TaqMan® Assays

...not just Primers and Probes

Dr. Uta Möllers
Support
The Applera Genomic Initiative

Integrating high value genomic content with state of the art technology solutions
**Integrated Workflow**

<table>
<thead>
<tr>
<th>DNA Extraction &amp; RT reaction</th>
<th>Whole Genome Survey Hypothesis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Data Interpretation</td>
</tr>
<tr>
<td></td>
<td>Biological validation or Candidate Gene Analysis</td>
</tr>
<tr>
<td></td>
<td>Data interpretation</td>
</tr>
<tr>
<td></td>
<td>Final Target List for Routine Testing</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th># genes</th>
</tr>
</thead>
<tbody>
<tr>
<td>„ALL“</td>
</tr>
<tr>
<td>50-500</td>
</tr>
<tr>
<td>1-100</td>
</tr>
</tbody>
</table>
Real Time PCR
Chemistry Basics

Signal generation with TaqMan® MGB Probes
5’ Nuclease Assay using TaqMan® probes

Forward Primer

Reverse Primer

TaqMan® Probe
Why short Amplicons?

• Recommendation: Assays with short Amplicons
  – robust PCR-Reactions = high sensitivity!
  – high efficiency = no standard curve necessary!

but...

- Also long Amplicons can be used
  🙁 PCR-efficiency of longer Amplicons drops
- Quantification requires standard curves
5’ Nuclease Assay
using TaqMan® probes

Diagram showing the 5’ nuclease assay using TaqMan® probes, with a green arrow indicating the reaction from 3' to 5'.
5’ Nuclease Assay
using TaqMan® probes

Strand Displacement
5’ Nuclease Assay using TaqMan® probes

Strand Displacement and Cleavage
5’ Nuclease Assay using TaqMan® probes

Finishing Synthesis
TaqMan® MGB probes

MGB: Minor Groove Binder
NFQ: Non-Fluorescent Quencher

Stabilization of last 5-6 bp on 3’ end
⇒ Short and more specific probes for a given Tm
⇒ high signal intensity
TaqMan® Assays

How to get the assay for YOUR gene?
Manual Approach in the Past

Gold Standard Quantitation Technology but One-by-One Approach
TaqMan® Assays remove Bottlenecks

Gold Standard Quantitation Technology
Effort and cost decreased

www.appliedbiosystems.com
TaqMan® Assays
Custom TaqMan® Assays
TaqMan® Low Density Array

Applied Biosystems
7900HT, 7500 and 7300
Real-time PCR Systems
TaqMan® Assays
Massive Bioinformatics Pipeline - High Content

RefSeq
GenBank
Celera Data

Mask
Repeats, SNPs & discrepancies (CDS/Public Db)

Map Gene to Genome (CDS/Public Db), mark Exons

Genome QC BLAST

Transcript QC BLAST

Assay Design

High-scoring Assays to Web

Manufacturing QC

Constant Remapping (CDS/Public Db)
TaqMan® Assays
Massive Bioinformatics Pipeline - High Content

- RefSeq
- GenBank
- Celera Data

RefSeq → GenBank → Celera Data → Mask repeats, SNPs & discrepancies (CDS/Public Db) → Map Gene to Genome (CDS/Public Db), mark Exons → Map Gene to Genome (CDS/Public Db), mark Exons

Genome QC BLAST → Transcript QC BLAST → Assay Design → High-scoring Assays to Web

Assay Design → Manufacturing QC → Constant Remapping (CDS/Public Db) → Manufacturing QC → Assay Design

High-scoring Assays to Web → Map Gene to Genome (CDS/Public Db), mark Exons → Map Gene to Genome (CDS/Public Db), mark Exons
TaqMan® Assays

Source of transcripts

• Start with **NCBI Reference Sequence transcripts**
  (http://www.ncbi.nlm.nih.gov/LocusLink/refseq.html)
  – Best known, highly curated and non-redundant transcripts in the public domain

• Manually curated by NCBI scientist

• Good evidence available: cDNA clones
TaqMan® Assays

Regularly updated data from up to 40 total databases, both proprietary and public
Value of the Celera Discovery System\textsuperscript{SM} Platform in the Assay Pipeline

Locus Link Genes (Homo sapiens)

2817

51261

RefSeq Transcripts (Homo sapiens)

NM_002081

NM_017844

NM_016552

NM_023089

Celera Genes (Homo sapiens)

PCG20534

hCG1811573

hCG2013451

hCG32867

Celera Transcripts (Homo sapiens)

hCT11513

hCT2304937

hCT2304939

hCT1953155

hCT23847

hCT2304496

PUBLIC

plus proprietary

PUBLIC
TaqMan® Assays
Massive Bioinformatics Pipeline - High Content

RefSeq
GenBank
Celera Data

Mask
Repeats, SNPs
& discrepancies
(CDS/Public Db)

Map Gene
to Genome
(CDS/Public Db),
mark Exons

Genome QC
BLAST

Transcript QC
BLAST

Assay Design

High-scoring
Assays to Web

Manufacturing
QC

Constant
Remapping
(CDS/Public Db)
Transcript Pre-Processing 1

- Design only in a region of **unambiguous sequence**
  - No repeats (poly-nucleotide repeats, ALU repeats etc)
- Mapping the masked transcripts to the Celera and Public Genome assembly
  - Sequence discrepancies
  - BLAST against **Celera SNP database**
    - Most comprehensive SNP database, integrates both public and Celera-proprietary SNPs
    - Over 4 million known human SNPs
    - More than 3 million mouse SNPs
Transcript Pre-Processing 2

- Exon-Exon boundaries of multi-exon genes are marked
- Identification of single-exon transcripts

**cDNA specific Assays!**
TaqMan® Assays
Massive Bioinformatics Pipeline - High Content

- RefSeq
- GenBank
- Celera Data
  
  Mask
  Repeats, SNPs
  & discrepancies
  (CDS/Public Db)

- Genome QC
  BLAST

  Transcript QC
  BLAST

- Map Gene
to Genome
(CDS/Public Db),
mark Exons

- High-scoring
Assays to Web

- Manufacturing
QC

- Constant
Remapping
(CDS/Public Db)

Assay Design
Assay Design

• Long experience with probe and primer design of 5’ nuclease assays

• TaqExpress
  – Automated process
  – Significant enhancement of the algorithms resident in AB Primer Express® Software
TaqMan® Assays
Massive Bioinformatics Pipeline - High Content

1. RefSeq, GenBank, Celera Data
2. Mask Repeats, SNPs, & discrepancies (CDS/Public Db)
3. Map Gene to Genome (CDS/Public Db), mark Exons
4. Genome QC BLAST
5. Transcript QC BLAST
6. Assay Design
7. High-scoring Assays to Web
8. Manufacturing QC
9. Constant Remapping (CDS/Public Db)
In Silico Quality Scoring 1

Penalty score for

• Assay designs that are not highly specific for the gene of interest

• Assay designs that may not accurately report the quantitative expression results for a particular target
In Silico Quality Scoring 2

• Transcript BLAST scoring
  – Penalty is assigned if an assay detects any closely homologous transcript(s)

• Genome BLAST scoring
  – Penalty is assigned if an assay hits a second location on the genome

• Intron Size scoring
  – Penalty is assigned if an assay is designed across an exon-exon boundary that spans a small intron (<2kb)
TaqMan® Assays
Massive Bioinformatics Pipeline - High Content

RefSeq
GenBank
Celera Data

Mask
Repeats, SNPs
& discrepancies
(CDS/Public Db)

Map Gene
to Genome
(CDS/Public Db), mark Exons

Genome QC
BLAST

Transcript QC
BLAST

Assay Design

High-scoring
Assays to Web

Manufacturing
QC

Constant
Remapping
(CDS/Public Db)
TaqMan® Assays
Massive Bioinformatics Pipeline - High Content

RefSeq
GenBank
Celera Data

Mask
Repeats, SNPs & discrepancies
(CDS/Public Db)

Map Gene to Genome (CDS/Public Db), mark Exons

Genome QC
BLAST

Transcript QC
BLAST

Assay Design

High-scoring Assays to Web

Manufacturing QC

Constant Remapping (CDS/Public Db)
Remapping

- Genomic information is still changing
- New transcripts are being discovered
- Constant remapping after a new transcript database is released (i.e. when RefSeq is updated, approx. Every 4 weeks)
TaqMan® Assays

Universal conditions, no optimization!

www.allgenes.com
TaqMan® Assays

How to find the ones I need
myScience
research environment

Genomic Products

GENE EXPRESSION  |  learn more

- TaqMan® Gene Expression Assays
- TaqMan® Low Density Array
- Custom TaqMan® Gene Expression Assays
- Related Gene Expression Products

GENOTYPING  |  learn more

- TaqMan® SNP Genotyping Assays
- Custom TaqMan® Gene Expression Assays
- SNPlex™ Genotyping System
- Related Genotyping Products

RESEQUENCING  |  learn more

- VariantSEQ™ Resequencing System
- Related Products

Research Tools

- Gene Search
  Find information and genomic products for your human, mouse or rat gene or genomic region of interest.
  - Human
  - Mouse
  - Rat

- myScience Map View
  Browse the human, mouse and rat genomes and genomic products using our interactive graphical viewer.

- Panther™
  Browse or search for proteins organized by biological processes and molecular functions.
<table>
<thead>
<tr>
<th>Locus Link ID</th>
<th>Species</th>
<th>Locus Link Gene Name</th>
<th>Locus Link Gene Symbol</th>
<th>Panther Molecular Function</th>
<th>Panther Biological Process</th>
<th>Public Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>4313</td>
<td>Homo sapiens</td>
<td>matrix metalloproteinase 9 (gelatinase B, 92kDa gelatinase, 92kDa type IV collagenase)</td>
<td>MMPS</td>
<td>2 more...</td>
<td>2 more...</td>
<td>chr. 20 45,322,989 - 45,330,521</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Celera Gene ID**
- hCG40015
  - 11 assays
  - Ambion siRNAs
  - 31 assays
  - 3 RBS IDs
  - 3 transcripts

**Public RefSeq and GenBank Transcripts**

**TaqMan Assays**
- Gene Expression

**siRNAs**

**TaqMan Assays**
- SNP Genotyping
human Chromosome 20
Location displayed graphically

Reference: NCBI - Homo sapiens - Build34 - chr. 20

Public RefSeq
Public GenBank
Gene Expression Assays available
Click on assay to see which sequences are interrogated

Mouse over to get Assay information
Integrated Workflow

- **RNA Extraction & RT reaction**
- **Whole Genome Survey Hypothesis**
  - Data Interpretation
- **Biological validation or Candidate Gene Analysis**
  - Data interpretation
- **Final Target List for Routine Testing**

# genes
- „ALL“: 50-500
- 1-100

- CELERA DISCOVERY SYSTEM®
- Applied Biosystems
TaqMan® Low Density Array

384 spots

TaqMan® Assays pre-loaded

1 to 8 Samples

12 to 380 Targets (per one array)

No. ofReplicates
### Configuring my Custom Array

<table>
<thead>
<tr>
<th>Number of Genes</th>
<th>Number of Replicates</th>
<th>Number of Samples per Array</th>
</tr>
</thead>
<tbody>
<tr>
<td>12</td>
<td>4</td>
<td>8</td>
</tr>
<tr>
<td>16</td>
<td>3</td>
<td>8</td>
</tr>
<tr>
<td>24</td>
<td>2</td>
<td>8</td>
</tr>
<tr>
<td>32</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>48</td>
<td>1</td>
<td>8</td>
</tr>
<tr>
<td>64</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>96</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>96</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>381</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>
Trademarks

- For Research Use Only. Not for use in diagnostic procedures.

- Applied Biosystems and Celera are registered trademarks and AB (Design), Applera, Celera Discovery System, iScience, and iScience (Design) are trademarks of Applera Corporation or its subsidiaries in the US and/or certain other countries.

- TaqMan is a registered trademark of Roche Molecular Systems, Inc.

- SYBR is a registered trademark of Molecular Probes, Inc.

- All other trademarks are the sole property of their respective owners.

- © 2005 Applied Biosystems. All rights reserved.