence of genomic DNA. RRN18S is a part of the ribosomal RNA and is not amplified since the gene already is present in a large amount. An amplification of RRN18S could inhibit the other genes' amplification. The used reference genes can be seen in *Table 7*.

Table 7. Reference Genes from TATAA Biocenter's Reference Gene Panel Human

Assay	Full Gene Name	Alias
ACTB	Actin, beta	BRWS1, PS1TP5BP1
B2M	Beta-2-microglobulin	IMD43
GAPDH	Glyceraldehyde-3-phosphate dehydrogenase	G3PD, GAPD, HEL-S-162eP
GUSB	Glucuronidase beta	BG, MPS7
HPRT1	Hypoxanthine phosphoribosyltransferase 1	HPRT, HGPRT
PPIA	Peptidylprolyl isomerase A	CYPA, CYPH, HEL-S-69p
RPLP	60S acidic ribosomal protein P0	-
RRN18S	18s rRNA	-
TBP	TATA-box binding protein	HDL4; GTF2D; SCA17; TFIID; GTF2D1
TUBB	Tubulin beta class I	M40, TUBB1, TUBB5, CDCBM6, CSCSC1, OK/SW-cl.56
UBC	Ubiquitin C	HMG20
YWHAZ	Tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein zeta	HEL4, YWHAD, KCIP-1, HEL-S-3, HEL-S-93, 14-3-3-zeta

These reference genes are designed to be expressed in the most part of samples. This is no guarantee but usually at least two reference genes are expressed in the sample and can be used for normalization of the remaining genes.

Results from the preamplification have been successful. Almost every gene has efficiency over 90 %. There are some exceptions, where EYA1, EPHA7 and EPHB3 are examples. In particular, there have been some problems with EPHA7 earlier in the validation. The gene has been redesigned because of low efficiency. The problem has followed to the preamplification. Furthermore, the genes with low efficiency are reproducible and are therefore acceptable. *Figure 15* and *Figure 16* show the variation of ΔC_q and the efficiency of a selection of the assays.

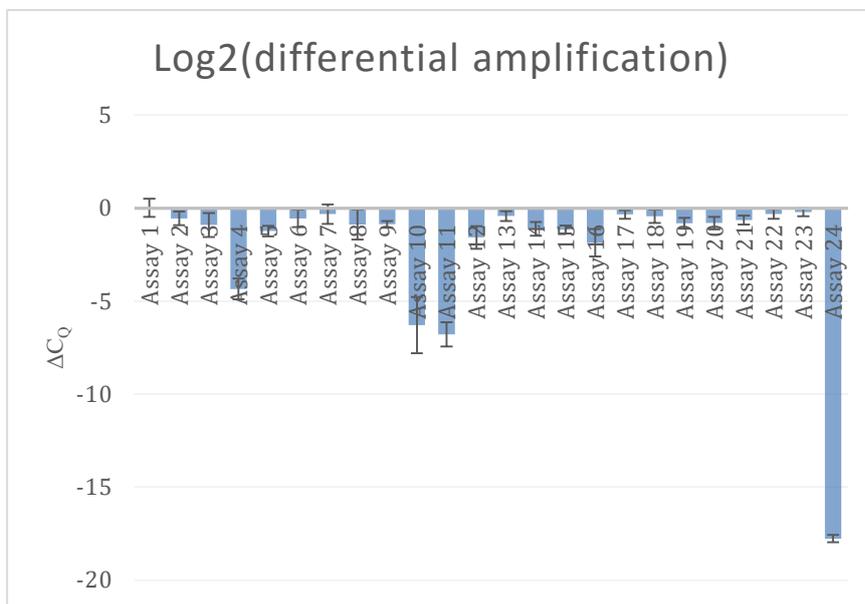


Figure 15. Reproducibility and error bars for some assays

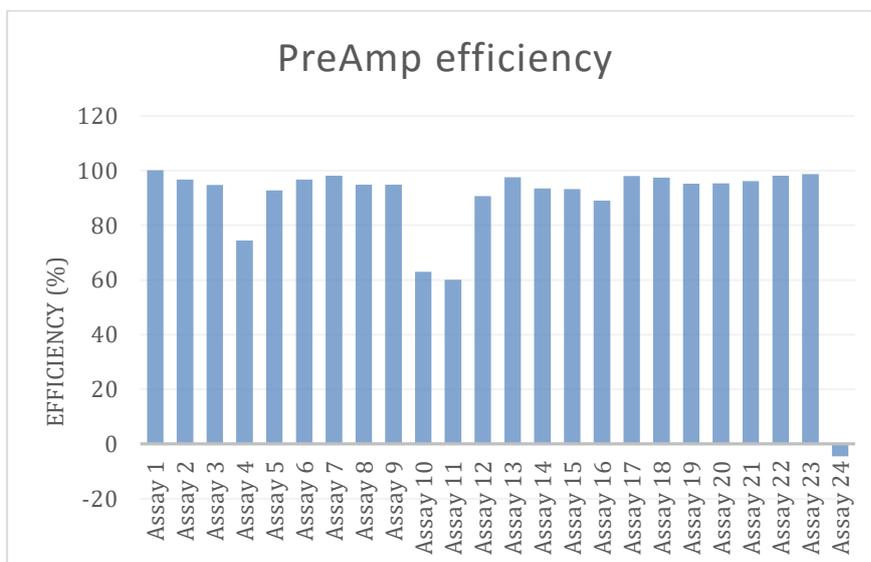


Figure 16. The variation of efficiency between some assays

Assay 24 is RRN18S and has a negative efficiency. It depends on RRN18S is not preamplified. The other assays have similar efficiencies except assay 4, 10 and 11. These assays are reproducible according to the blue staples in *Figure 15*. Error bars in *Figure 15* are a representation of the variability of ΔC_q between the preamplified and the non-preamplified. Dilution factors and cycles have been taken into account. The error bars are used to indicate error or uncertainty in the measurement. In this experiment the error bars are regarding at the standard deviation. The error bars are low for all assays except assay 10. The higher error bar for assay 10 depends on wide spread between the replicates. Generally, the C_q values are high for the non-preamplified reactions compared to the preamplified. The non-preamplified reactions consists of fewer copies and the result will become haphazard. This can lead to incorrect efficiencies that can not be trusted. For example, OC90 have an efficiency at 133 % and had only one result for the non-preamplified reaction. Standard deviations for the assays are calculated by average of the standard deviations between the preamplified and non-preamplified. In other words, a high standard deviation for the non-preamplified reaction increase the overall standard deviation. This causes that all raw data must be inspected before drawing a conclusion. In some cases, raw data must be removed to obtain a correct result.

There has been some problem with the delivery of the probes and five probes have not arrived. This has contributed to that no preamplification with probe have been performed. Therefore, an preamplification with probe will be performed when all probes have arrived. An advantage with analysis with probe is the minimization of non-specific products. A preamplification with probe is usually used in the analysis of real samples. Therefore, it is important that this part will be performed.

6. CONCLUSION

The purpose of the project was to design and validate a qPCR panel for stem cells differentiation for human cells. The development implied design of assays, qPCR analysis with SYBR Green and probe, control of PCR product and finally a preamplification. The design and qPCR analysis of the assays have been successful. Some of the genes have been redesigned due to primer dimers and low efficiency. The theoretical and practical lengths of the amplicons were the same for all assays. The final step was the preamplification of all assays including reference genes and ValidPrime. The preamplification with SYBR Green was efficient. Almost all assays had good efficiencies and all assays were reproducible. The assays with low efficiencies were reproducible and were approved.

The result of this project is a developed qPCR panel for stem cells differentiation for human cells. The qPCR based panel can now be used of scientist to detect the differentiation state of added stem cells. So far, the panel is only validated for preamplification with SYBR Green. In the near future, the panel also will be validated with probe. This will increase the demand of the panel since probe increases the efficiency and nonspecific products are not formed.

REFERENCES

- Andersson, D. et al., 2015. Properties of targeted preamplification in DNA and cDNA quantification. *Expert review of molecular diagnostics*, 15(8), pp.1085–100. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4673511&tool=pmcentrez&rendertype=abstract> [Accessed May 10, 2016].
- Anon, Dictionary of Science (6th Edition) - PCR - Knovel. Available at: https://app.knovel.com/web/view/swf/show.v/rcid:kpDSE00001/cid:kt007X8VYC/viewerType:pdf/root_slug:dictionary-science-6th?cid=kt007X8VYC&page=9&b-q=PCR&sort_on=default&b-subscription=TRUE&b-group-by=true&b-search-type=tech-reference&b-sort-on=default&scrollto=PCR See https://app.knovel.com/web/view/swf/show.v/rcid:kpDSE00001/cid:kt007X8VYC/viewerType:pdf/root_slug:dictionary-science-6th?cid=kt007X8VYC&page=9&b-q=PCR&sort_on=default&b-subscription=TRUE&b-group-by=true&b-search-type=tech-reference&b-sort-on=default&scrollto=PCR [Accessed May 10, 2016].
- Biocenter, T., TATAA Biocenter. Available at: <http://www.tataa.com/> [Accessed May 10, 2016].
- Britannica, polymerase chain reaction -- Britannica Online Encyclopedia. Available at: <http://academic.eb.com.proxy.lib.chalmers.se/EBchecked/topic/468736/polymerase-chain-reaction> [Accessed May 10, 2016].
- Bustin, S.A., 2004. *A-Z of Quantitative PCR*, IUL Biotechnology Series.
- Conrad, S. et al., 2016. Expression of Genes Related to Germ Cell Lineage and Pluripotency in Single Cells and Colonies of Human Adult Germ Stem Cells. *Stem Cells International*, 2016, pp.1–17. Available at: <http://www.hindawi.com/journals/sci/2016/8582526/> [Accessed May 10, 2016].
- Kubista, M., 2010. Real-time PCR where we are and where are we heading. *Drug Discovery World*, 11(2), pp.59–65. Available at: https://www.researchgate.net/publication/289877155_Real-time_PCR_where_we_are_and_where_are_we_heading [Accessed May 10, 2016].
- Lakshmi, V. et al., 2011. Application of Polymerase Chain Reaction to Detect HIV-1 DNA in Pools of Dried Blood Spots. *Indian journal of microbiology*, 51(2), pp.147–52. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3209897&tool=pmcentrez&rendertype=abstract> [Accessed April 2, 2016].
- Nationalencyklopedin, genteknik - Uppslagsverk - NE. Available at: <http://www.ne.se.proxy.lib.chalmers.se/uppslagsverk/encyklopedi/l%C3%A5ng/genteknik> [Accessed May 10, 2016a].
- Nationalencyklopedin, stamcell - Uppslagsverk - NE. Available at: <http://www.ne.se.proxy.lib.chalmers.se/uppslagsverk/encyklopedi/l%C3%A5ng/stamcell> [Accessed May 10, 2016b].
- Noaksson, K. et al., Monitoring differentiation of human embryonic stem cells using real-time PCR. *Stem cells (Dayton, Ohio)*, 23(10), pp.1460–7. Available at:

<http://www.ncbi.nlm.nih.gov/pubmed/16081663> [Accessed April 5, 2016].

Nolan, T. & Bustin, S., 2013. *PCR Technology* Third Edit., CRC Press Taylor & Francis Group.

R&D, Cardiac Stem Cell Markers Research Areas : R&D Systems. Available at: <https://www.rndsystems.com/research-area/cardiac-stem-cell-markers> [Accessed May 10, 2016].

Ståhlberg, A. & Kubista, M., 2014. The workflow of single-cell expression profiling using quantitative real-time PCR. *Expert review of molecular diagnostics*, 14(3), pp.323–31. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4819576&tool=pmcentrez&rendertype=abstract> [Accessed May 10, 2016].

ThermoFisher SCIENTIFIC, TaqMan® Assays vs. SYBR® Green Dye for qPCR. Available at: <https://www.thermofisher.com/se/en/home/life-science/pcr/real-time-pcr/qpcr-education/taqman-assays-vs-sybr-green-dye-for-qpcr.html> [Accessed May 10, 2016].

Tsankov, A.M. et al., 2015. A qPCR ScoreCard quantifies the differentiation potential of human pluripotent stem cells. *Nature biotechnology*, 33(11), pp.1182–92. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4636964&tool=pmcentrez&rendertype=abstract> [Accessed April 9, 2016].

Tymoczko, J.L., Berg, J.M. & Stryer, L., 2013. *BIOCHEMISTRY A Short Course* Second Edi., W.H. Freeman and Company.

APPEENDIX

Table I. The selected genes for stem cells differentiation

Assay	Full Gene Name	Alias	Amplicon length (bp)
Kidney Markers			
PROM1	Prominin 1	RP41, AC133, CD133, MCDR2, STGD4, CORD12, PROML1, MSTP061	196
EYA1	EYA transcriptional coactivator and phosphatase 1	BOP, BOR, OFC1	209
Nervous system Markers			
EPHA7	EPH receptor A7	EHK3, EK11, EHK-3, HEK11	144
EPHB3	EPH receptor B3	ETK2, HEK2, TYRO6	124
EPHA6	EPH receptor A6	EHK2, EK12, EPA6, EHK-2, HEK12, PRO57066	145
PAX2	Paired box 2	FSGS7, PAPRS	148
EPHB2	EPH receptor B2	DRT, EK5, ERK, CAPB, Hek5, PCBC, EPHT3, Tyro5	67
Cardiomyocyte Markers			
ISL1	ISL LIM homeobox 1	Isl-1, ISLET1	136
NKX2-5	NK2 homeobox 5	CSX, CSX1, VSD3, CHNG5, HLHS2, NKX2E, NKX2.5, NKX4-1	105
GATA4	GATA binding protein 4	TOF, ASD2, VSD1, TACHD	95
GATA6	GATA binding protein 6	-	80
Hematopoietic Stem cell/Early Endothelial Cell Markers			
FLII	Flightless I actin binding protein	FLI, FLIL, Fli1	95
TAL1	T-cell acute lymphocytic leukemia 1	SCL, TCL5, tal-1, bHLHa17	89
TEK	TEK receptor tyrosine kinase	TIE, VMCM, TIE-2, VMCM1, CD202B	149
GATA2	GATA binding protein 2	DCML, IMD21, NFE1B, MONOMAC	121
ETV2	Ets variant 2	ER71, ETSRP71	83
Early Smooth Muscle Cell Markers			
ACTA2	Actin, alpha 2, smooth muscle, aorta	AAT6, ACTSA, MYMY5	116
TAGLN	Transgelin	SM22, SMCC, TAGLN1, WS3-10	134
Mesenchymal Stem Cell Markers			
CD44	CD44 molecule (Indian blood group)	IN, LHR, MC56, MDU2, MDU3, MIC4, Pgp1, CDW44,	137

		CSPG8, HCELL, HUTCH-I, ECMR-III	
ENG	Endoglin	END, HHT1, ORW1	140
THY1	Thy-1 cell surface antigen	CD90, CDw90	135
NT5E	5'-nucleotidase ecto	NT, eN, NT5, NTE, eNT, CD73, E5NT, CALJA	116
Hepatic Tissue Markers			
TAT	Tyrosine aminotransferase	-	137
ALB	ALBUMIN	FDAH, ANALBA, PRO0883, PRO0903, PRO1341	121
Pluripotency Markers			
AICDA	Activation-induced cytidine deaminase	AID, ARP2, CDA2, HIGM2, HEL-S-284	118
CCNA2	Cyclin A2	CCN1, CCNA	138
DPPA2	Developmental pluripotency associated 2	CT100, PESCRG1, ECAT15-2	140
DPPA3	Developmental pluripotency associated 3	STELLA	137
DPPA4	Developmental pluripotency associated 4	2410091M23Rik	117
DPPA5	Developmental pluripotency associated 5	ESG1	89
NR5A2	Nuclear receptor subfamily 5 group A member 2	B1F, CPF, FTF, B1F2, LRH1, LRH-1, FTZ-F1, hB1F-2, FTZ-F1beta	108
NR6A1	Nuclear receptor subfamily 6 group A member 1	RTR, GCNF, NR61, hRTR, CT150, GCNF1, hGCNF	121
OTX2	Orthodenticle homeobox 2	CPHD6, MCOPS5	107
RIF1	Replication timing regulatory factor 1	-	147
SCN1A	Sodium voltage-gated channel alpha subunit 1	FEB3, FHM3, NAC1, SCN1, SMEI, EIEE6, FEB3A, HBSCI, GEFSP2, Nav1.1	146
SOX15	SRY-box 15	SOX20, SOX26, SOX27	103
TCL1A	T-cell leukemia/lymphoma 1A	TCL1	126
CDC42	Cell division cycle 42	TKS, G25K, CDC42Hs	192
CDK1	Cyclin-dependent kinase 1	CDC2, CDC28A, P34CDC2	163
HSPA9	Heat shock protein family A (Hsp70) member 9	CSA, MOT, MOT2, CRP40, GRP75, PBP74, GRP-75, HSPA9B, MTHSP75, HEL-S-124m	138
MYBL2	V-myb avian myeloblastosis viral oncogene homolog-like 2	BMYB, B-MYB	86
ZFP42	ZFP42 zinc finger protein	REX1, REX-1, ZNF754, zfp-42	88
FGF2	Fibroblast growth factor 2	BFGF, FGFB, FGF-2, HBGF-2	97
SOX2	SRY (sex determining region Y)-box 2	ANOP3, MCOPS3	157
LIN28A	Lin-28 homolog A (C. elegans)	CSDD1, LIN28, LIN-28,	216

		ZCCHC1, lin-28A	
Embryonic Development Markers			
TDGF1	Teratocarcinoma-derived growth factor 1	CR, CRGF, CRIPTO	186
GDF3	Growth differentiation factor 3	KFS3, MCOP7, MCOPCB6	174
AFP	Alpha-fetoprotein	AFPD, FETA, HPAFP	156
DES	Desmin	CSM1, CSM2, LGMD2R	117
DNMT3B	DNA (cytosine-5-)-methyltransferase 3 beta	ICF, ICF1, M.HsaIIIB	213
POU5F1	POU class 5 homeobox 1	OCT3, OCT4, OTF3, OTF4, OTF-3, Oct-3, Oct-4	146
NANOG	Nanog homeobox	-	166
Other Stem Cell Markers			
WT1	Wilms Tumor 1	GUD, AWT1, WAGR, WT33, NPHS4, WIT-2, EWS-WT1	202
SIX1	SIX homeobox 1	BOS3, TIP39, DFNA23	79
SIX2	SIX homeobox 2	-	89
SALL1	Spalt-like transcription factor 1	TBS, HSAL1, Sal-1, ZNF794, HEL-S-89	127
CITED1	Cbp/p300-interacting transactivator with Glu/Asp rich carboxy-terminal domain, 1	MSG1	198
DLK1	Delta-like 1 homolog (Drosophila)	DLK, FA1, ZOG, pG2, DLK-1, PREF1, Delta1, Pref-1	124
EPCAM	Epithelial cell adhesion molecule	ESA, KSA, M4S1, MK-1, DIAR5, EGP-2, EGP40, KS1/4, MIC18, TROP1, EGP314, HNPCC8	145
LHX1	LIM homeobox 1	LIM1, LIM-1	116
FOXD3	FOXD3 forkhead box D3	AIS1, HFH2, VAMAS2, Genesis	131
OC90	Otoconin 90	PLA2L	150

Table II. Results from qPCR analysis and the Fragment Analyzer

Assay	Efficiency (%) SYBR/Probe	R ² SYBR/Probe	Amplicon length Theoretical/Practical
EYA1	96/96	0.9997/0.9992	209/213
EPHA7	84/89	0.9986/0.9934	144/145
EPHB3	96/97	0.9995/0.9982	124/125
EPHA6	97/94	0.9988/0.9994	145/147
PAX2	89/92	0.9985/0.9989	148/148
ISL1	96/99	0.9984/0.9992	136/142
NKX2-5	95/100	0.9985/0.9993	106/111
GATA4	97/97	0.9995/0.9993	95/95
GATA6	95/101	0.9994/0.9995	80/79
FLII	97/100	0.9989/0.9957	95/98
TAL1	98/	0.9974/not analyzed	89/91
TEK	97/98	0.9986/0.9994	149/154
GATA2	92/93	0.9982/0.9978	121/123
ETV2	96/not analyzed	0.9992/not analyzed	83/88
ACTA2	98/97	0.9990/0.9992	116/123
TAGLN	99/97	0.9986/0.9996	134/139
CD44	98/96	0.9991/0.9995	137/143
ENG	99/100	0.9996/0.9976	140/145
THY1	97/95	0.9992/0.9982	135/140
NT5E	97/100	0.9995/0.9988	116/119
TAT	97/98	0.9993/0.9994	137/143
ALB	96/98	0.9989/0.9994	121/128
AICDA	99/99	0.9993/0.9995	118/121
CCNA2	95/100	0.9993/0.9987	138/137
DPPA2	97/96	0.9995/0.9995	140/143
DPPA3	97/not analyzed	0.9995/not analyzed	137/139
DPPA4	97/100	0.9993/0.9990	117/115
DPPA5	98/not analyzed	0.9992/not analyzed	89/96
NR5A2	97/98	0.9995/0.9995	108/116
NR6A1	98/not analyzed	0.9992/not analyzed	121/125
OTX2	97/100	0.9995/0.9992	107/116
RIF1	99/97	0.9996/0.9987	147/151
SCN1A	99/99	0.9988/0.9984	146/151
SOX15	98/98	0.9993/0.9968	103/111
TCL1A	100/101	0.9993/0.9989	126/130
CDC42	97/96	0.9992/0.9989	192/190
CDK1	95/97	0.9992/0.9992	163/163
HSPA9	102/96	0.9991/0.9990	138/144
MYBL2	101/101	0.9993/0.9995	86/94
ZFP42	98/not analyzed	0.9995/not analyzed	88/95
FGF2	97/97	0.9987/0.99949	97/105
WT1	95/94	0.9983/0.9989	202/198
SIX1	96/97	0.9989/0.9981	79/86
SIX2	95/99	0.9969/0.9991	89/94
SALL1	96/101	0.9958/0.9976	127/134
CITED1	99/95	0.9994/0.9993	198/198
DLK1	100/99	0.9963/0.9948	124/126
LHX1	96/99	0.9987/0.9986	116/118

Table III. Results from the preamplification for SYBR Green

Assay	Efficiency (%) SYBR Green
WT1	100
SIX1	97
SIX2	95
EYA1	74
SALL1	93
CITED1	97
PAX2	98
LHX1	95
DLK1	95
EPHA7	63
EPHB3	60
EPHA6	91
PROM1	98
EPHB2	93
EPCAM	93
VALIDPRIME	89
ACTB	98
B2M	97
GAPDH	95
GUSB	95
HPRT1	96
PPIA	98
RPLP	99
18S	-4
TBP	96
TUBB	95
UBC	100
YWHAZ	97
ISL1	99
NKX2-5	104
GATA4	102
GATA6	104
FLII	98
TAL1	96
TEK	98
GATA2	98
ETV2	96
ACTA2	97
TAGLN	97
CD44	96
ENG	99
THY1	98
NT5E	97
TAT	103
ALB	99
AICDA	78
CCNA2	96
DPPA2	96
DPPA3	95
DPPA4	95
DPPA5	98
NR5A2	97
NR6A1	99
OTX2	101

RIF1	98
SCN1A	99
SOX15	99
TCL1A	100
CDC42	98
CDK1	97
HSPA9	97
MYBL2	101
ZFP42	95
FGF2	93
SOX2	98
LIN28A	97
TDGF1	98
GDF	100
AFP	97
DES	96
DNMT3B	96
POU5F1	97
NANOG	97
FOXD3	89
OC90	133 (not credible)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
A	NTC	cDNA																						
B	NTC	cDNA																						
C	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	gDNA							
D	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	gDNA							
E	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2								
F	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2								
G	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3								
H	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3								
I	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4								
J	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4								
K	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5								
L	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5								
M	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6								
N	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6								
O	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7								
P	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7								

Figure I. Example of plate setup for qPCR analysis with SYBR Green

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
A	NTC																							
B	NTC																							
C	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
D	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
E	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
F	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
G	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
H	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
I	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4
J	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4
K	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5
L	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5
M	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6
N	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6
O	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7
P	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7

Figure II. Example of plate setup for qPCR analysis with probe

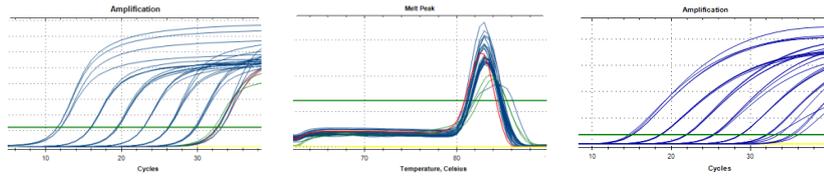
	1	2	3	4	5	6	7	8	9	10	11	12
A	NTC	gBlock	cDNA	gDNA	NTC	gBlock	cDNA	gDNA	NTC	gBlock	cDNA	Ladder
B	gDNA	NTC	gBlock	cDNA	gDNA	NTC	gBlock	cDNA	gDNA	NTC	gBlock	Ladder
C	cDNA	gDNA	NTC	gBlock	cDNA	gDNA	NTC	gBlock	cDNA	gDNA	NTC	Ladder
D	gBlock	cDNA	gDNA	NTC	gBlock	cDNA	gDNA	NTC	gBlock	cDNA	gDNA	Ladder
E	NTC	gBlock	cDNA	gDNA	NTC	gBlock	cDNA	gDNA	NTC	gBlock	cDNA	Ladder
F	gDNA	NTC	gBlock	cDNA	gDNA	NTC	gBlock	cDNA	gDNA	NTC	gBlock	Ladder
G	cDNA	gDNA	NTC	gBlock	cDNA	gDNA	NTC	gBlock	cDNA	gDNA	NTC	Ladder
H	gBlock	cDNA	gDNA	NTC	gBlock	cDNA	gDNA	NTC	gBlock	cDNA	gDNA	Ladder

Figure III. Example of plate setup for analysis with Fragment Analyzer

PreAmp 1-3	A																							
	B																							
	C																							
	D																							
	E																							
	F																							
PreAmp NTC 1-2	G																							
	H																							
	I																							
	J																							
non-PreAmp	K																							
	L																							
	M																							
	N																							
NTC	O																							
	P																							

Figure IV. Example of plate setup for preamplification

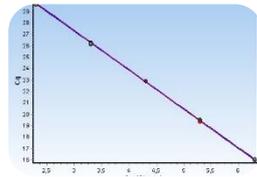
EYA1



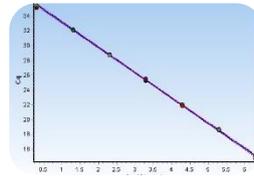
SYBR Green

Melt curve

Probe

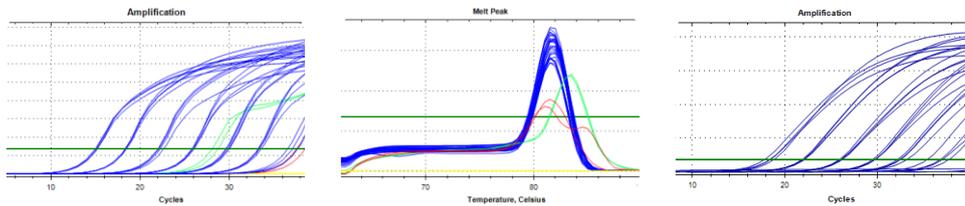


SYBR Green



Probe

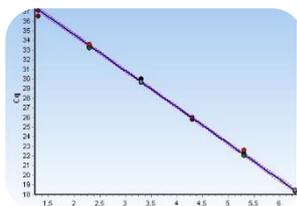
EPHA7



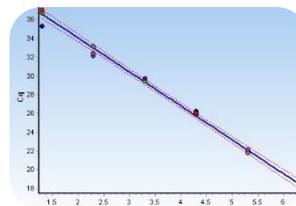
SYBR Green

Melt curve

Probe

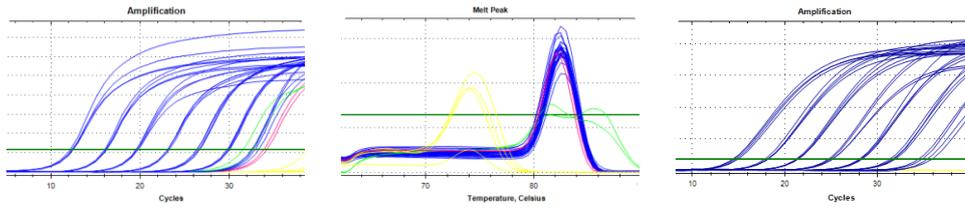


SYBR Green



Probe

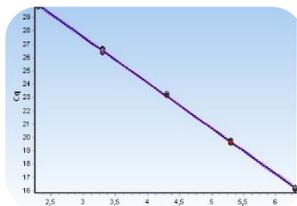
EPHB3



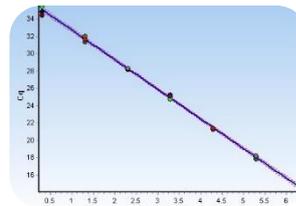
SYBR Green

Melt curve

Probe

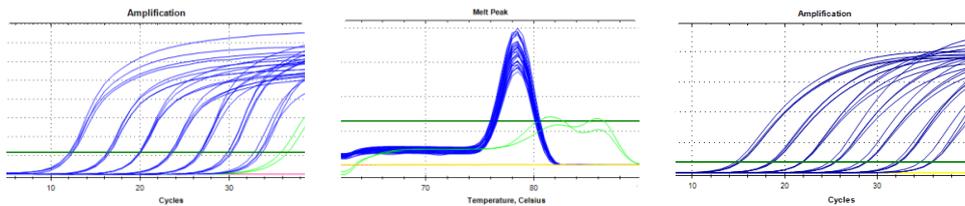


SYBR Green



Probe

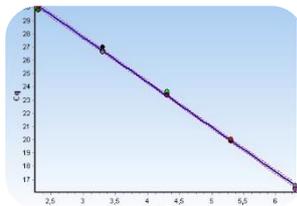
EPHA6



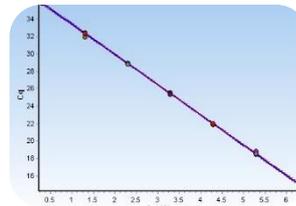
SYBR Green

Melt curve

Probe

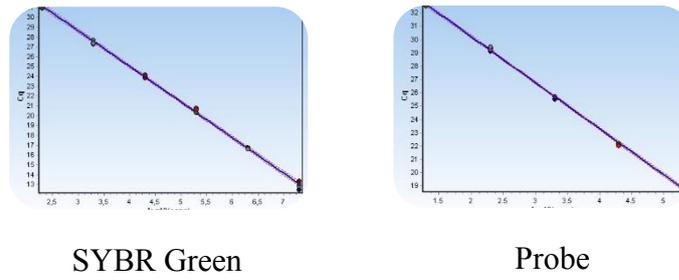
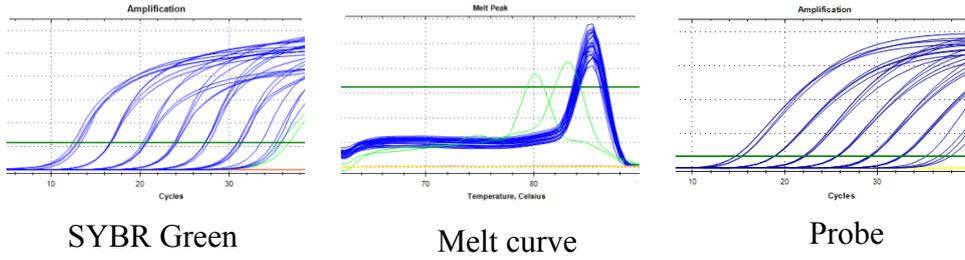


SYBR Green

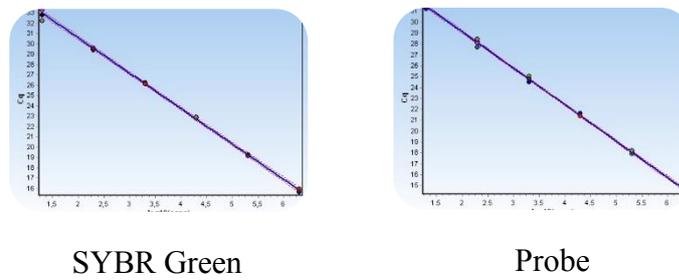
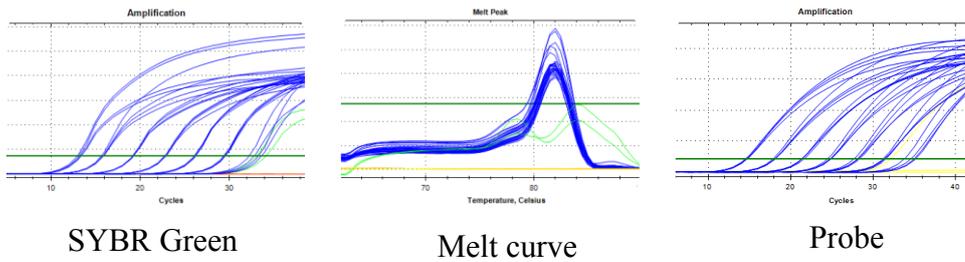


Probe

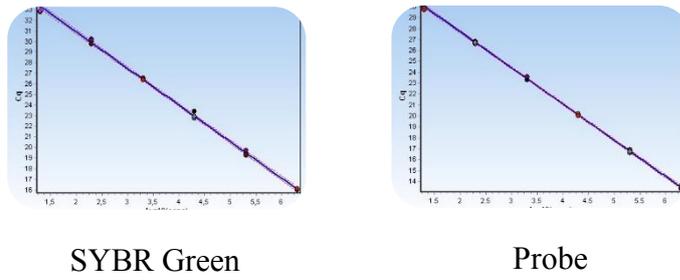
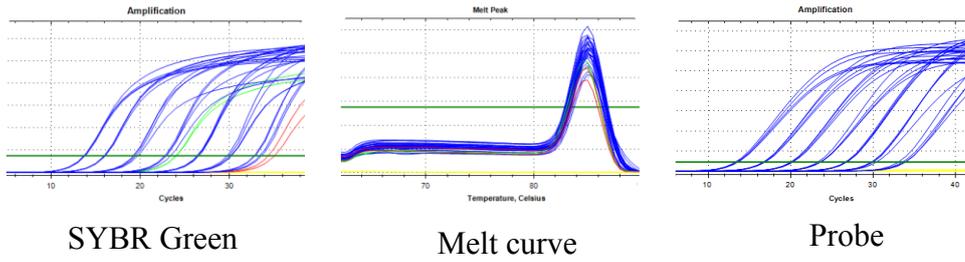
PAX2



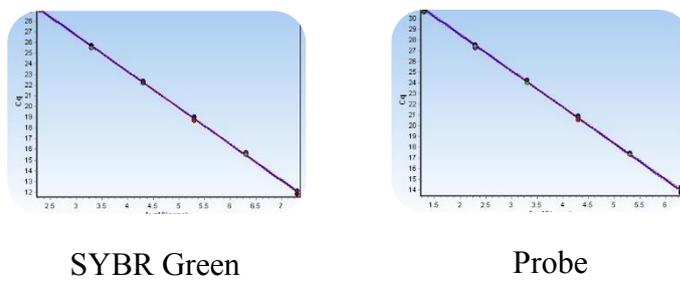
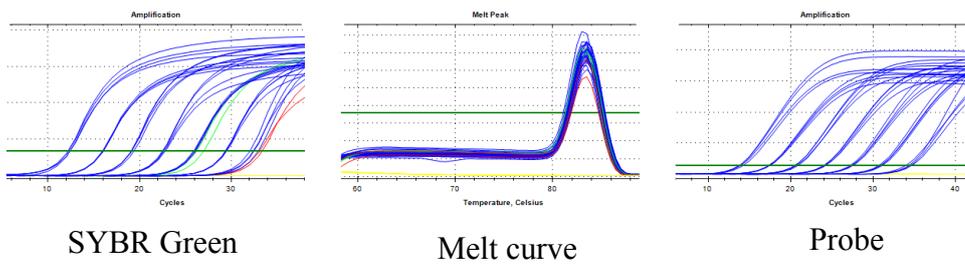
ISL1



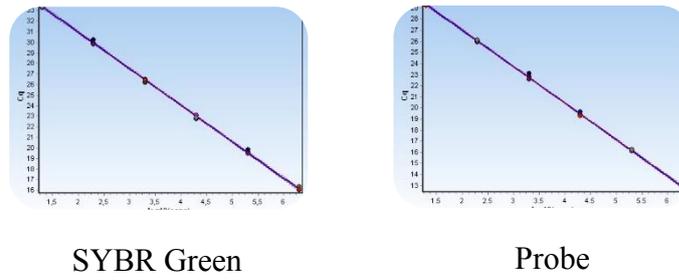
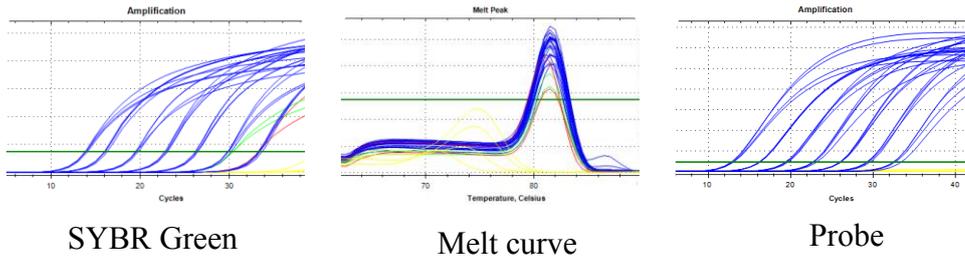
NKX2-5



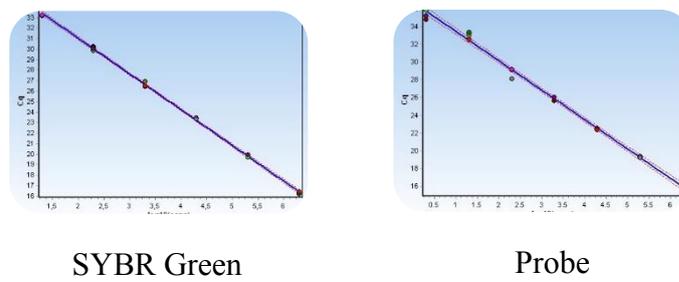
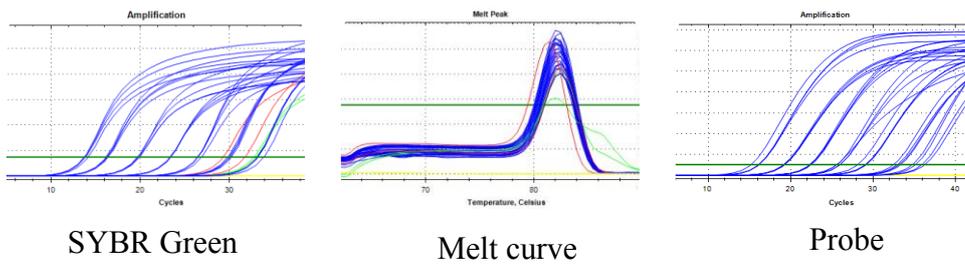
GATA4



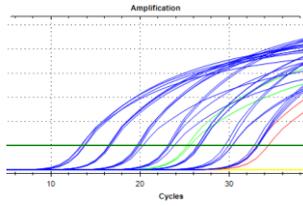
GATA6



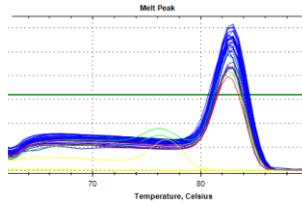
FLII



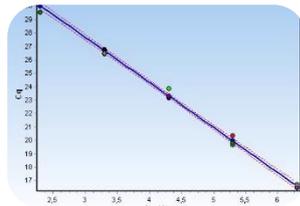
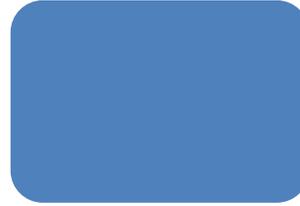
TAL1



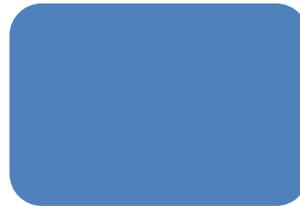
SYBR Green



Melt curve

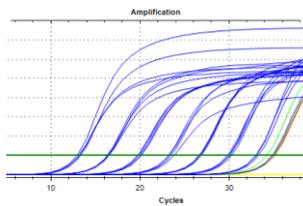


SYBR Green

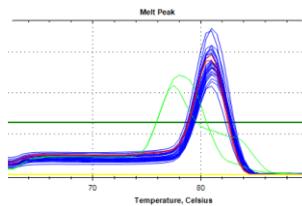


Probe

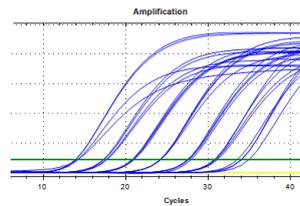
TEK



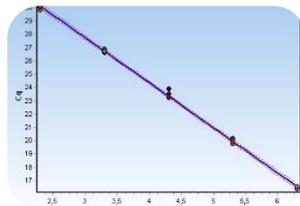
SYBR Green



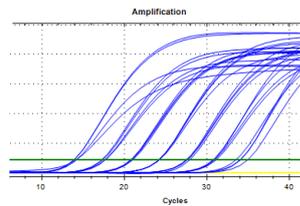
Melt curve



Probe

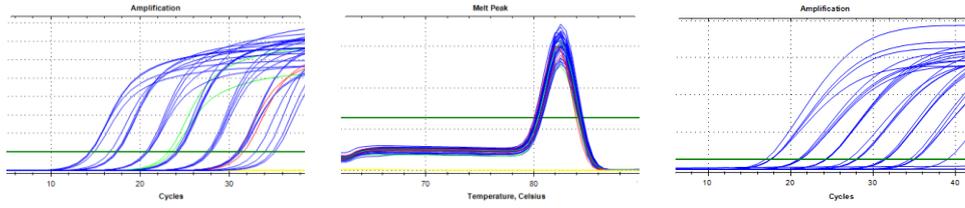


SYBR Green



Probe

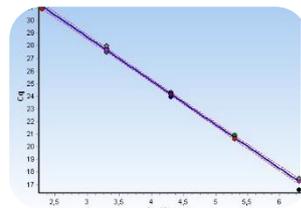
GATA2



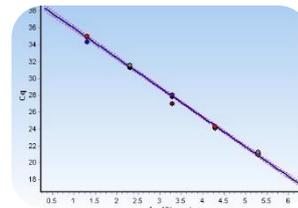
SYBR Green

Melt curve

Probe

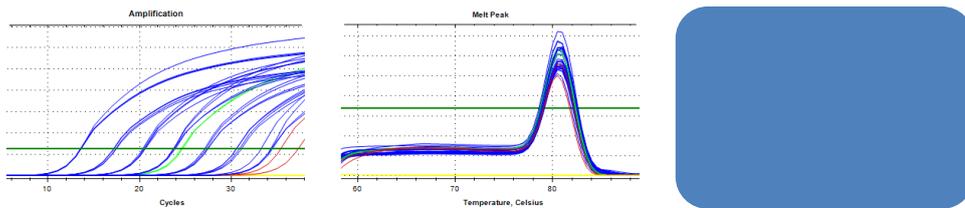


SYBR Green



Probe

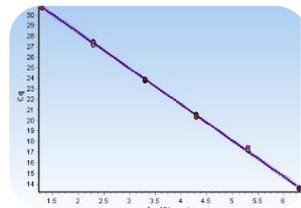
ETV2



SYBR Green

Melt curve

Probe

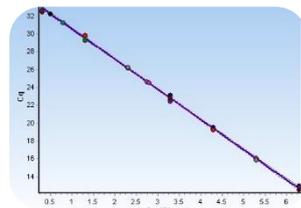
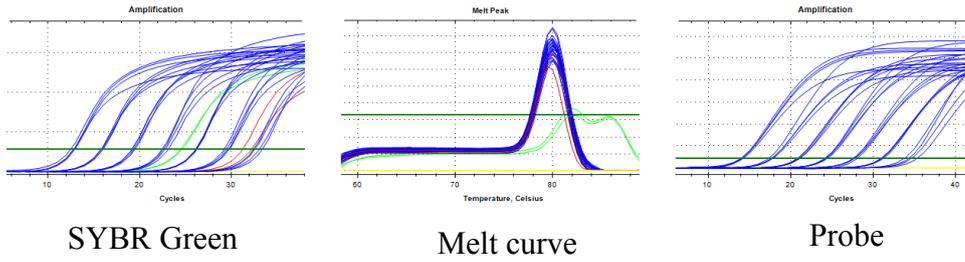


SYBR Green

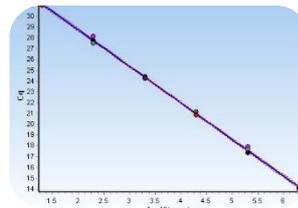


Probe

ACTA2

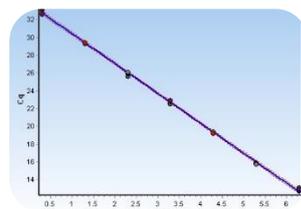
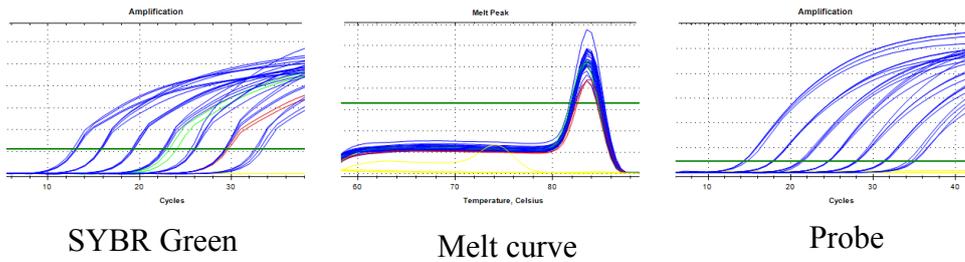


SYBR Green

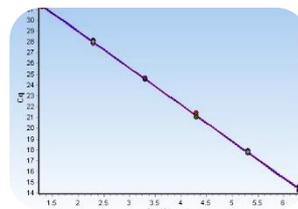


Probe

TAGLN

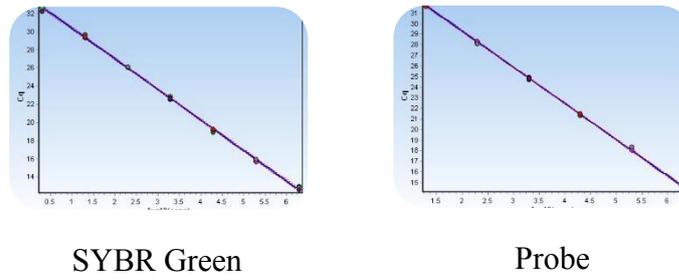
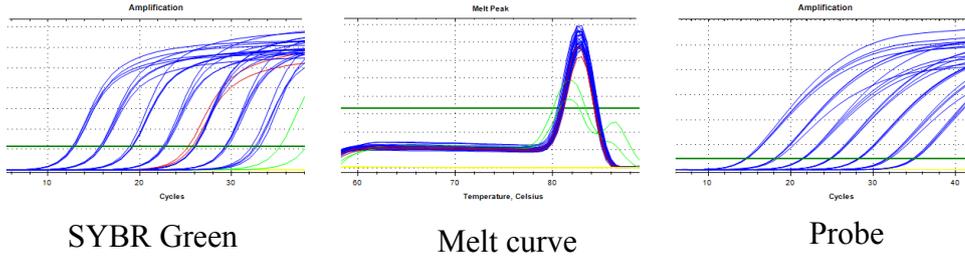


SYBR Green

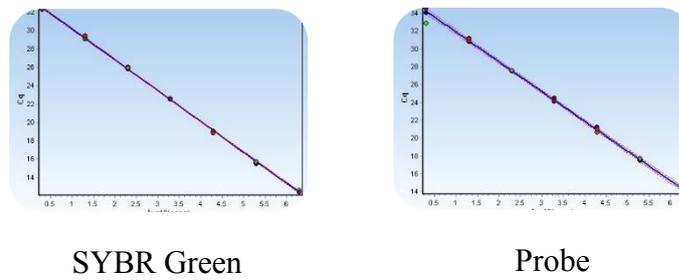
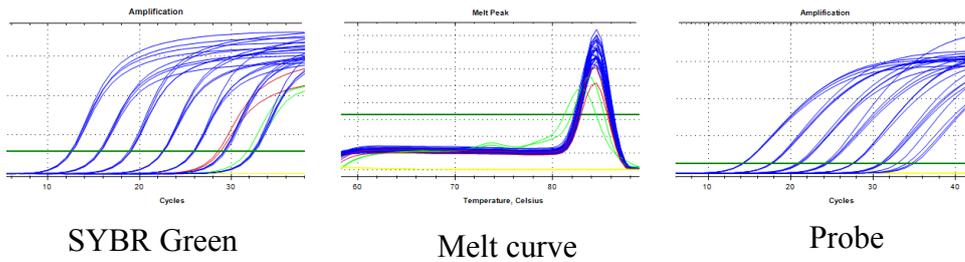


Probe

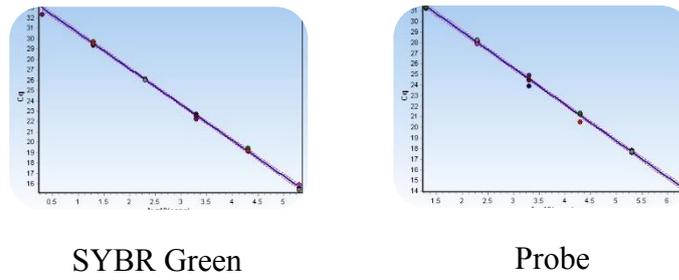
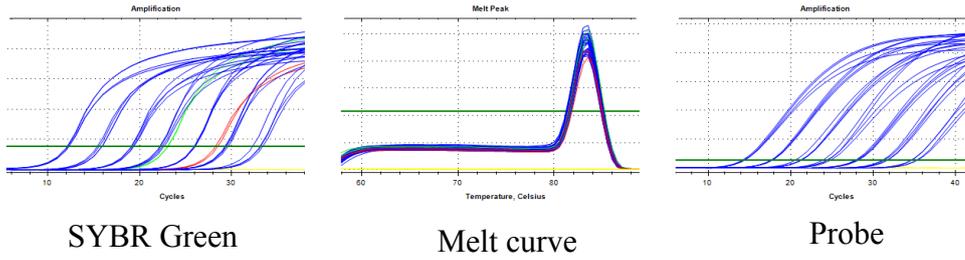
CD44



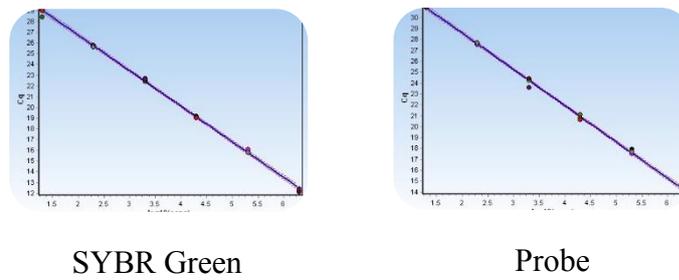
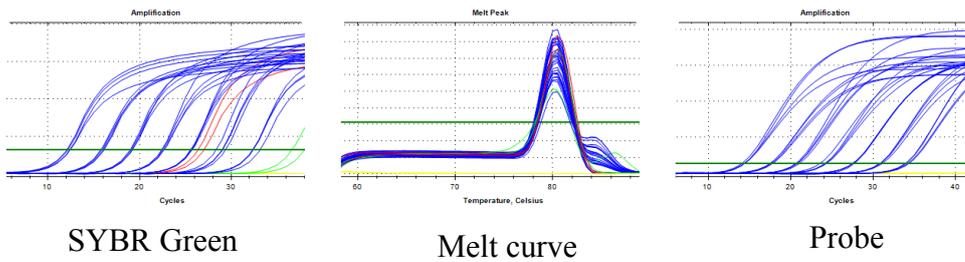
ENG



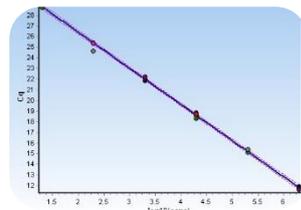
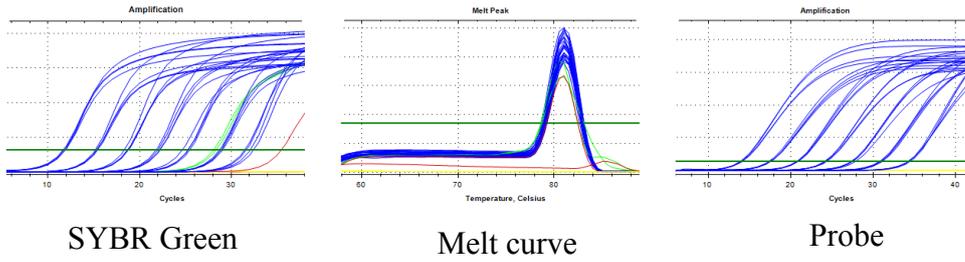
THY1



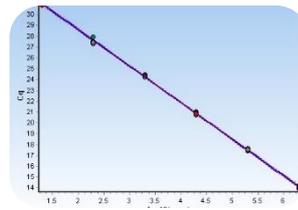
NT5E



TAT

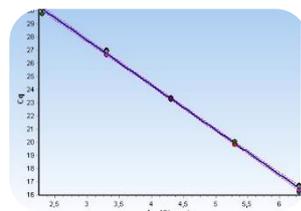
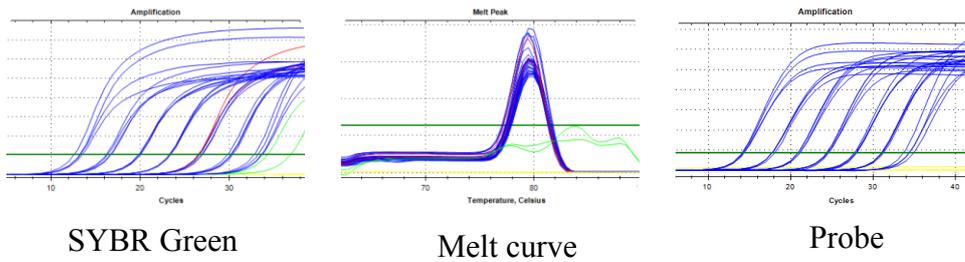


SYBR Green

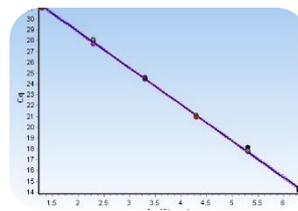


Probe

ALB

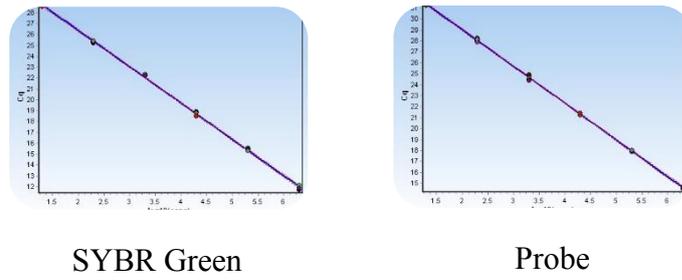
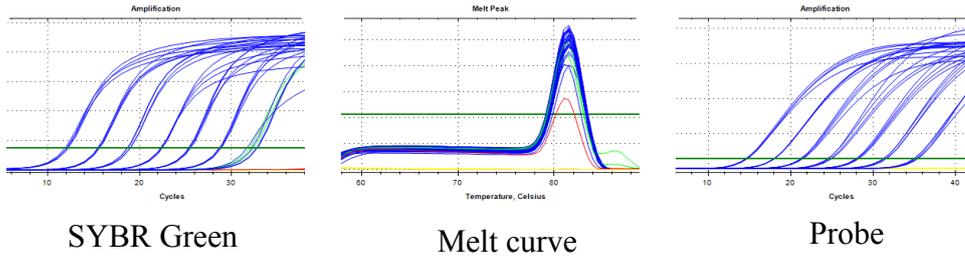


SYBR Green

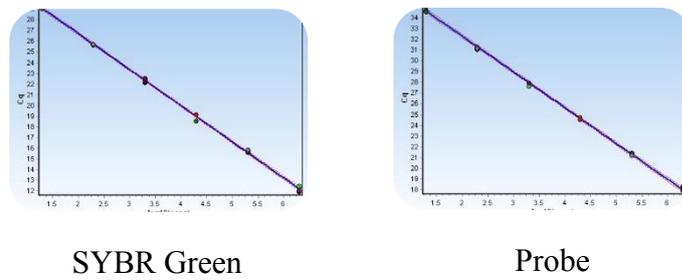
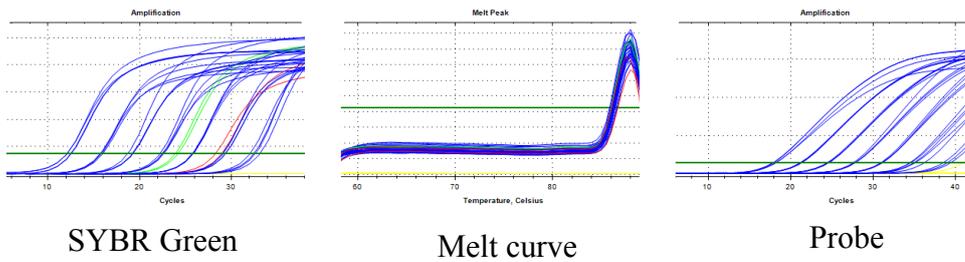


Probe

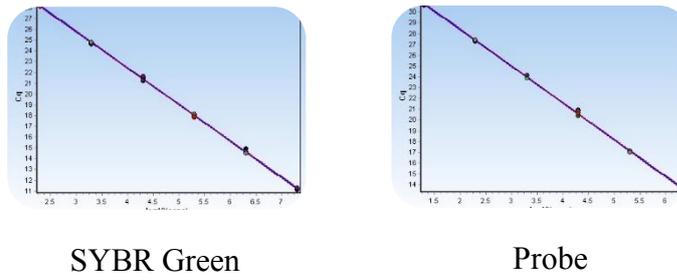
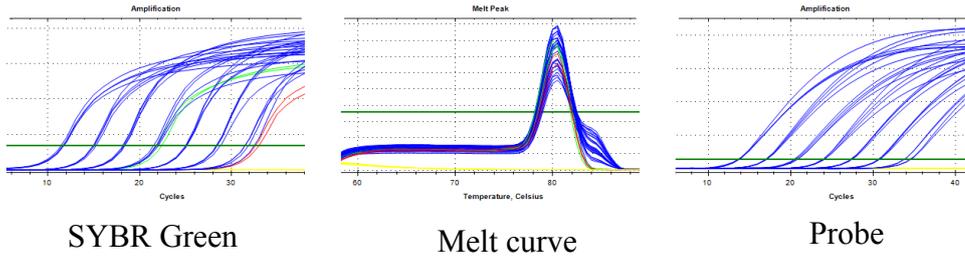
AICDA



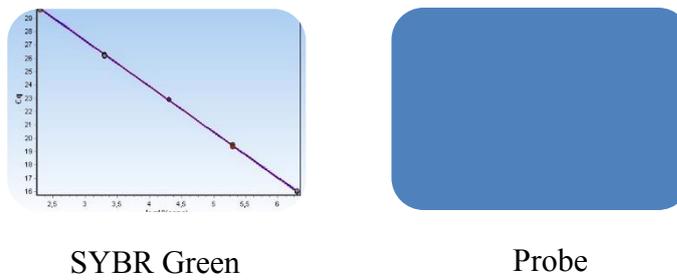
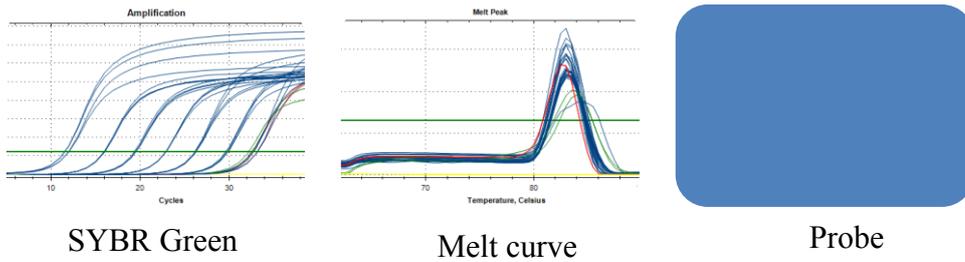
CCNA2



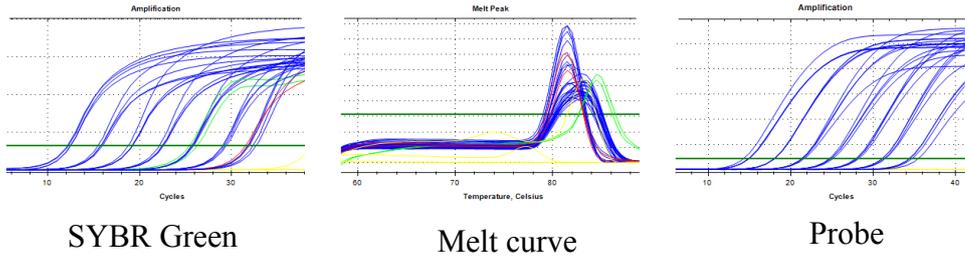
DPPA2



DPPA3



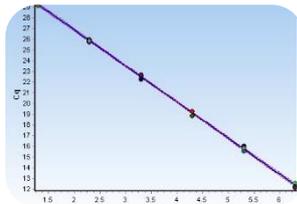
DPPA4



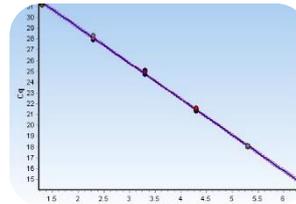
SYBR Green

Melt curve

Probe

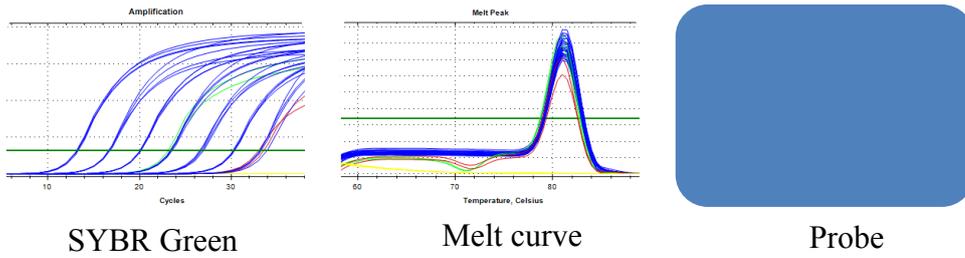


SYBR Green



Probe

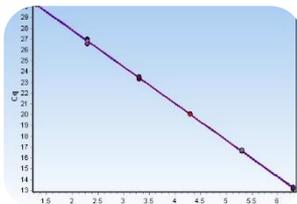
DPPA5



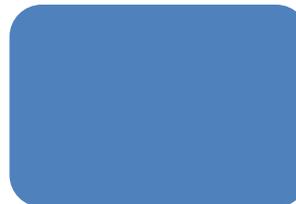
SYBR Green

Melt curve

Probe

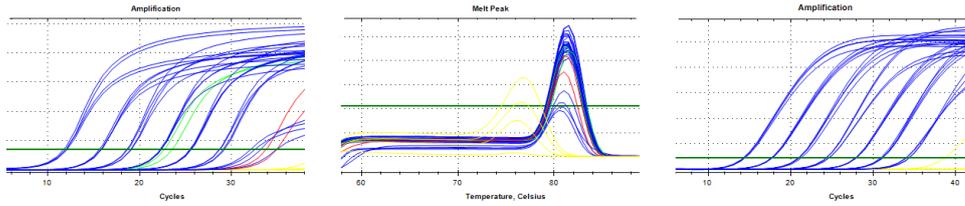


SYBR Green



Probe

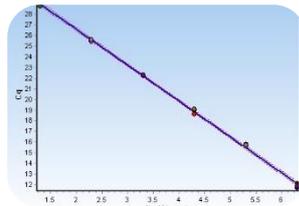
NR5A2



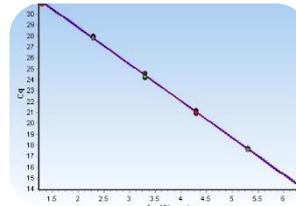
SYBR Green

Melt curve

Probe

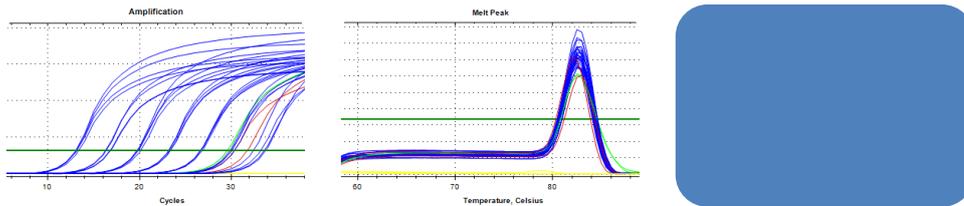


SYBR Green



Probe

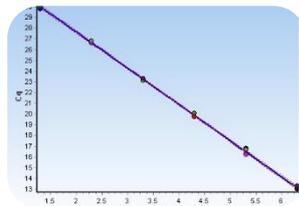
NR6A1



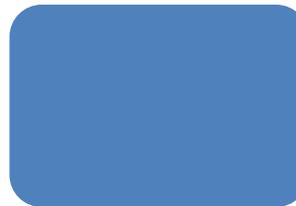
SYBR Green

Melt curve

Probe

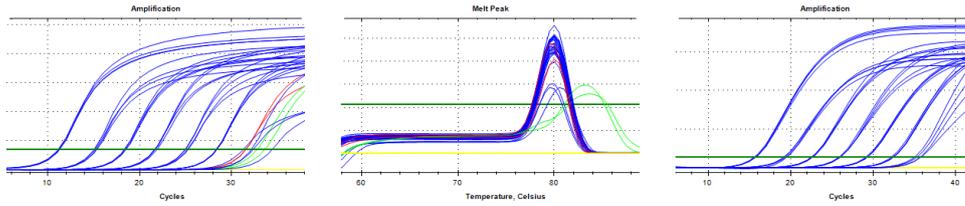


SYBR Green



Probe

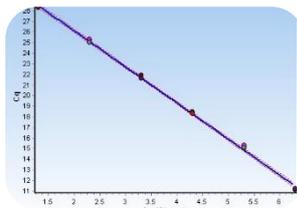
OTX2



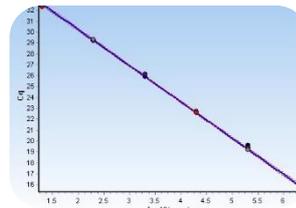
SYBR Green

Melt curve

Probe

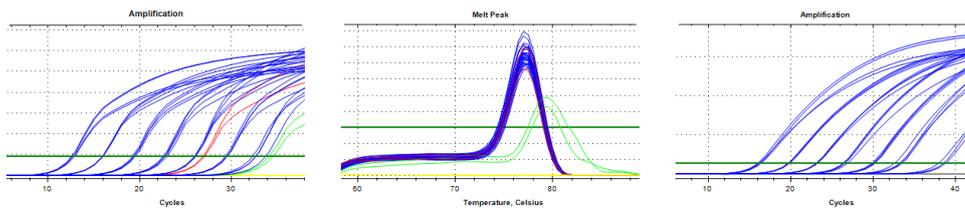


SYBR Green



Probe

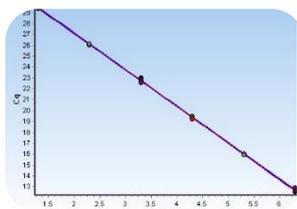
RIF1



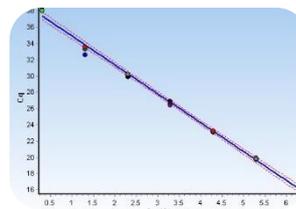
SYBR Green

Melt curve

Probe

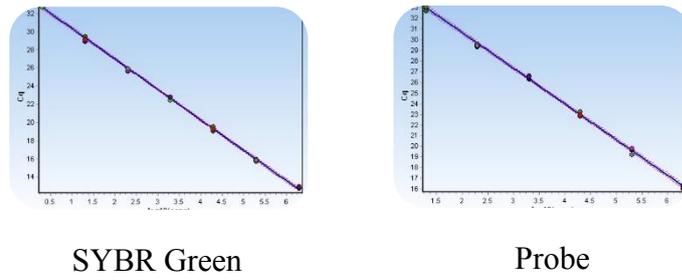
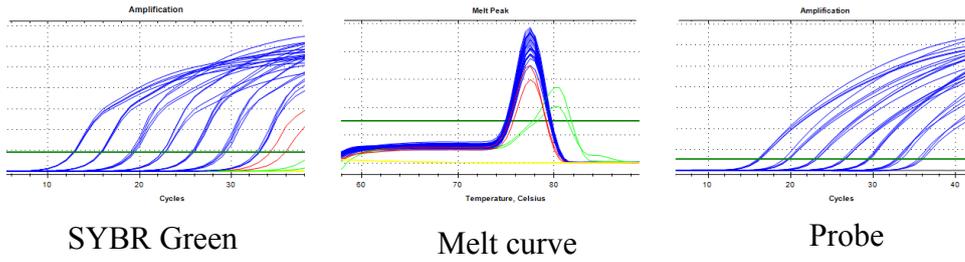


SYBR Green

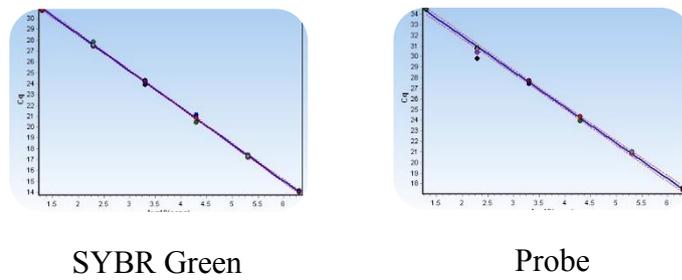
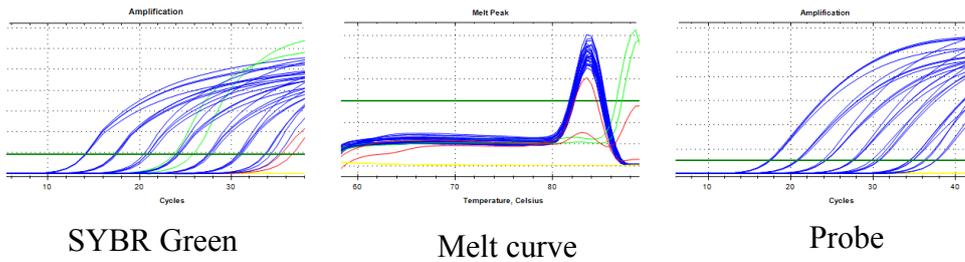


Probe

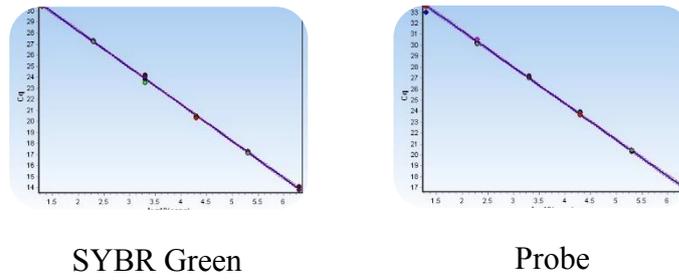
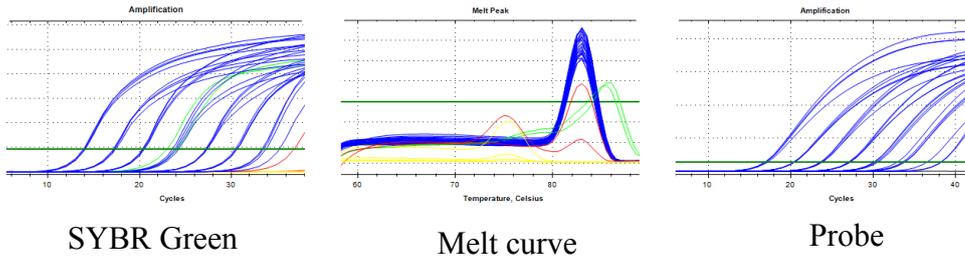
SCN1A



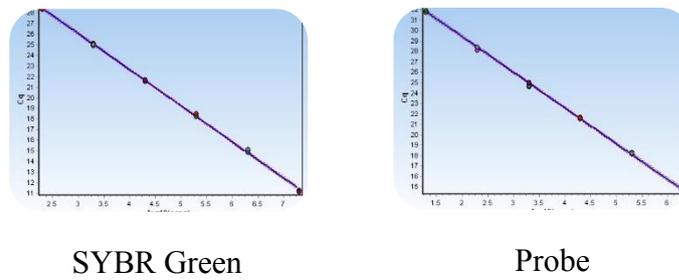
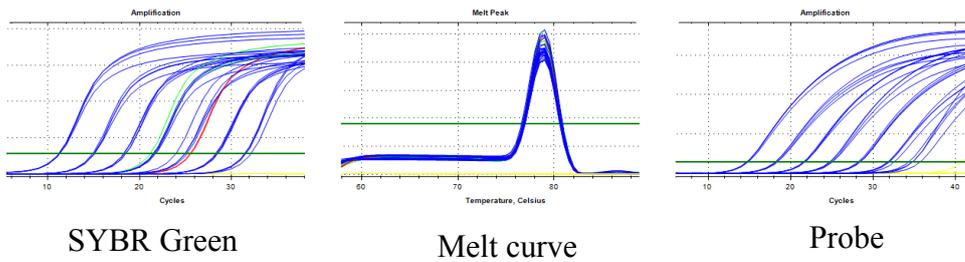
SOX15



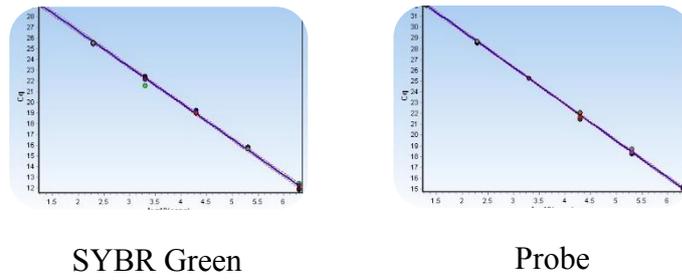
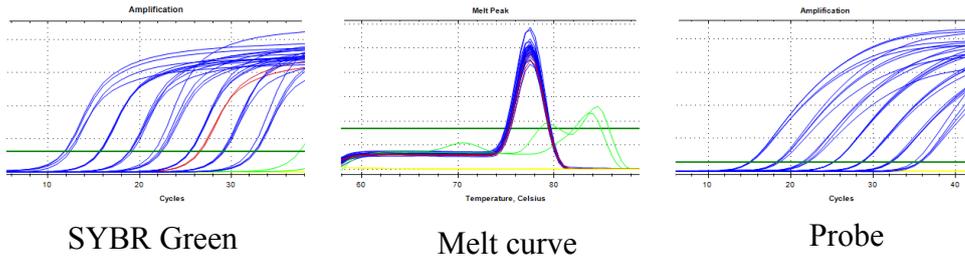
TCL1A



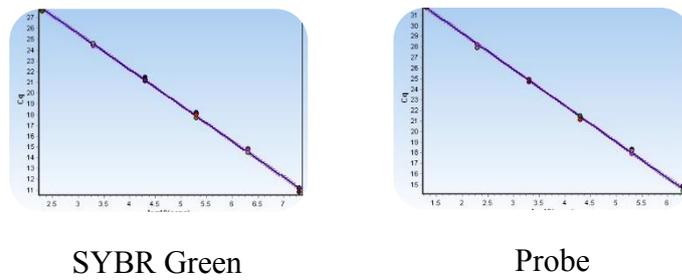
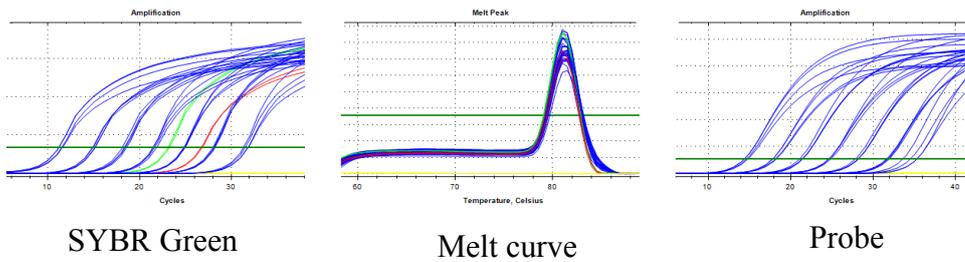
CDC42



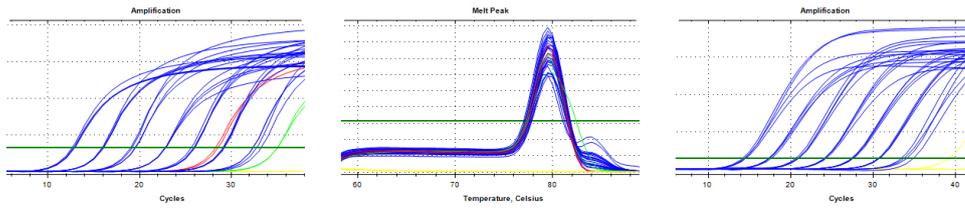
CDK1



HSPA9



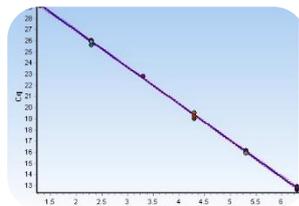
MYBL2



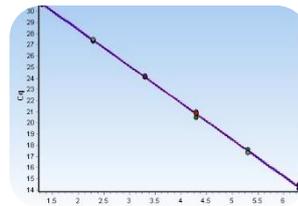
SYBR Green

Melt curve

Probe

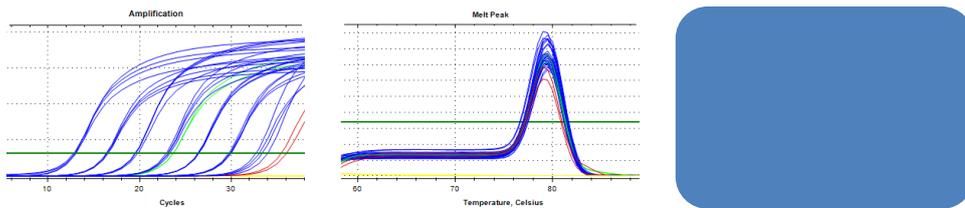


SYBR Green



Probe

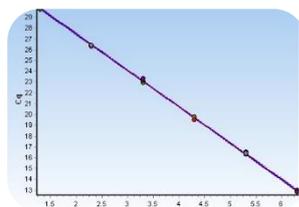
ZFP42



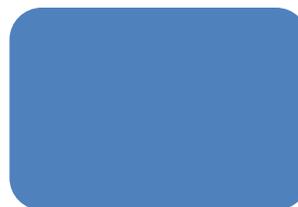
SYBR Green

Melt curve

Probe

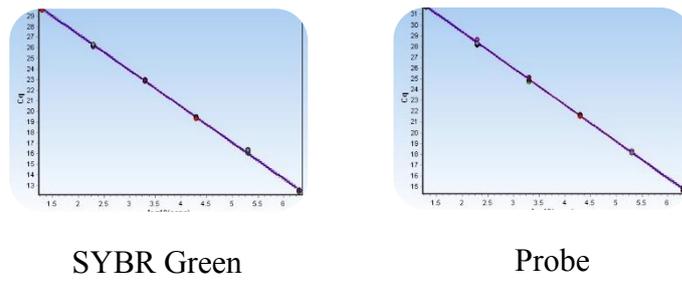
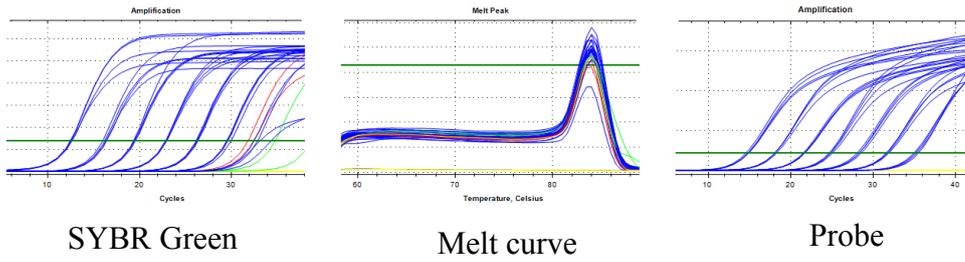


SYBR Green

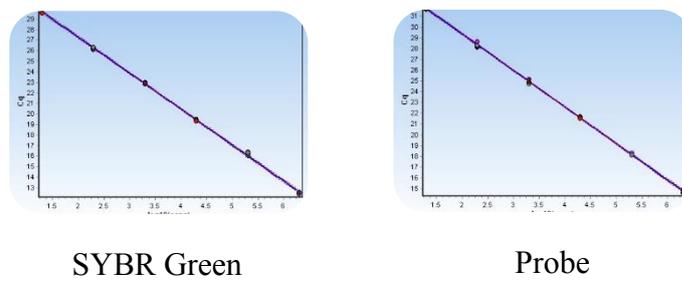
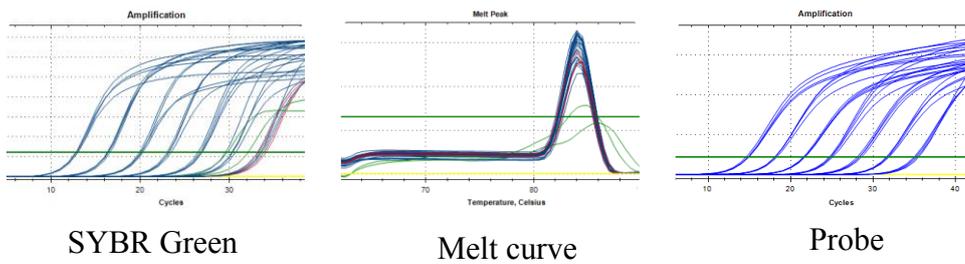


Probe

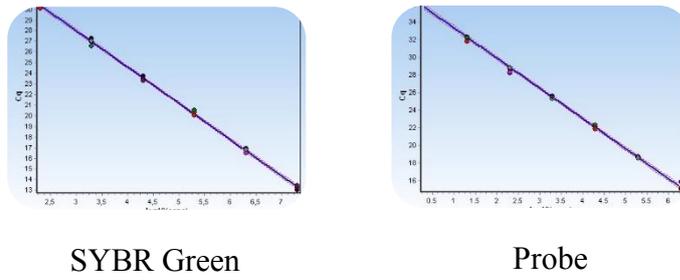
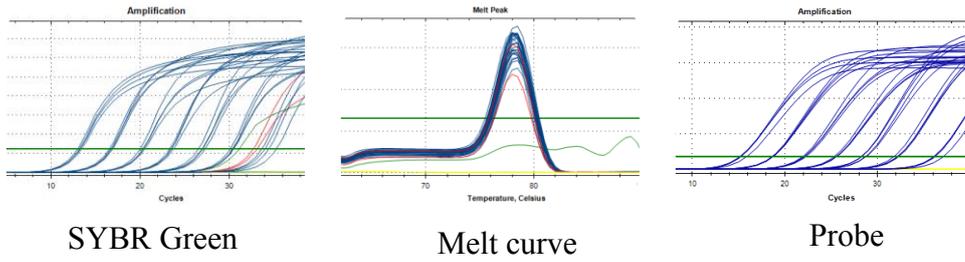
FGF2



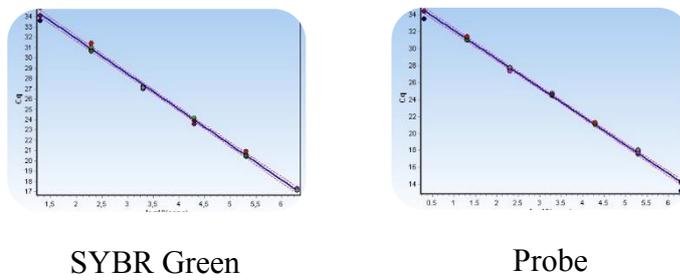
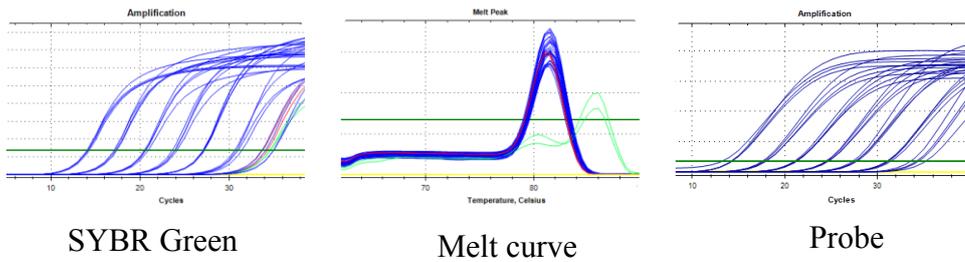
WT1



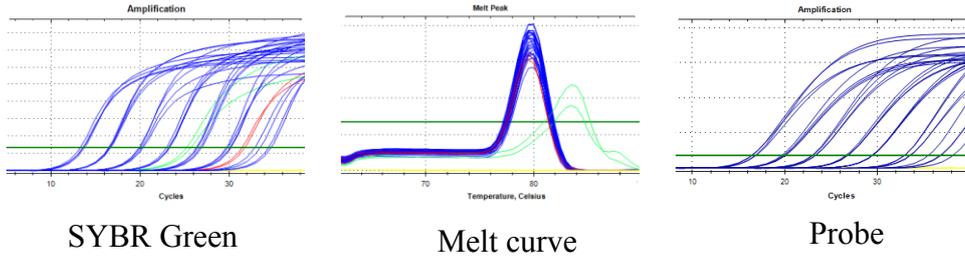
SIX1



SIX2



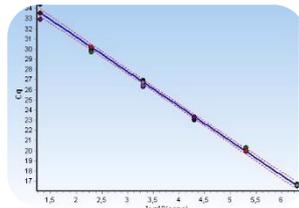
SALL1



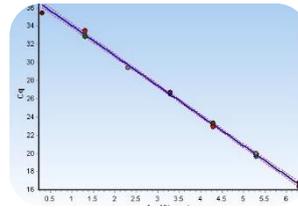
SYBR Green

Melt curve

Probe

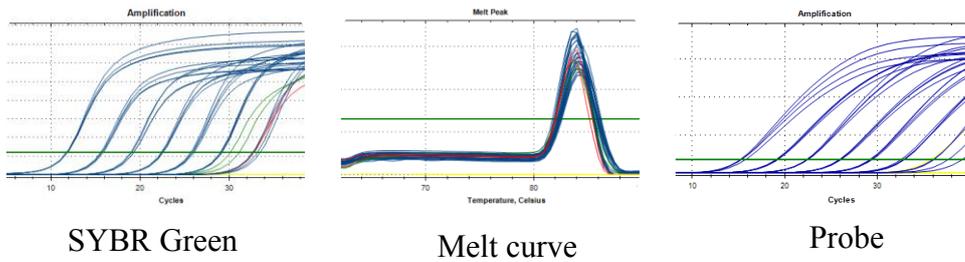


SYBR Green



Probe

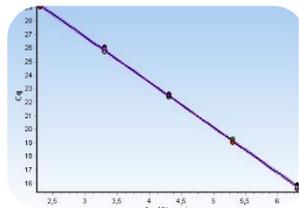
CITED1



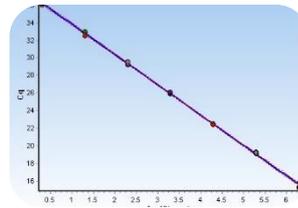
SYBR Green

Melt curve

Probe

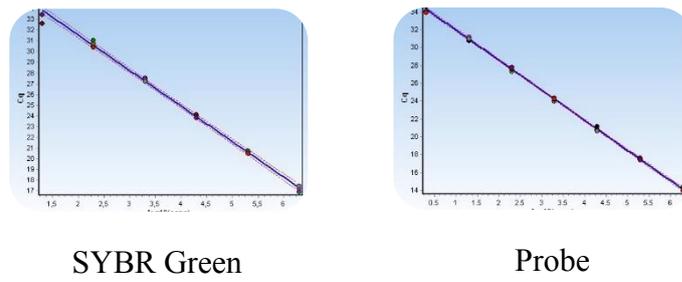
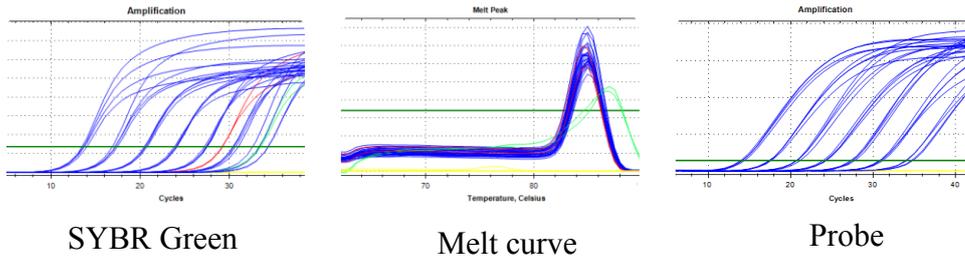


SYBR Green



Probe

DLK1



LHX1

