

Referansegen i qPCR-analyser av humane endotelceller fra koronararterier (HCAEC)

MED-3950

5.-årsoppgaven

Profesjonsstudiet i medisin ved Uniersitetet i Tromsø

Stine Skoglund Machiedo

MK-09

Tromsø, mai 2014

Forord

I 1969 beskrev Thomas D. Brock og Hudson Freeze en termofil bakterie som de ga navnet *Thermus aquaticus*. Bakterien hadde de funnet i en varm kilde i Yellowstone nasjonalpark. Den trivdes ved ca 70°C - noe som var en sensasjon. Oppdagelsen av bakterien ga oss tilgang til varmemestabile biologiske enzymer, deriblant DNA-polymerase som katalyserer dannelsen av DNA.

I 1970 oppdaget Howard Temin og David Baltimore et enzym fra et retrovirus som kunne lage DNA-kopier av RNA. Enzymet ble kalt revers transkriptase, og dette fikk de en Nobelpris for i 1975.

I 1971 beskrev nordmannen Kjell Kleppe prinsippet for polymerase chain reaction (PCR) sammen med Har Gobind Khorana. PCR er en metode for å lage mange kopier av DNA. Kary B. Mullis og medarbeidere videreutviklet PCR på 80-tallet, og fikk Nobelpris for dette i 1993.

Det var oppdagelsen av den varmemestabile DNA-polymerasen som åpnet for å syntetisere DNA i et laboratorium, noe bare levende organismer hadde kunnet gjøre inntil da. Revers transkriptase muliggjorde oppkopiering av RNA ved å «oversette» det til DNA.

I 1992 koplet Higuchi og kollegaer et fluorescerende fargestoff til PCR. Dette gjorde at man kunne monitorere reaksjonen mens den pågikk (real-time PCR, nå kjent som qPCR). Dette åpnet døren for å kvantifisere DNA og RNA.

I dag er qPCR en av de mest brukte metodene i molekylærbiologi. Det brukes i alt fra matvarekontroll til forskning.

I et studie som gjennomføres i forbindelse forskerlinjen ved medisinstudiet ved UiT, ser jeg på forskjeller i genuttrykk hos koronare endotelceller når de utsettes for lavt nivå av oksygen og inflammasjon sammenliknet med en kontroll. Genuttrykk bestemmes ved hjelp av qPCR. Denne oppgaven har som formål å bestemme hvilke referansegene som er best egnet for mitt studie ut fra et kommersielt tilgjengelig referansegenpanel.

Jeg vil med dette rette en takk til ansatte ved gastrokirurgisk..... ved UiT for praktisk tilretteleggelse, kontorplass og hjelp til tusenvis av pipetteringer og samt merking av uendelige mengder plastrør! En spesiell takk til Trine Kalstad!

Veileder for denne oppgaven, Trine Lund ved Kardiovaskulær Forskningsgruppe, IMB, UiT fortjener en stor takk for teoretisk påfyll og rettleiding. Det har heller ikke vært mangel på hjelp til løsninger av praktiske problemstillinger. Hovedveileder for min forskerlinjeoppgave, Truls Myrmedal, takkes i denne omgang særdeles for å ha sendt meg på et praktisk kurs i qPCR - til stor inspirasjon (med påfølgende frustrasjon) og nytte!

Tromsø, mai 2014. Stine Skoglund Machiedo.

Sammendrag

Denne oppgaven er en del av et forskerlinjeprosjekt som ser på forskjeller i genuttrykk hos koronare endotelceller utsatt for lavt nivå av oksygen og inflammasjon sammenliknet med kontroller. Et av forsøkene hadde som hensikt å se utvikling i genuttrykk etter ulik behandlingstid.

Denne oppgaven skal bestemme hvilke referansegene som er best egnet for dette forsøkssettet ved qPCR ved hjelp av de statistiske hjelpemidlene geNorm og NormFinder.

Forsøkene ble gjort på humane endotelceller fra koronare arterier (HCAEC) med kontroller og TNF- α -stimulerte celler ved 20% og 2% O₂ i tidsrommene 1, 2, 4, 6, 8 og 12 timer, n = 6. Et humant referansegenpanel med 12 gen fra TATAA Biocenter ble benyttet. Forsøkene ble utført ved bruk av humane endotelceller fra koronare arterier (HCAEC). Forsøksoppsettet bestod av fire hovedgrupper: 1) Kontroller normoksi (20% O₂); 2) Kontroll hypoksi (2% O₂); 3) TNF- α -stimulerte celler ved normoksi (20% O₂); 4) TNF- α stimulerte celler ved hypoksi. Gruppene ble inkubert ved i førnevnte forholdene i 1, 2, 4, 6, 8 og 12 timer, n = 6. Totalt utgjorde dette oppsettet 24 ulike behandlingsgrupper. Det er derfor interessant å kartlegge hvilke kjente referansegene som er stabilt uttrykt under disse forholdene.

Et utvalg på 72 av totalt 144 prøver ble tatt med i qPCR, tilsvarende 3n av de 24 gruppene. Et humant referansegenpanel med 12 gen fra TATAA Biocenter ble benyttet. Etter qPCR ble 6 gen ekskludert fra studien grunnet dårlig effektivitet, dannelse av primer dimerer og usikker spesifisitet på PCR-produkt. GenEx fra MultiD Analyses AB ble benyttet til databehandlingen. geNorm og NormFinder er innbakt i denne softwaren. Begge metodene rangerte genene likt med RPLP og YWHAZ som de beste kandidatene. Det bør vurderes om ytterligere arbeid med å optimalisere qPCR for de ekskluderte kandidatene skal utføres, slik at de kan inkluderes i en ny analyse. I mellomtiden ansees RPLP og YWHAZ som gode nok referansegene til dette formålet.

Innhold

Innledning	1
Bakgrunn	1
Kort om qPCR	1
Normalisering	3
Referanseggen	3
geNorm	4
NormFinder	4
Materiale og metode	5
Forsøk	5
Isolering og rensing av RNA	5
Kvantifisering og kvalitetskontroll av RNA	6
Revers transkripsjon (RT)	6
No-reverse transcription-kontroll (NoRT)	6
Referanseggen-panel	7
qPCR	7
Smeltekurve	8
SYBR Green	8
Primer-dimer	8
Effektivitet og fortynningskurve	8
Databehandling	9
Arbeidsprosessen	9
Resultater	11
Diskusjon	14
Konklusjon	17
Referanser	18
Vedlegg 1	20

Innledning

Bakgrunn

I forskerlinjeprosjektet mitt ser jeg på forskjeller i genuttrykk hos koronare endotelceller utsatt for lavt nivå av oksygen og inflammasjon sammenliknet med kontroller. Genuttrykk bestemmes ved hjelp av qPCR (kvantitativ real-time polymerase chain reaction). Referansegene benyttes i qPCR for å korrigere for teknisk variasjon. Referansegene må bestemmes for hvert unike studie. Et av forsøkene hadde som hensikt å se utvikling i genuttrykk etter ulik behandlingstid. Denne oppgaven skal bestemme hvilke referansegene som er best egnet for dette forsøkssettet ut fra et kommersielt tilgjengelig referansegenepanel (TATAA Biocenter??).

Kort om qPCR

qPCR er den mest brukte metoden til å påvise og kvantifisere små mengder nukleinsyrer. Det brukes i alt fra matkontroll til medisinsk diagnostikk og forskning (1). I et typisk biologiskstudie vil man bruke qPCR til å påvise en forskjell i genuttrykk mellom to ulike biologiske grupper som har vært utsatt for ulik behandling(2).

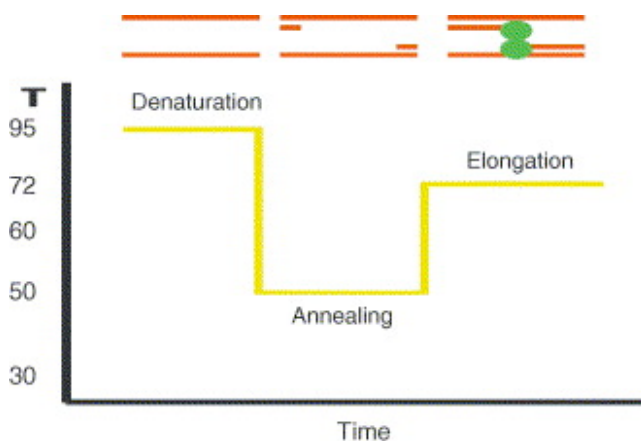


Fig. 1. Temperatursyklus PCR: [1] temperatur ved ca. 95°C smelter dobbeltrådet DNA slik at de to trådene separeres, [2] primerene fester seg til templatet når temperaturen senkes, [3] temperatur ved ca 72°C tillater DNA-polymerasen å forlenge primerene (3).

I prinsippet er det enkelt å utføre qPCR. Foruten prøvemateriale trenger man to primere, DNA-byggesteiner (dNTP), en varmestabil DNA-polymerase, et fluorescerende fargestoff og en qPCR-maskin. Primerene er korte nukleinsyre kjeder (18-20 nt) som binder seg spesifikt i hver sin ende på hver sin komplementære DNA-tråd av genet man vil påvise. Deretter bygger DNA-polymerasen opp en ny DNA-tråd ved å forlenge primerne, slik at man ender opp med to kopier av genet. Det hele foregår ved hjelp av en temperatursyklus. Høy temperatur smelter DNAet slik at de to komplementære trådene skiller lag. Ved å senke temperaturen til ca 60°C tillater man primerene å bindes til DNA-trådene. Til slutt justeres temperaturen til 72°C som er temperaturen der DNA-polymerasen jobber optimalt. Ved å gjenta temperatursyklusen, kan flere kopier (PCR-produkt) dannes (3).

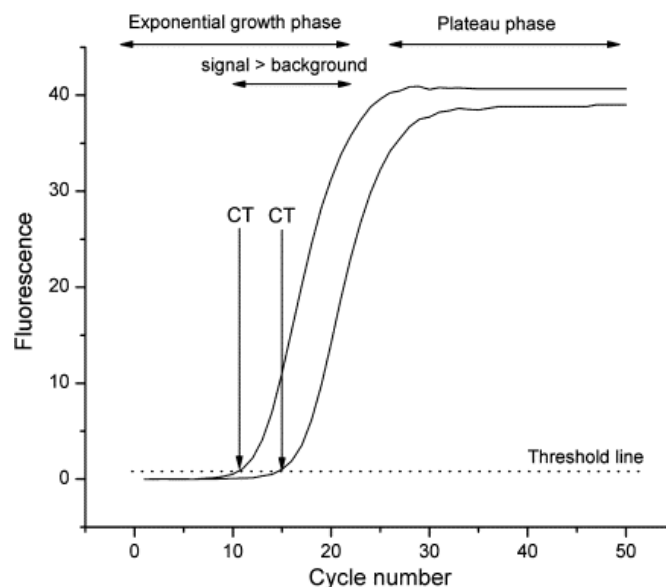


Fig. 2. qPCR-kurve. Fluorescenssignalet øker eksponensielt. En grenseverdi (Threshold line) settes over bakgrunnssignalet. Cq (CT på figuren) er punktet der et signal krysser grenseverdien. Den eksakte Cq-verdien er ofte ikke interessant i seg selv. Hvor grenseverdien settes er av liten betydning så lenge den settes i den eksponentielle fasen og den samme grenseverdien benyttes til alle prøver som skal sammenliknes(3).

Det fluorescerende fargestoffet (SYBR green) binder seg kort fortalt til dobbelttrådet DNA, endrer konformasjon og sender ut et fluorescerende signal som kan detekteres. Dette signalet leses av ved slutten av hver syklus. Fluorescensen er proporsjonal med startkonsentrasjonen av det aktuelle genet, og vil øke eksponentielt for hver syklus inntil reaksjonen begrenses av mangel på reagenser og opphopning av produkt etc (platåfase). Antall sykluser som behøves for å nå et bestemt

fluorescensenivå - en grenseverdi - kalles Cq-verdi. Prøver med ulik startkonsentrasjon av et bestemt gen vil gi ulike Cq-verdier (3).

I et typisk studie av genuttrykk er man ikke interessert i den eksakte konsentrasjonen (eller Cq-verdi) av et gen, men heller differansen mellom to prøver.

Selv om qPCR i prinsippet er enkelt, er mange trinn involvert: Isolering av RNA fra prøvematerialet, revers transkripsjon av RNA til cDNA og amplifisering av cDNA ved hjelp av qPCR. Alle disse trinnene bidrar på hver sin måte til (teknisk betinget) variasjon i analyseresultatet som kan dekke over den reelle, biologiske variasjonen (1,2,3,4).

Normalisering

Kilder til variasjon er blant annet ulik startmengde av prøvematerialet, ulikt utbytte av RNA-isolering og ulik effektivitet av omdanning fra RNA til cDNA. I tillegg kan prøvene ha ulik grad av forurensing og ulik RNA-kvalitet. For å øke analysens presisjon, må man normalisere. Oppsummert av Kubista med medarbeidere (årstall) er det vanlig å basere normaliseringen på fysiske parametere (masse, volum, størrelse, celleantall), mengde RNA (totalt, ribosomalt, eksternt tilsatt) og ikke minst mot referanseggen (3).

Referanseggen

Et referanseggen er per definisjon et gen som er konstant uttrykt i alle prøvene som er inkludert i en studie. Det analyseres i tillegg til målgenet på alle prøver. Variasjon i prøvematerialet elimineres ved å kalkulere differansen i Cq-verdi mellom målgen og referanseggen (ΔCq) for hver prøve (2). Denne differansen kan igjen brukes direkte for å sammenlikne det relative uttrykket av målgenet mellom ulike prøver ($-\Delta\Delta Cq$). Det forutsettes at referanseggen ikke ko-reguleres med målgen eller med hverandre. Det finnes ingen universelle referanseggen. Dette må valideres for hvert enkelt studie (5,6). Det er vanlig å teste 10-12 ulike gen. Disse bør representere ulike biologiske funksjoner i en celle, samt være uttrykt på ulike nivå (høyt, lavt). Det er enighet om at normalisering bør skje mot 2-3 referanseggen (4). Flere metoder for å beregne hvilke gen som er mest stabilt uttrykt i et referanseggenpanel er utviklet. To av de mest brukte i dag er geNorm og NormFinder (4).

geNorm

geNorm sammenlikner det relative genuttrykket til referansegenene parvis i de ulike prøvene involvert i studien. For hver sammenlikning, elimineres det genet som viser størst variasjon i uttryksprofilen i prøveutvalget, sammenliknet med de andre genene. Til slutt sitter man igjen med to gener som er best egnet som referansegen til den aktuelle studien (7). Det kalkuleres en M-verdi som bilde på variasjonen. Genet som får den høyeste M-verdien blir eliminert for hver sammenlikning.

NormFinder

NormFinder er basert på ANOVA, og ser på den samlede variasjonen i genuttrykk mellom de aktuelle referansegenene. I tillegg kan man estimere variasjonen i og mellom grupper i prøveutvalget (behandlet/kontroll, syk/frisk).(8).

Materiale og metode

Forsøk

Forsøkene ble gjort på humane endotelceller fra koronare arterier (HCAEC) fra antatt frisk donor kommersielt tilgjengelig fra Lonza. Cellene ble oppkultivert i 75 cm² cellekulturflasker med vekstmediet EGM-2 (Lonza) tilsatt 10% fetal bovine serum (FBS) fra i inkubator ved 37°C, 20% O₂, 5% CO₂ og mettet luftfuktighet (normalforhold). Ved konfluent oppvekst, ble cellene splittet og oppkultivert videre slik at man ved forsøkets start hadde 3 ulike passasjer (3n) til rådighet. Oppkultiveringen tok ca. 7 uker. Forsøket ble utført på 12 stk. 12-brønnsplater (vekstareal 3,8 cm²). På hver plate ble det sådd ut 2 brønner fra hver n , til sammen $2 \times 3n = 6$ prøver. Hver brønn inneholdt 150 000 celler, talt med automatisk celleteller (Millipore, 60 µm tellespiss). Cellene på platene ble inkubert ca. 16 timer under samme forhold som tidligere nevnt. Deretter 4 timer med serumsupresjon i cellemediet EBM-2 (Lonza). Under forsøket ble EBM-2 tilsatt 2% FBS benyttet. 1 prøve ble beholdt som kontroll og 1 prøve ble stimulert med 1 ng/ml TNF-α (Sigma) for å indusere en inflammasjonstilstand. Halvparten av platene ble så inkubert under normalforhold. Den andre halvparten ble inkubert ved samme temperatur og luftfuktighet, men med 2% O₂ (hypoksi). Inkuberingstiden var henholdsvis 1, 2, 4, 6, 8 og 12 timer. Prøvene representerte således en tidskurve for stimulerte- og kontrollceller ved normale og hypoksiske forhold (24 grupper). Det ble gjort to forsøksserier, tilsammen 6n. Dette utgjorde 144 enkeltprøver. Ved endt inkubasjonstid, ble cellene vasket og lysert med 1,0 ml lyseringsbuffer (fra RNA-isolerings kit). Lysatet ble umiddelbart satt på is og deretter frosset ned til -70°C.

Isolering og rensing av RNA

RNA ble isolert fra lysatet og rensert ved hjelp av et kommersielt tilgjengelig kolonnebasert kit (PerfectPure RNA Cultured Cell Kit, Applied Biosystems). 12 prøver ble isolert i hver omgang. Avvik fra leverandørens protokoll: 45 µl DNase ble benyttet etter erfaring fra tidligere forsøk. Det rensede RNAet ble eluert fra rensokolonnen med 50 µl Elution Solution, og umiddelbart satt på is og frosset ned til -70°C. 4 µl ble frosset ned separat til kvalitetskontroll av RNA.

Kvantifisering og kvalitetskontroll av RNA

Total mengde RNA ble bestemt i hver prøve ved hjelp av spektrofotometri (NanoDrop ND-1000) . Konsentrasjonen i prøvene var i gjennomsnitt 41,4 ng/μl (min. 19,1, maks. 64,1, median 41,9). Spektrofotometri av nukleinsyrer gir typisk en absorbanse topp ved 260 nm og proteiner ved 280 nm. Ratioen (A_{260}/A_{280}) mellom disse gir en indikasjon på renheten av nukleinsyren (1), og bør være ca. 2.0 for RNA. Ratioen i prøvene var i gjennomsnitt 2,09 (min. 1,88, maks. 2,57, median 2,06).

Revers transkripsjon (RT)

RNA ble omdannet til cDNA ved bruk av et kommersielt tilgjengelig kit (High Capacity cDNA Reverse Transcription Kit) fra Applied Biosystems. Reaksjonen er basert på random hexamer priming. Totalt 200 ng RNA fra hver prøve ble pipettert i hver sin brønn på en 96-brønners PCR-plate. Fortynnet med PCR-grade H₂O til totalt volum på 10 μl. 5 prøver med konsentrasjon < 22 ng/μl ble ikke fortynnet. Alle prøver sto på is fram til pipettering. 5 prøver ble tint av gangen. PCR-plate sto på kjøleblokk med temperaturindikator under hele forløpet fram til temperatursyklus. Etter tilsetning av RNA ble 10 ul reaksjonsmix (random primer, nukleotider, cDNA buffermix, reverse transcriptase). Brønner ble fortløpende forseglet med lokk (8-cap strip fra Applied Biosystems). De 144 prøvene ble fordelt på 2 plater. Temperatursyklus ble gjennomført på PCR-maskin. Alle prøver fortynnet 1:8 med PCR-grade H₂O. cDNA oppbevart ved 4°C. Ved lengre tids lagring ble cDNA oppbevart ved -20°C.

No-reverse transcription-kontroll (NoRT)

NoRT er en prøve som har vært gjennom revers transkripsjon uten at enzymet RT er tilsatt. Dette er en kontroll på genomisk DNA (gDNA) kontaminering i RNA prøven og er en indikasjon på at prøvene er tilstrekkelig behandlet med DNase (1). At disse prøvene har C_q-verdier > 5 sykluser fra cDNA prøvene aksepteres (9). 16 ble prøver tilfeldig valgt som NoRT, hvorav minst 1 fra hver omgang RNA-isolering. Prøvene ble pipettert på samme måte og som, og samtidig med, RT. NoRT ble fordelt på de to PCR-platene sammen med RT.

Referansegen-panel

Kommersielt tilgjengelig humant referansegen-panel (Reference Gene Panel Human, TATAA Biocenter) ble benyttet. Hvert sett inneholdt reagenter nok til 100 qPCR-reaksjoner (100 rxn).

Tabell 1. Gen inkludert i det humane referansegenpanelet fra TATAA Biocenter.
*B2M og 18S rRNA er designet innen et exon og kan amplifisere genomisk DNA.

GAPDH	Glyceraldehyde-3-phosphate dehydrogenase
TUBB	Tubulin, beta polypeptide
PPIA	Peptidylpropyl isomerase A, Cyclophilin A
ACTB	Actin, beta
YWHAZ	Tyrosine 3/Tryptophan 5 - monooxygenase activation protein, zeta polypeptide
RRN18S*	18s rRNA
B2M*	Beta-2-microglobulin
UBC	Ubiquitin C
TBP	TATAA-box binding protein
RPLP	60S acidic ribosomal protein P0
GUSB	Beta-glucuronidase
HPRT1	Hypoxanthine-guanine phosphoribosyltransferase

qPCR

PCR ble utført samsvar med den anbefalte protokollen fra produsenten av referansegenpanelet. Denne protokollen forutsattes optimalisert fra produsentens side. PCR-reaksjonene hadde et totalt volum på 25 µl hvorav 9,5 µl PCR-grade H₂O, 1 µl primermix (TATAA Biocenter), 12,5 µl mastermix (FAST SYBR Green Master Mix fra Applied Biosystems) og 2 µl cDNA. qPCR ble utført på 7900HT Fast Real-Time PCR-System (Applied Biosystems) med et initialt trinn på 95°C i 20 sekunder, etterfulgt av 40 sykluser med denaturering (95°C i 20 sekunder), annealing (60°C i 20 sekunder) og elongation (72°C i 20 sekunder). Til slutt ble det kjørt en smeltekurve (95°C i 15 sekunder, 60°C i 15 sekunder, 95°C i 15 sekunder) med en temperaturforandring på 2°C per sekund). Det ble brukt PCR-plater med 96 brønner (MicroAmp Fast 96-Well Reaction Plate 0,1 ml) forseglet med adherende film (MicroAmp Optical Adhesive Film) fra Applied biosystems. 72 av de 144 prøvene ble tatt med i analysen. Dette utgjorde 3n av de 24 ulike gruppene inkludert i studien. 8 NoRT, 1 NTC (no target control) og 1 positiv kontroll (DNA, levert sammen med referansegenpanel) ble også analysert i tillegg til en fortynningskurve bestående av 5 fortytningstrinn. Kun et replikat av hver prøve og kontroll ble kjørt.

Smeltekurve

Temperaturen som må til for å skille to DNA-tråder kalles smeltetemperatur. Smeltetemperaturen vil være bestemt av DNA-molekylets lengde. Etter endt qPCR er alle PCR-produktene like store, gitt at kun den ønskede DNA-sekvensen er amplifisert. En smeltekurve vil bestemme smeltetemperaturen og avsløre om det finnes et eller flere PCR-produkt i prøven. En smeltekurve vil være et mål på reaksjonens spesifisitet (1) og vil være avhengig av primer designet. Dersom primerne «tar opp» flere produkt vil det gi flere topper på smeltekurven. Videre er PCR-produkt fra gDNA eller primer dimere er andre eksempler på hva som kan gi PCR-produkt av ulik størrelse i en reaksjon. Spesifisiteten kan også kontrolleres ved å kjøre PCR-produktet på en gel (1).

SYBR Green

SYBR Green er et fargestoff som binder seg uspesifikt til dobbeltrådet DNA. Det vil si at det binder seg, og dermed gir et fluorescerende signal, fra alle DNA-molekyler i qPCR. Et uønsket PCR-produkt vil derfor ha direkte innvirkning på C_q i den aktuelle prøven (3).

Primer-dimer

Primerene som avgrensner DNA-sekvensen som skal amplifiseres i qPCR er selv korte DNA-tråder. I enkelte tilfeller vil disse kunne gå sammen og danne et dobbeltrådet DNA-molekyl, en såkalt primer-dimer. En primer-dimer vil binde SYBR Green og avgi et fluorescerende signal (3).

Effektivitet og fortynningskurve

I teorien doubles antall DNA-kopier for hver temperatursyklus (2^n). Det gir en PCR-effektivitet på 100%. I praksis vil effektiviteten variere (4). Effektivitet på 90% vil gi færre DNA-kopier ($1+0,90^n$). For å korrigere for dette, må effektiviteten bestemmes. Effektivitet (E) bestemmes ut fra stigningstallet (slope) på en standardkurve (3) hvor Y-aksen er C_q-verdi og X-aksen er logaritmen av konsentrasjon.

$$E = 10^{(-1/\text{slope})} - 1$$

10 µl fra hver cDNA-prøve inkludert i studien (144 prøver) ble brukt som standardløsning (10). Cq-verdier for en fortyningsserie på 5 fortyningstrinn utgjorde standardkurven. Det ble fortynt 1:5 mellom hvert trinn. Hvert fortyningstrinn ble analysert i duplikat. Dette ble gjort for hvert gen i referansegenpanelet.

Databehandling

qPCR-rådata ble behandlet med software SDS version 2.4.1 (Applied Biosystems). Baseline ble satt automatisk. Threshold ble satt manuelt til 0,20, unntatt på HPRT1, TBP, TUBB, GUSB hvor det ble satt til 0,18. Data for standardkurve beregnet automatisk, med unntak av effektivitet som ble regnet ut manuelt. Rådata ble deretter eksportert til Excel (Microsoft) for klargjøring av datafil til videre behandling i GenEx 5.4.2.128 (MultiD Analyses AB). GenEx er et verktøy for statistisk behandling av qPCR-data, og inkluderer geNorm og NormFinder. Datasettet ble først analysert for statistiske parametere som gjennomsnitt og 95% KI. Deretter ble Cq-verdiene justert for effektivitet med påfølgende analyse med geNorm og NormFinder. For NormFinder ble det valgt å se på variasjonen innad i en gruppe hvor alle prøvene var inkludert. Alternativt kunne man i tillegg sett på intra- og intervariasjonen mellom alle 24 grupper i studien (kontroll, behandlet, normalt oksygennivå, hypoksi eller for hvert tidspunkt i tidskurven) .

Arbeidsprosessen

Denne oppgaven er en del av forskerlinjeprosjektet mitt, og er utført over en tidsperiode på 1 år og 4 måneder. Det meste av praktisk arbeid har vært utført av meg selv, med hjelp og veiledning underveis fra andre på forskergruppen. Trine Kalstad har vært i første rekke til praktisk hjelp: Oppkultivering av celler; avlastning med celledell, forsøk; bistått i det store arbeidet med å så ut celler til forsøk, cDNA; hjelp til pipettering, qPCR; hjelp til pipettering, innkjøp, kompetanse på og ansvar for utstyr og andre fasiliteter. Andre ansatte har også tidvis hjulpet meg med merking av prøverør til forsøkene samt annen praktisk tilrettelegging. Oppkultivering av celler til forsøk tok ca. 7 uker. Forsøk tok ca. 1,5 døgn for hver runde. Isolering, rensing og kvantifisering av RNA krevde ca. 8 dager. RT 3 dager. I gjennomsnitt brukte jeg 1 dag på qPCR og analyse av rådata per gen. Forbehandling av data før statistisk

behandling 5 dager. Databehandling i GenEx 1 dag. Resten av tiden har vært brukt til forberedelser og etterarbeid i forbindelse med laboratoriearbeid, prøving og feiling, litteraturgjennomgang og innlæring av bruk av software. Med unntak av noen uforventede dårlige resultater, har arbeidet med oppgaven gått som forventet.

Resultater

Tabell 1: Oppsummering av rådata fra qPCR.

Gen	Smelte-kurve prøver	Smeltekurve fortynnings- kurve	Cq gj.snitt	Cq NoRT	Cq NTC	Slope	R ²	Effektivite t
UBC	Dobbeltopp 81-82°C		22	>36	-	-3,55	0,998	0,91
YWHAZ	Enkeltopp 78°C		23	>32	-	-3,54	0,999	0,92
GAPDH	Enkeltopp 79°C		21	>33	>38	-3,48	0,999	0,94
RPLP	Enkeltopp 80°C		22	>30	>34	-3,39	0,999	0,97
B2M	Enkeltopp 76°C		22	>35	-	-3,37	0,999	0,98
PPIA	Enkeltopp 76°C		21	>32	-	-3,36	0,998	0,98
RRN18S	Enkeltopp 81°C		9	>28	>37	-3,28	0,999	1,02
GUSB	Enkeltopp 81°C	Primer dimer?	28	>33	>37	-3,18	0,997	1,06
HPRT1	Enkeltopp 77°C	81-82°C Primer dimer?	27	>35	>37	-2,88	0,988	1,22
ACTB	Antydning til dobbeltopp 83-84°C		21	>32	-	-2,66	1,000	1,38
TUBB	Enkeltopp 82°C		28	>34	>35	-2,46	0,997	1,55
TBP	Enkeltopp 79°C		28	>34	>35	-2,29	0,995	1,73

Kontroller: NoRT og NTC hadde i enkelte tilfeller amplifisering av et produkt. Uten unntak var alle Cq-verdiene > 5 syklene fra prøvene. Eventuell kontaminering av genomisk DNA antas da å ikke ha betydning for resultatet. Positiv kontroll var tilfredsstillende.

Smeltekurver: Alle prøver hadde en enkelt topp på smeltekurven med lik smeltetemperatur. Dette gjaldt alle gen med unntak av UBC og ACTB. Her hadde smeltekurvene en sammenhengende topp med to maksimumspunkter. Dette kan skyldes at PCR-produktene har en AT-rik og en CG-rik ende og dermed ulikt

smeltepunkt for de to endene. Det kan likevel ikke utelukkes at to PCR-produkt ble amplifisert. Gel burde ha blitt kjørt for å verifisere antall PCR-produkt. Enkelte NoRT og NTC ble amplifisert. Samtlige hadde smeltepunkt lavere enn prøvene og positiv kontroll. Det antas at produktene var primer dimere. På fortynningskurvene til GUSB og til en viss grad HPRT1 var det to topper på smeltekurven. Den ene falt sammen med smeltepunkt for prøver og positiv kontroll, den andre med smeltepunkt for NTC. Det antas å være primer dimere.

R2: Alle fortynningskurver oppnådde $R2 > 0,995$.

Effektivitet: En effektivitet fra 0,91 til 1,73 ble oppnådd på de ulike genene. Stor variasjon i effektivitet er vanlig i biologiske prøver på grunn av inhibisjon (10), men dette gir oftest lave effektiviteter. Effektiviteter mellom 0,95 og 1,05 anses å være bra, men effektiviteter mellom 0,90 og 1,05 blir også godtatt .

Vurdering: GUSB ble ekskludert på grunn av problem med primer dimer. HPRT1, ACTB, TUBB og TBP ble ekskludert for sin uforklarlige høye effektivitet. UBC og ACTB ble ekskludert på grunn av usikker spesifisitet. Gjenværende 6 gen, YWAHZ, GAPDH, RPLP, B2M, PPIA og RRN18S, ble ansett som aktuelle kandidater og ble analysert videre.

geNorm validerte RPLP og YWHAZ til de to best egnede referansegenene. M-verdien var 0,65. NormFinder utpekte YWHAZ til den beste kandidaten med SD 0,37, med RPLP som andrevalg. Øvrige gen ble rangert likt.

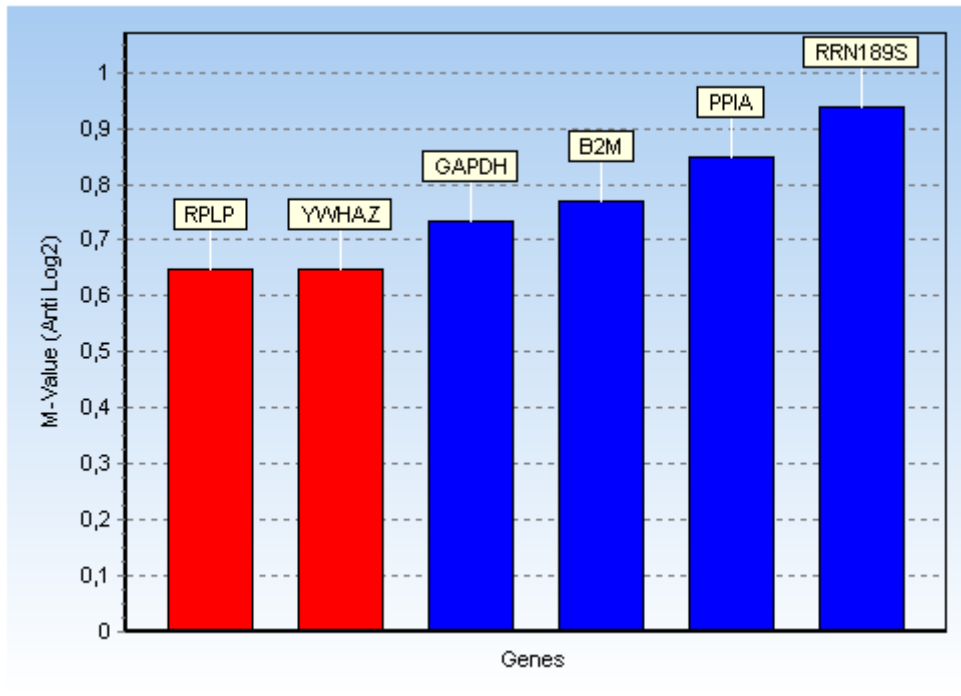


Fig. 3. geNorm. Beste kandidater som referansegen RPLP og YWHAZ med M-verdi 0,65. Etterfulgt av GAPDH 0,73, B2M 0,77, PPIA 0,85 og RRN18S.

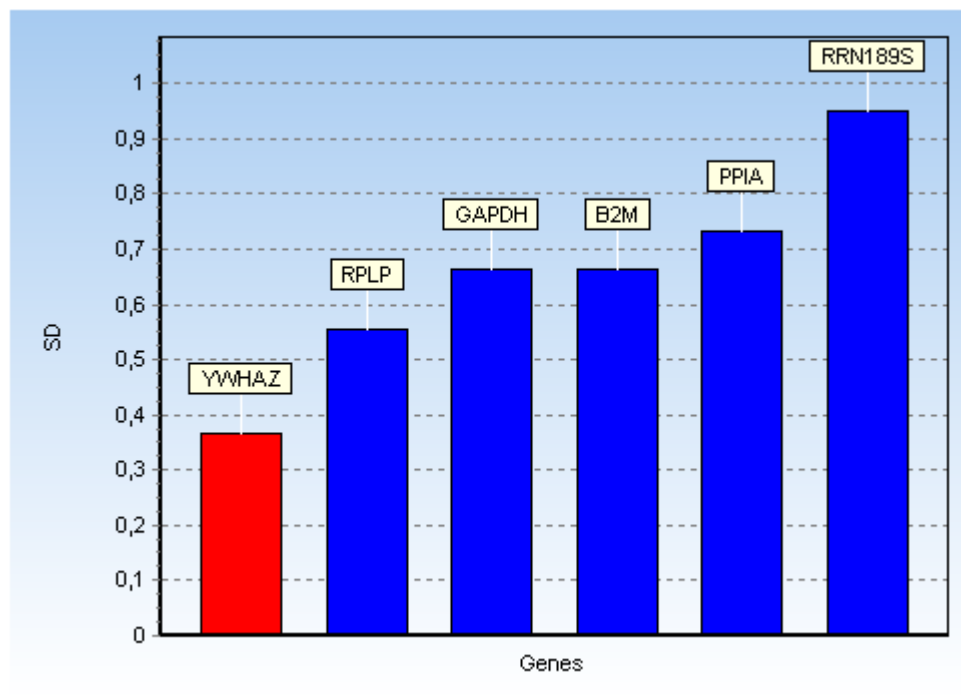


Fig. 4. NormFinder. Beste kandidat som referansegen YWHAZ med SD 0,71. Nest beste RPLP 0,55, deretter GAPDH 0,66, B2M 0,66, PPIA 0,73 og RRN18S 0,95.

Diskusjon

I denne studien ble kun renhet på RNA målt ved A_{260}/A_{280} -ratioen. Metoden gir lite annet enn en indikasjon på renhetsgrad, og RNA integrity number (RIN) burde blitt bestemt (1,4,11,12). RNA er et ustabil molekyl, og RIN sier noe om hvor intakt RNA er. RNA fra dette eksperimentet ble utvunnet fra rene cellekulturer på en skånsom måte, og det er derfor ikke urimelig å anta at RNA holdt en tilfredsstillende kvalitet. Allikevel bør det vurderes om RIN skal bestemmes i ettertid.

For å øke presisjonen er det vanlig praksis å kjøre 2 til 3 qPCR-replikater for hver prøve (3). Mine forsøk besto av i alt 24 ulike grupper for hver n . Dersom jeg i tillegg hadde valgt å analysere med 2 eller 3 qPCR-replikater ville jeg kun hatt reagenser til å inkludere $1n$ i referansegenstudien. Det er vist at flere RT-replikater (13) og replikat av prøver (antall n) (2) er en vel så god løsning, sistnevnte spesielt for cellekulturer. Dette dannet grunnlaget for å velge å bruke $3n$ uten RT- og qPCR-replikater, siden vi mener hoved variasjonen i våre prøver ligger her. Å øke antall replikater på de andre trinnene ville i tillegg ha medført en større belastning økonomisk og tidsmessig, uten nødvendigvis å gi bedre resultater i dette studiet.

Fordelen med å benytte seg av et kommersielt tilgjengelig referansegenpanel er mange, eksempelvis et relevant utvalg av gen og gode primer-design. Protokollen validert og optimalisert med tanke på spesifisitet og effektivitet. Likevel kom fire av de inkluderte genene (HPRT1, ACTB, TUBB og TBP) ut med svært høy effektivitet i denne studien. Biologiske prøver vil alltid ha ulikt innhold av stoffer som kan ha innvirkning på PCR-analyser (3,10,14). Dette dreier seg stort sett om inhibitorer. C_q for den høyeste konsentrasjonen i en fortynningsrekke vil typisk være høyere enn forventet ved inhibisjon. Fortynningene med lavere konsentrasjon vil bli inhibert i mindre grad, siden inhibitorene også blir fortynnet, og C_q -verdiene vil bli som forventet. På en standardkurve vil dette dermed bli synlig ved at distansen mellom de ulike kurvene blir mindre ettersom prøven fortynnes. I sum vil dette ha innvirkning på standardkurvens slope og gi høyere effekt enn 100%. Urenheter i prøvene, som proteiner, vil kunne virke inhiberende. A_{260}/A_{280} - ratio ved spektrofotometri av RNA var mellom 1,88 og 2,57 med median 2,06. Selv om de fleste prøvene var rene nok, kan enkeltprøver ha bidratt til inhibering da utgangsløsningen for fortynningsrekken

besto av komponenter fra alle 144 prøver. At effektiviteten for disse genene var såpass mye høyere enn 100 % (>122 %) taler også for inhibisjon. Valgte metode for bestemmelse av effektivitet (og også den mest brukte) tenderer å gi for høy effektivitet (15). Forsøk på å *double* mengden cDNA og ble gjort for genet TBP (resultater ikke vist). Det resulterte i en effektivitet på 90 %. Det taler ikke for inhibisjon, da *fortynning* av mulige inhibitorer vil gi bedre effektivitet. Feilpipettering er likeledes en kilde til uønsket effektivitet. I motsetning til prøvene, ble fortynningskurven kjørt med 2 qPCR-replikater. Alle standardkurver oppnådde $R^2 > 0,99$, en indikasjon på tilfredsstillende presisjon ved pipettering. En systematisk pipetteringsfeil ved tillaging av fortynningsløsninger er ikke utelukket av den grunn. Systematisk bruk av for lavt volum fortynningsvæske (eks. 70 μ l i stedet for 80 μ l) eller for høyt volum standardløsning vil typisk gi for høy effektivitet. Både inhibisjon og feilpipettering kan være sannsynlig. qPCR ble kjørt samme dag for genene HPRT1, ACTB, TUBB og TBP, og samme fortynningsrekke ble benyttet. Øvrige gen ble analysert andre dager med fortynningsrekker laget respektive dager. Det sannsynliggjør en systematisk pipetteringsfeil som grunnlag for den høye effektiviteten - eventuelt andre systematiske feil som intrådte den dagen. Det faktum HPRT1, ACTB, TUBB og TBP ble kjørt samme dag utelukker i midlertid ikke at inhibisjon kan være årsaken. Flere fortynningsrekker ble kjørt for de fire genene uavhengig av hverandre (resultat ikke vist). Eksempelvis ble to nye kurver for ACTB kjørt med effektivitet på henholdsvis 96 % og 99%. Det styrker teorien om feilpipettering for dette genet. For TUBB forble effektiviteten 157 %. Dette svekker teorien om feilpipettering. Det var heller ikke tegn til primer dimer-problem i smeltekurvene. HPRT1, ACTB, TUBB og TBP ble derfor ekskludert som mulig referansegen på grunn av usikkerhet rundt effektivitet i denne oppgaven.

GUSB ble ekskludert på grunnlag av primer dimere som forstyrret analysen. Det ble gjort forsøk på å øke og redusere mengde prøve, primermix, SYBR Green og totalt reaksjonsvolum (ikke publisert) uten å redusere problemet med primer dimere. Det var også antydning til primer dimer i siste fortynningstrinn HPRT1.

UBC og ACTB ble ekskludert da det ikke kunne utelukkes at det ble dannet flere enn det ønskede PCR-produktet. Ved å kjøre PCR-produktene på en agarosegel kunne

man med sikkerhet ha bestemt antall PCR-produkt. Dette ble det dessverre ikke anledning til å gjennomføre.

De 6 gjenværende gen, YWHAZ, GAPDH, RPLP, B2M, PPIA og RRN18S, ble rangert i geNorm og NormFinder. De ulike statistiske metodene rangerte samtlige gen likt. Som Andersen et al. påpeker er dette ikke gitt: geNorm heller i retning av å velge gen med lik uttryksprofil over hele prøvesettet - noe som ikke er gitt er sammenfallende med genene med lavest variasjon i genuttrykket (8). Ved den parvise analysen i geNorm, oppnådde de to beste kandidatene med en M-verdi på 0,65. Det er tidligere vist at stabile referansegene i homogene prøveutvalg oppnår M-verdi $< 0,5$ og i mer heterogene prøveutvalg M-verdi < 1 (16). I utgangspunktet utgjør cellekulturer et homogent prøveutvalg, men det kan diskuteres hvorvidt de ulike behandlingene og ulike behandlingstidsrommene har innvirkning på homogeniteten. Det er heller ikke utelukket at de beste kandidatene til dette prøvesettet befinner seg blant de 6 genene som ble ekskludert fra rangeringen på grunn av dårlig effektivitet, primer dimerer og usikker spesifisitet. Dette kan antydes ved å se på 95% KI for de ulike genene i prøvesettet og av innledende analyser i geNorm og Normfinder der alle 12 gen ble inkludert (vedlegg 1 og 2).

Det er i dag anbefalt å bruke minimum 3 referansegene i qPCR-analyser (4), og for prøvegrupper med større variasjon kan flere være påkrevd (1,7,8). Jeg velger likevel å trekke den slutningen at 2 referansegene vil være godt nok i mitt studie. Dette begrunnes ut fra en helhetsvurdering av kostnader, tidsbruk og gevinst i form av pålitelige data. Dette prøvesettet skal brukes til å avgjøre ved hvilket tidspunkt HCAEC har størst respons på de ulike stimuli de er utsatt for, og er ment som et utgangspunkt for videre studier.

Konklusjon

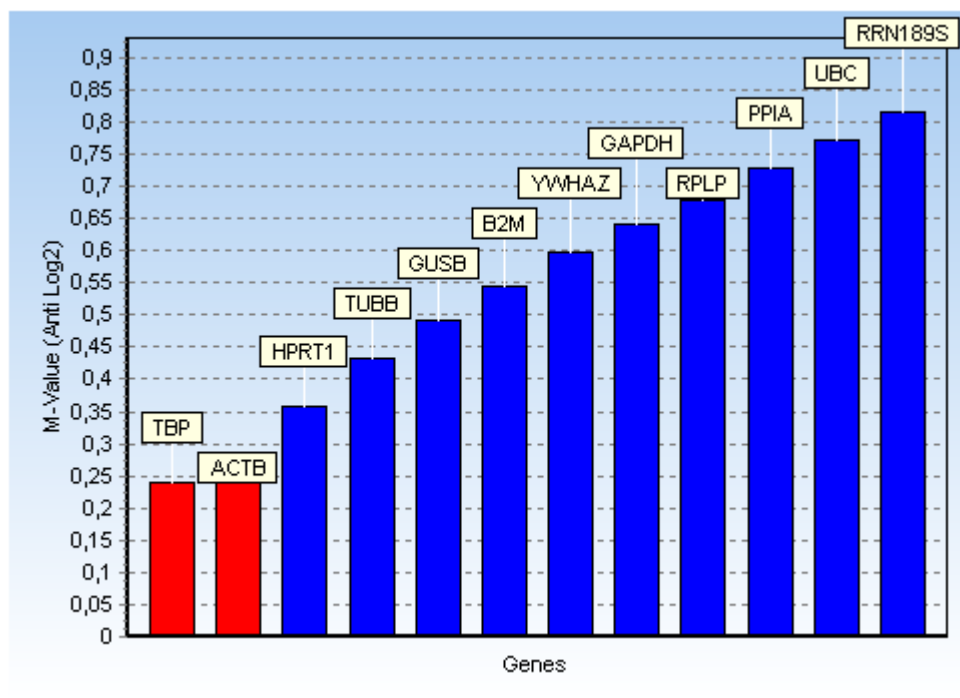
YWHAZ og GAPDH ansees som gode og mange nok referansegen til bruk i prøvesettet fra forsøkene for å se utvikling i genuttrykk etter ulik behandlingstid. I ettertid bør kvaliteten på RNA bestemmes om dette er praktisk mulig. PCR-produkt fra UBC og ACTB bør kjøres på gel for å avgjøre spesifisiteten, og om mulig inkluderes i en ny analyse i geNorm og NormFinder - gitt at ATCBs effektivitet på 96% - 99% kan reproduseres. Dersom dataene skal publiseres, bør en gjøre en full optimalisering av effektivitet på TUBB og TBC. Det er usikkert om videre arbeid med GUSB og HPRT1 vil være hensiktsmessig for dette prøvesettet på grunn av primer dimer-problemet.

Referanser

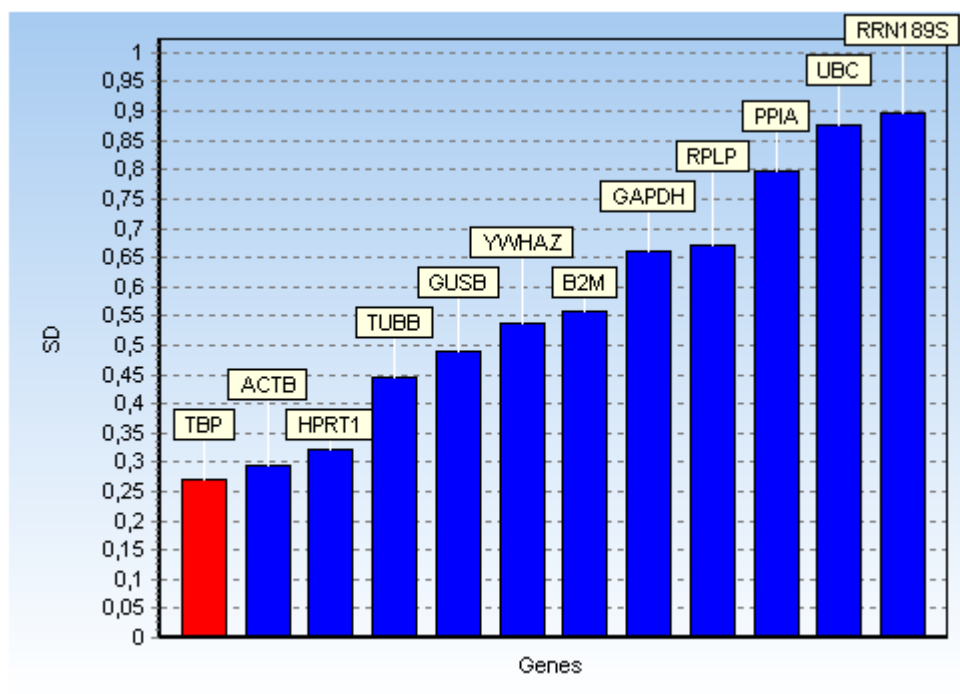
1. Bustin SA, Benes V, Garson JA, Hellemans J, Huggett J, Kubista M, et al. *The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments*. *Clinical chemistry*. 2009;55(4):611-22.
2. Tichopad A, Kitchen R, Riedmaier I, Becker C, Stahlberg A, Kubista M. *Design and optimization of reverse-transcription quantitative PCR experiments*. *Clinical chemistry*. 2009;55(10):1816-23.
3. Kubista M, Andrade JM, Bengtsson M, Forootan A, Jonak J, Lind K, et al. *The real-time polymerase chain reaction*. *Molecular aspects of medicine*. 2006;27(2-3):95-125.
4. Derveaux S, Vandesompele J, Hellemans J. *How to do successful gene expression analysis using real-time PCR*. *Methods*. 2010;50(4):227-30.
5. Bustin SA. *Absolute quantification of mRNA using real-time reverse transcription polymerase chain reaction assays*. *Journal of molecular endocrinology*. 2000;25(2):169-93.
6. Gibbs PJ, Cameron C, Tan LC, Sadek SA, Howell WM. *House keeping genes and gene expression analysis in transplant recipients: a note of caution*. *Transplant immunology*. 2003;12(1):89-97.
7. Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A, et al. *Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes*. *Genome biology*. 2002;3(7):RESEARCH0034.
8. Andersen CL, Jensen JL, Orntoft TF. *Normalization of real-time quantitative reverse transcription-PCR data: a model-based variance estimation approach to identify genes suited for normalization, applied to bladder and colon cancer data sets*. *Cancer research*. 2004;64(15):5245-50.
9. Laurell H, Iacovoni JS, Abot A, Svec D, Maoret JJ, Arnal JF, et al. *Correction of RT-qPCR data for genomic DNA-derived signals with ValidPrime*. *Nucleic acids research*. 2012;40(7):e51.
10. Stahlberg A, Aman P, Ridell B, Mostad P, Kubista M. *Quantitative real-time PCR method for detection of B-lymphocyte monoclonality by comparison of kappa and lambda immunoglobulin light chain expression*. *Clinical chemistry*. 2003;49(1):51-9.
11. Schroeder A, Mueller O, Stocker S, Salowsky R, Leiber M, Gassmann M, et al. *The RIN: an RNA integrity number for assigning integrity values to RNA measurements*. *BMC molecular biology*. 2006;7:3.

12. Fleige S, Pfaffl MW. *RNA integrity and the effect on the real-time qRT-PCR performance*. Molecular aspects of medicine. 2006;27(2-3):126-39.
13. Stahlberg A, Hakansson J, Xian X, Semb H, Kubista M. *Properties of the reverse transcription reaction in mRNA quantification*. Clinical chemistry. 2004;50(3):509-15.
14. Nolan T, Hands RE, Ogunkolade W, Bustin SA. *SPUD: a quantitative PCR assay for the detection of inhibitors in nucleic acid preparations*. Analytical biochemistry. 2006;351(2):308-10.
15. Pfaffl M W. Relative quantification. I: Dorak M T, red. Real-time PCR. Abingdon: Taylor & Francis group; 2006. 63.
16. Hellemans J, Mortier G, De Paepe A, Speleman F, Vandesompele J. *qBase relative quantification framework and software for management and automated analysis of real-time quantitative PCR data*. Genome biology. 2007;8(2):R19.

Vedlegg 1



geNorm. Analyse av beste referansegene med rådata som ikke er korrigert for effektivitet fra alle 12 gen fra referansegenpanelet. GUSB med sitt betydelige primer dimer-problem er også med. Legg merke til at samtlige gen ekskludert fra den endelige analysen rangeres bedre enn de inkluderte. M-verdien er også vesentlig bedre.



NormFinder. Analyse av beste referansegene med rådata som ikke er korrigert for effektivitet fra alle 12 gen fra referansegenpanelet. GUSB med sitt betydelige primer dimer-problem er også med. Legg merke til at samtlige gen ekskludert fra den endelige analysen rangeres bedre enn de inkluderte.