hogeschool Leiden

PCR primer and probe design, an introduction

August 2014

Willem van Leeuwen

TopLab

content

- What is a primer?
- Primer classification
- probes
- Probe classification
- Melting curve analysis

Primer: characteristics

- Short fragment of single-stranded DNA (ssDNA)
- Specific for target and delimiting target

- Start for DNA polymerase in PCR
- May be modified for special applications (stuffer sequence, label, etc)

Nucleic acis target; matrix for primer design

phage MS2 genome

GenBank: V00642.1

GenBank Graphics

```
>gi|15081|emb|V00642.1| phage MS2 genome
GGGTGGGACCCCTTTCGGGGTCCTGCTCAACTTCCTGTCGAGCTAATGCCATTTTTAATGTCTTTAGCGA
GACGCTACCATGGCTATCGCTGTAGGTAGCCGGAATTCCATTCCTAGGAGGTTTGACCTGTGCGAGCTTT
TAGTACCCTTGATAGGGAGAACGAGACCTTCGTCCCCTCCGTTCGCGTTTACGCGGACGGTGAGACTGAA
GATAACTCATTCTCTTTAAAATATCGTTCGAACTGGACTCCCGGTCGTTTTAACTCGACTGGGGCCAAAA
CGAAACAGTGGCACTACCCCTCTCCGTATTCACGGGGGGCGTTAAGTGTCACATCGATAGATCAAGGTGC
```

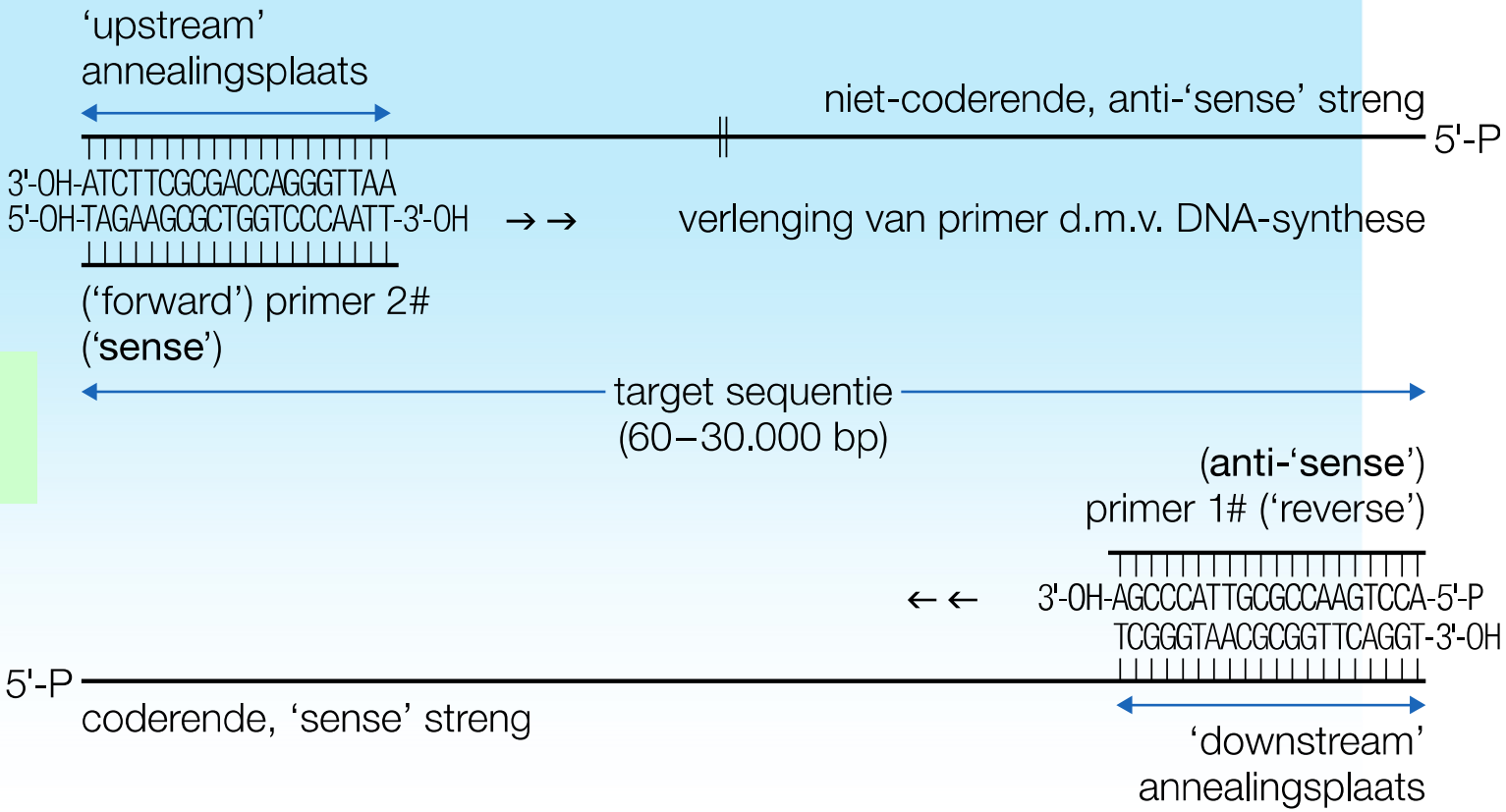
FASTA format: notation 5’ ➔ 3’; single-strand

Primer and target, amplification from ds configuration from 3’-OH

RNA targets cDNA; min-streng

Primers are complementary to target

Fasta from 5’ to 3’



phage MS2 genome: FASTA

5’-P-GGGTGGGACCCCTTTCGGGGTCCTGCTCAACTTCCTGTCG-3’OH

base nr 1……base nr 40

rules for primer design (q)PCR

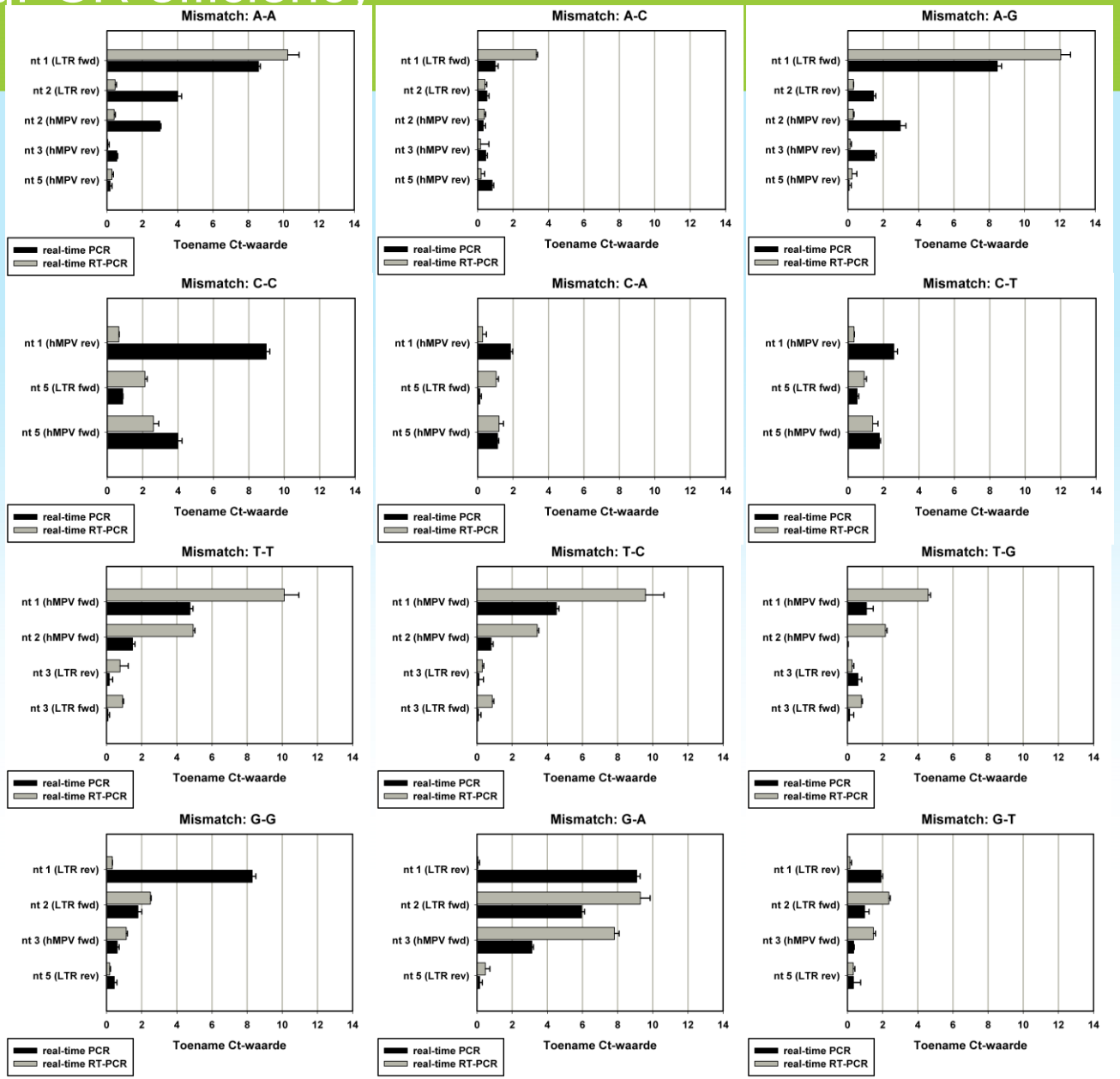
- Length primers 16-25 nt: oligonucleotide
- GC-level in primer 45-60%.
- Tm primers between 50^{o}C and 70^{o}C
- Variation of Tm among primers in primerpair, not exceeding 3^{o}C.
- G or C at 3'end of primers.
- No more than 2 Gs and/or Cs in the last 5 bases at 3'end.
- No G or C repeats of >3 nt in primer.

Influence of mastermix composition on primer bindingsite mismatch tolerance

HIV-1 LTR forward primer:

5'-AAGCCTCAATAAAGCTTGC-**C**$^{(5)}$-T-**T**$^{(3)}$-**G**$^{(2)}$-**A**$^{(1)}$-**3'**

Influemce of mismatches in the primer binding site on qPCR efficiency

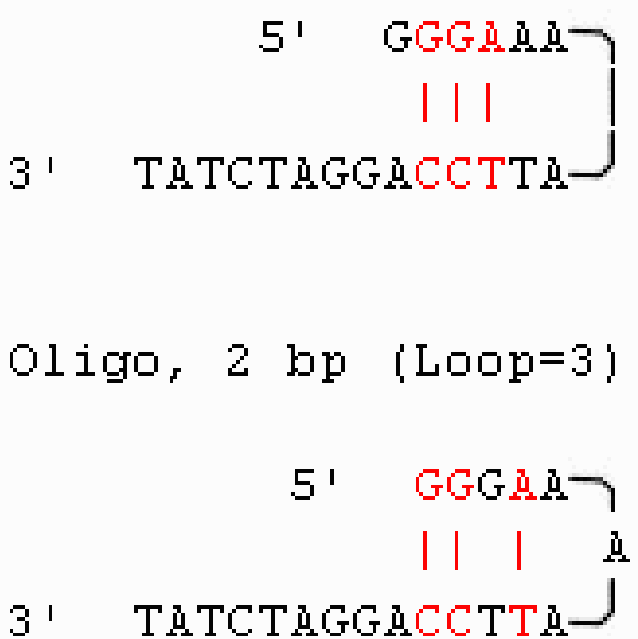


Preventing underquantification due to mutations in primers and probes

- Avoid mutations in the last 5 nucleotides of the 3‘ part of the primers
  - Avoid C-C!!, A-A, G-G, G-A en A-G
- Avoid mutations in the last nucleotide at the 3‘ part of the primers
  - Avoid T-C, C-T en T-T
- A-C, C-A, T-G en G-T mismatches are relatively well tolerated

Rules for primer design in (q)PCR

- avoid intramolecular complementary bases in primer (secondary structures).



http://bioweb.uwlax.edu/genweb/molecular/seq_anal/primer_design/primer_design.htm

Rules for primer design in (q)PCR

- no intermolecular complementary bases between primers. Avoid primer-dimers formation.

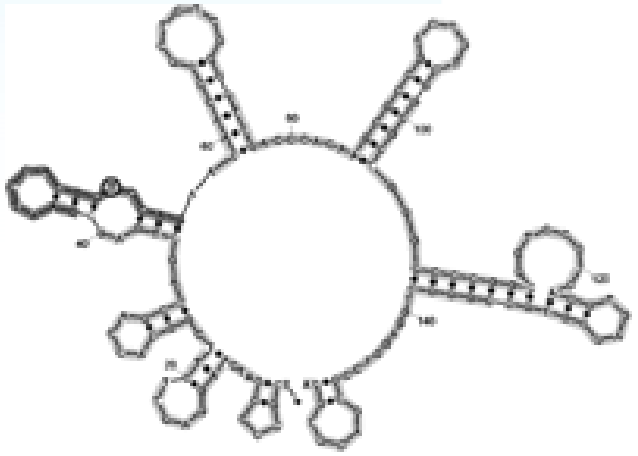
```
4 bp, delta G = -6.6 kc/m (bad!)  (worst= -36.6)
        5' GGGAAAATTCCAGGATCTAT 3'
             ||||  ||||
3' TATCTAGGACCTTAAAAGGG 5'


4 bp, delta G = -5.4 kc/m (bad!)  (worst= -36.6)
      5' GGGAAAATTCCAGGATCTAT 3'
             ||||
3' TATCTAGGACCTTAAAAGGG 5'
```

http://bioweb.uwlax.edu/genweb/molecular/seq_anal/primer_design/primer_design.htm

Rules for primer design in (q)PCR assay: requirements for the amplicon

- The shorter the amplicon, the more efficient the PCR. A PCR product of at least 75 bp is desirable to discriminate between specific product and primer-dimer (In validation state)
- Rather no templates with long single nucleotide stretches.
- No secundary structures in the (hairpins).
- If possible; amplicon with 50 – 60% GC-level.



http://nar.oxfordjournals.org/content/34/13/e90.full

T_m (melting temperature) primer

Definition Tm= temperature at which 50% of primer is hybridized to target DNA and 50% not

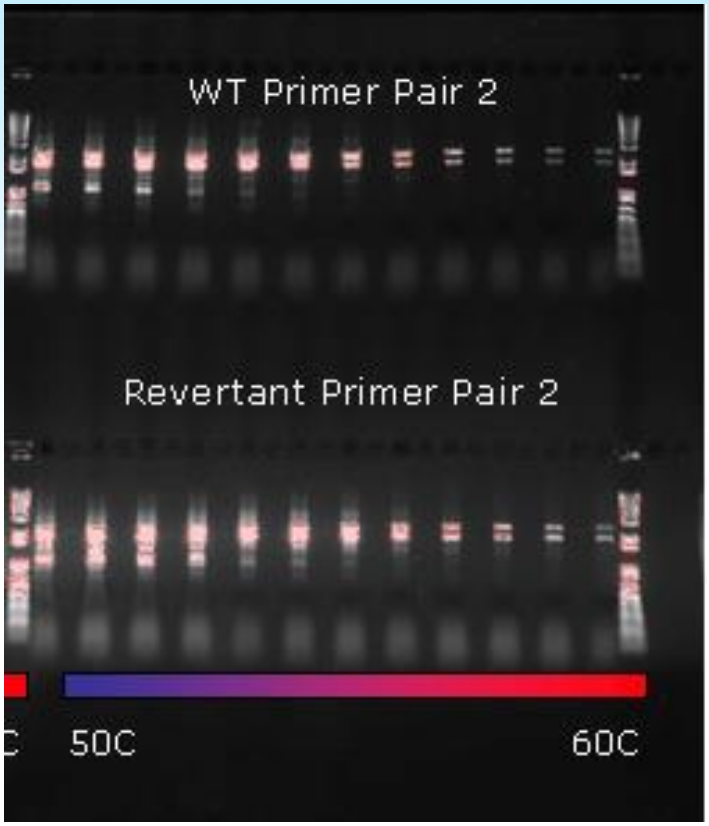
Depends on:

- primer length
- GC level
- mismatches with target?

- Tm= 2(A+T) + 4(G+C) ("wild guess")
- Ta=Tm – 5°C (guideline for Ta optim.)

Optimization Ta of primer in PCR

- Optimization of PCR using temperature gradient.



Optimization of Tanneal . A gradient PCR from 50 to 60°C was performed. At high temps, no PCR product was symthezised. Too low: by-products
In this example, Tanneal optimum: 58°C

http://www.stanford.edu/group/dahlia_genetics/2008_reports/lab_5/lab_5_meta.htm

Are those good primers?

```
   1 atgcagtttt actatgataa gtgtctccca ggcaacagca ccatgatgaa taattttgat
  61 gctgttacca tgaggttgac tgacatttca ttgaatgtca aagattgcat attggatatg
 121 tctaagtctg ttgctgcgcc taaggatcaa atcaaaccac taatacctat ggtacgaacg
 181 gcggcagaaa tgccacgcca gactggacta ttggaaaatt tagtggcgat gattaaaagg
 241 aactttaacg cacccgagtt gtctggcatc attgatattg aaaatactgc atctttagtt
 301 gtagataagt tttttgatag ttatttgctt aaagaaaaaa gaaaaccaaa taaaaatgtt
 361 tctttgttca gtagagagtc tctcaataga tggttagaaa agcaggaaca ggtaacaata
 421 ggccagctcg cagattttga ttttgtagat ttgccagcag ttgatcagta cagacacatg
 481 attaaagcac aacccaagca aaaattggac acttcaatcc aaacggagta cccggctttg
 541 cagacgattg tgtaccattc aaaaaagatc aatgcaatat ttggcccgtt gtttagtgag
 601 cttactaggc aattactgga cagtgttgat tcgagcagat ttttgttttt cacaagaaag
 661 acaccagcgc agattgagga tttcttcgga gatctcgaca gtcatgtgcc gatggatgtc
 721 ttggagctgg atatatcaaa atacgacaaa tctcagaatg aattccactg tgcagtagaa
 781 tacgagatct ggcgaagatt gggttttgaa gacttcttgg gagaagtttg gaaacaaggg
 841 catagaaaga ccaccctcaa ggattatacc gcaggtataa aaacttgcat ctggtatcaa
 901 agaaagagcg gggacgtcac gacgttcatt ggaaacactg tgatcattgc tgcatgtttg
 961 gcctcgatgc ttccgatgga gaaaataatc aaaggagcct tttgcggtga cgatagtctg
1021 ctgtactttc caaagggttg tgagtttccg gatgtgcaac actccgcgaa tcttatgtgg
1081 aattttgaag caaaactgtt taaaaaacag tatggatact tttgcggaag atatgtaata
1141 catcacgaca gaggatgcat tgtgtattac gatcccctaa agttgatctc gaaacttggt
1201 gctaaacaca tcaaggattg ggaacacttg gaggagttca gaaggtctct ttgtgatgtt
1261 gctgtttcgt tgaacaattg tgcgtattac acacagttgg acgacgctgt atgggaggtt
1321 cataagaccg cccctccagg ttcgtttgtt tataaaagtc tggtgaagta tttgtctgat
1381 aaagttcttt ttagaagttt gtttatagat ggctctagtt gttaa
```

Forward primer
aaagaaaaaa gaaaaccaaa

Reverse primer
gaaaataatc aaaggagcct

Free software to develop primers (and probes)

- NCBI/ Primer-BLAST

  http://www.ncbi.nlm.nih.gov/tools/primer-blast/index.cgi?LINK_LOC=BlastHome

- Primer3(plus)

  http://frodo.wi.mit.edu/primer3

hogeschool Leiden

Primer3

TopLab

Primer3 (v. 0.4.0) Pick primers from a DNA sequence.	Checks for mispriming in template.	disclaimer	Primer3 Home
	Primer3plus interface	cautions	FAQ/WIKI

Paste source sequence below (5'->3', string of ACGTNacgtn -- other letters treated as N -- numbers and blanks ignored). FASTA format ok. Please N-out undesirable sequence (vector, ALUs, LINEs, etc.) or use a Mispriming Library (repeat library): NONE

☑ Pick left primer, or use left primer below:	☐ Pick hybridization probe (internal oligo), or use oligo below:	☑ Pick right primer, or use right primer below (5' to 3' on opposite strand):

Pick Primers Reset Form

Sequence Id: A string to identify your output.

Targets: E.g. 50,2 requires primers to surround the 2 bases at positions 50 and 51. Or mark the source sequence with [and]: e.g. ...ATCT[CCCC]TCAT.. means that primers must flank the central CCCC.

Excluded Regions: E.g. 401,7 68,3 forbids selection of primers in the 7 bases starting at 401 and the 3 bases at 68. Or mark the source sequence with < and >: e.g. ...ATCT<CCCC>TCAT.. forbids primers in the central CCCC.

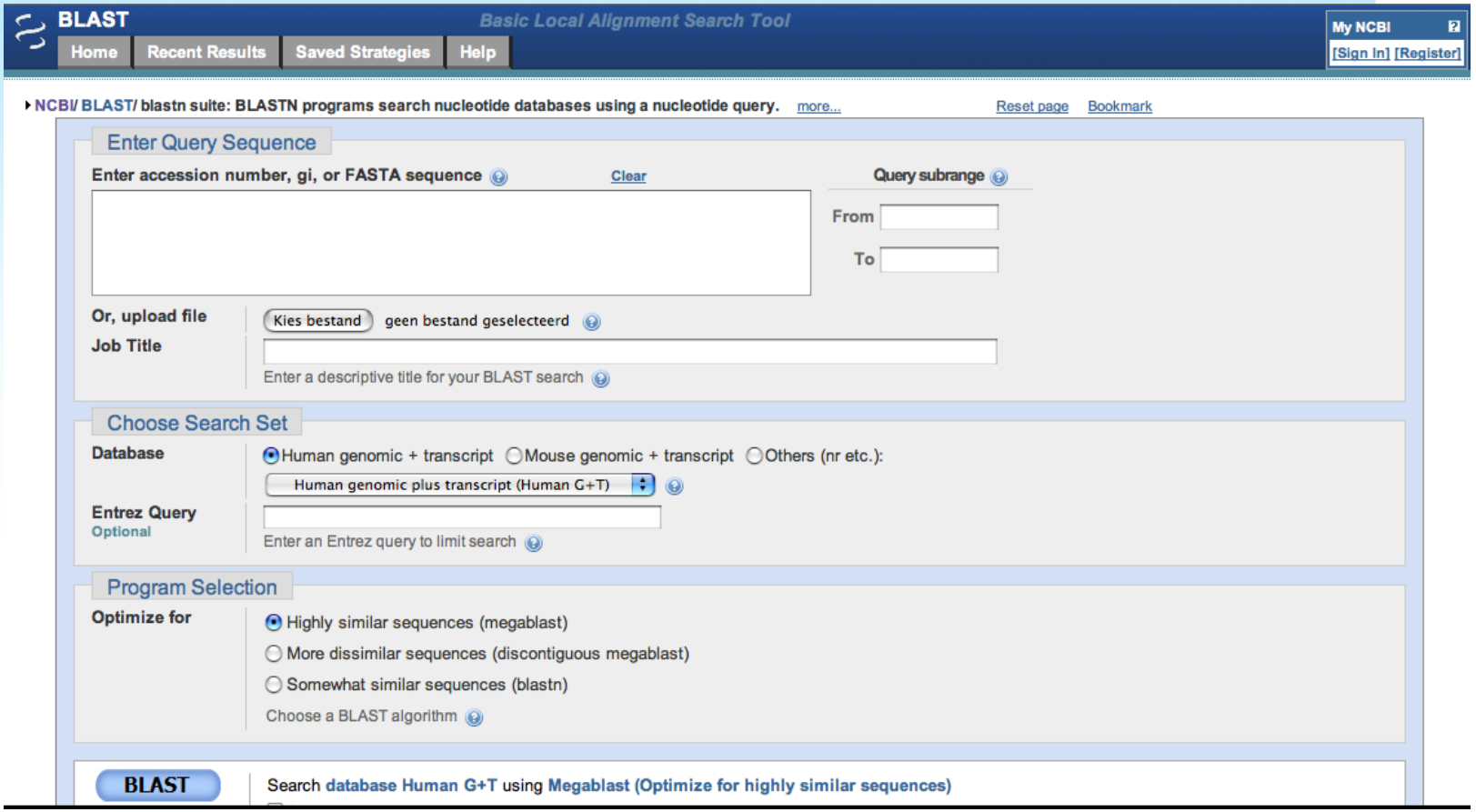
Product Size Ranges: 150-250 100-300 301-400 401-500 501-600 601-700 701-850 851-1000

Number To Return: 5 Max 3' Stability: 9.0

Max Repeat Mispriming: 12.00 Pair Max Repeat Mispriming: 24.00

BLAST (basic local alignment search tool); to check primer specificity

- E-value: calculation of frequency of this sequence in databaset
- The smaller, the better (preferabaly < 0.1)



Primerbank; a source for existing primers

MASSACHUSETTS GENERAL HOSPITAL | CCIB The Center for Computational and Integrative Biology | HARVARD MEDICAL SCHOOL

Primer Bank

PCR Primers for Gene Expression Detection and Quantification

Home/Search | PCR Protocol | Primer Statistics | Comments | Links | Citation Policy | Help/FAQ

Primer Search

Search for PCR Primers

Search where: GenBank Accession

Species: All Species

For text: Submit

You can blast your sequence against the primerbank sequence DB **here.**

Order Oligos

You can have primers synthesized and PCR reaction products sequenced at:

DNA Core Facility
Center for Computational and Integrative Biology

PrimerBank is a public resource for PCR primers. These primers are designed for gene expression detection or quantification (real-time PCR). PrimerBank contains over 306,800 primers covering most known human and mouse genes. There are several ways to search for primers: GenBank Accession, NCBI protein accession, NCBI Gene ID, Gene Symbol[New!], PrimerBank ID or Keyword (gene description) or you can blast your gene sequence against the primerbank Sequence DB[New!].

hogeschool Leiden

PCR primer classification

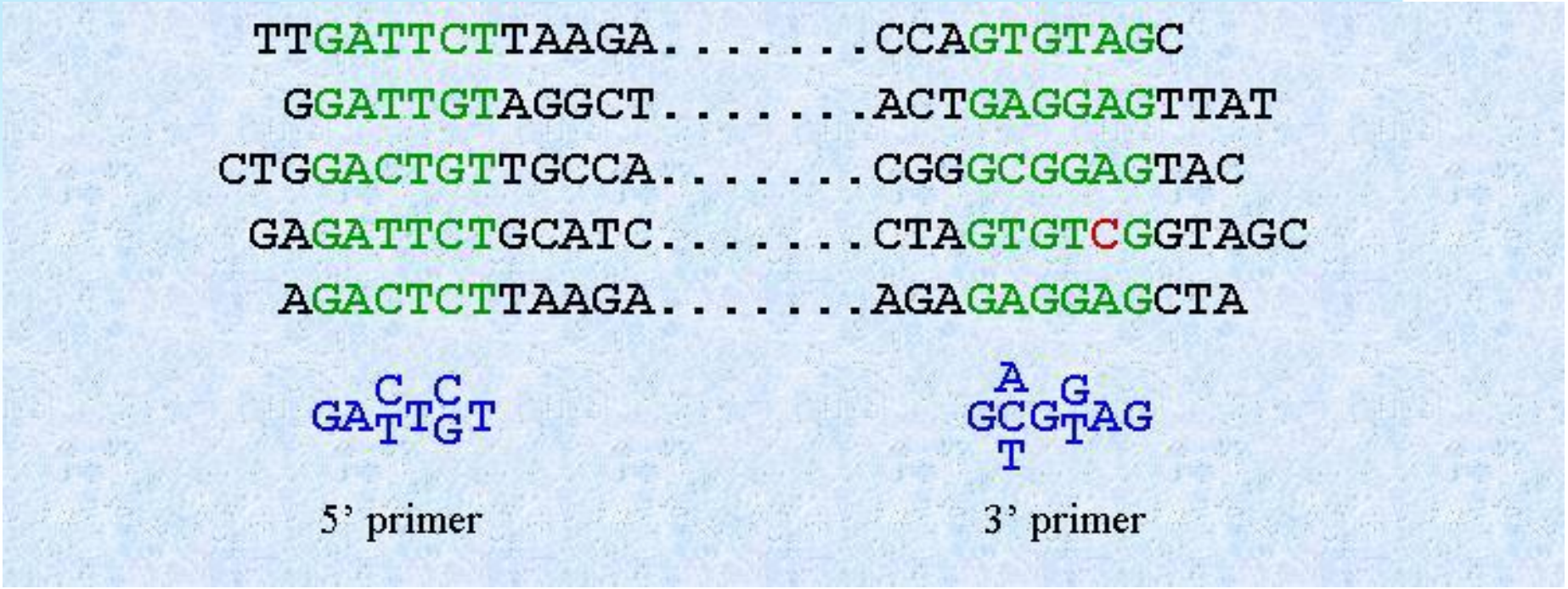
TopLab

primers

- Primers define the amplified target

- Special primers
  - Primers with an extension (Illumina NGS, MLPA)
  - Degenerate primers
  - …
  - …
  - …

Degenerate primers



What does this mean?

AWGTYRDCCTA

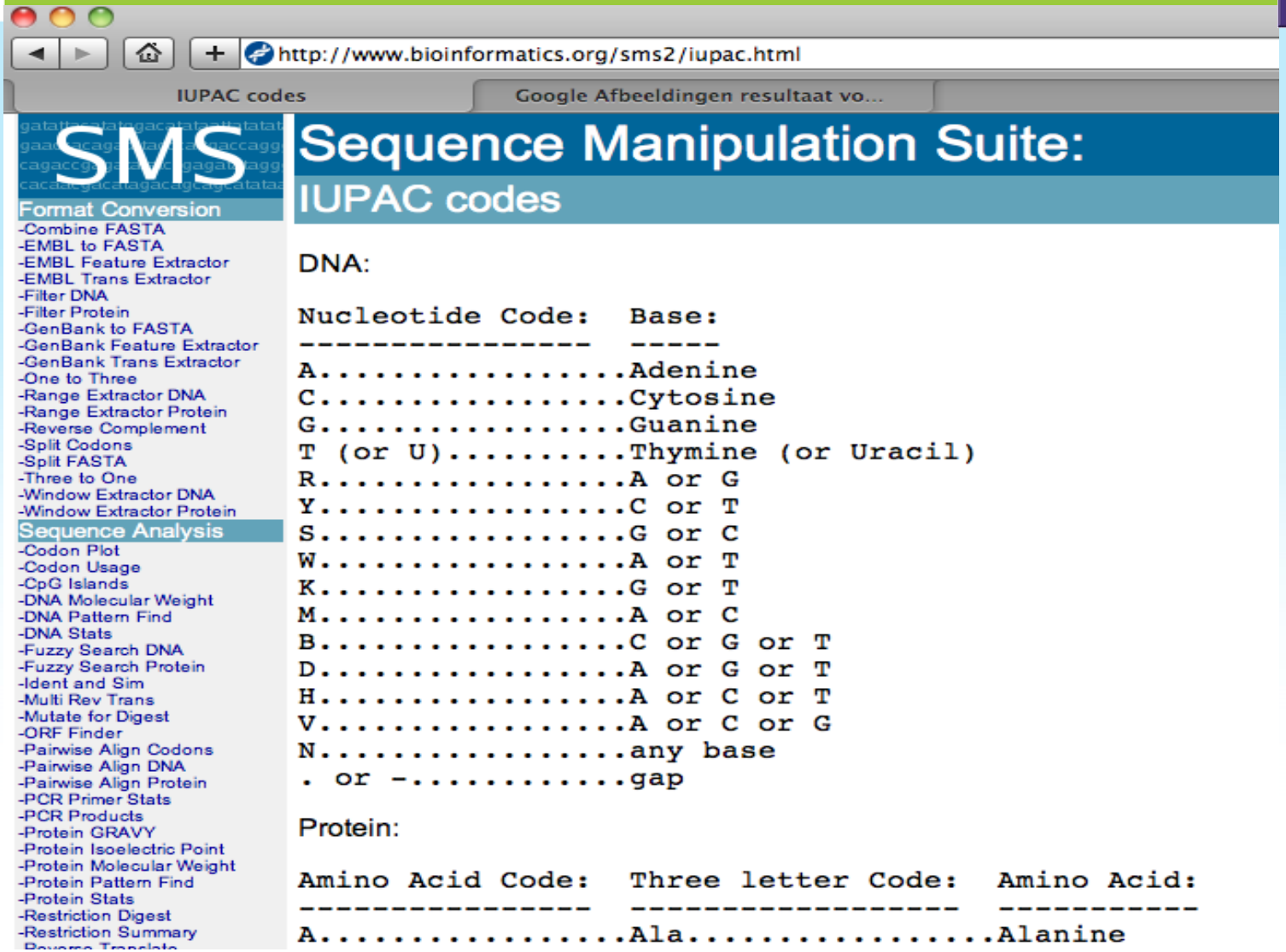
AAGTCRACCTA

T T G

T

Coding for primer syntheziser.
Randomly choose A or T; C or T; A of G or T at the indicated positions.
In this example, a “degenerated primer” is desired.

Your primer order: IUPAC code



hogeschool Leiden

Probes for qPCR

TopLab

Probes for qPCR

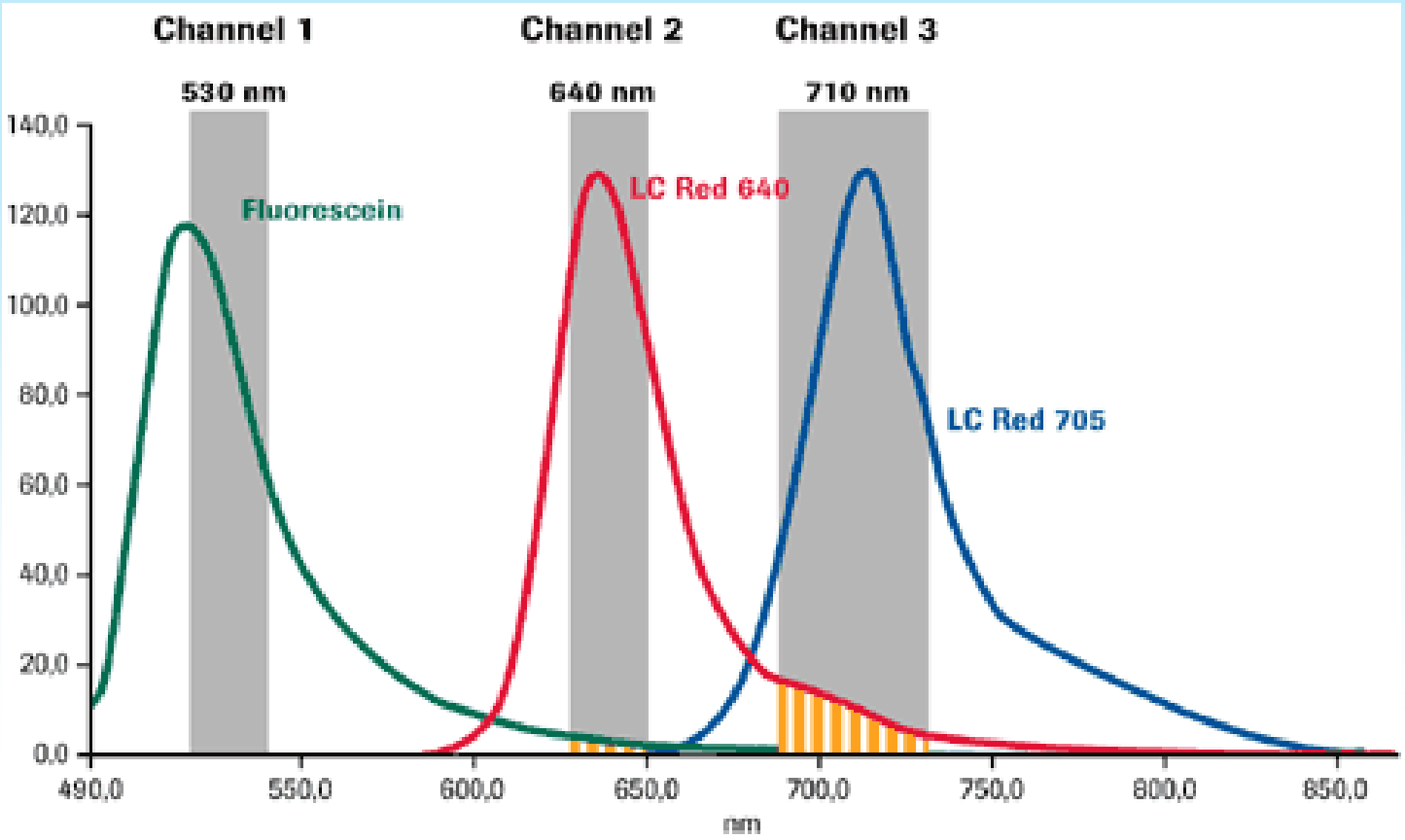
- Detection by fluorescence, labelled on probe
- Specificity of PCR wil be strongly elevated by using probes.

Design of hydrolisation (former TaqMan) probes

1. Tm probe is 5-10^{o}C higher thanTm primers.
2. Hydroization probes are often used in a 2-step PCR protocol (annela+elongation), Tm high 60-70^{o}C.
3. <30 nt (physical distance between primer – probe too large (non-optimal quenching).
4. No G at 5’-end (guanine may quench fluorochrome)

Spectral overlap fluorophores

(illustration Roche)



fluorophores in qPCR

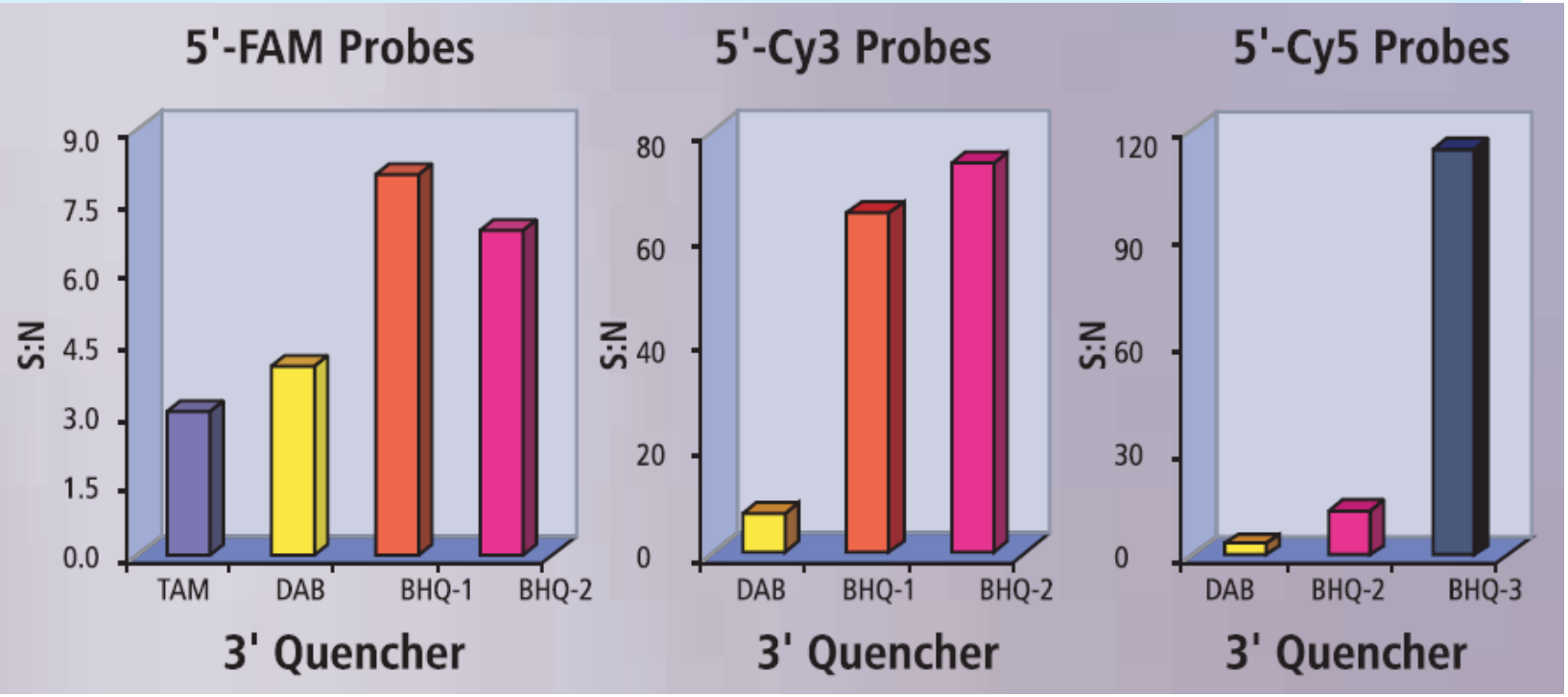
Fluorophore	Alternative Fluorophore	Excitation (nm)	Emission (nm)
FAM		495	515
TET	CAL Fluor Gold 540 [A]	525	540
HEX	JOE, VIC [B], CAL Fluor Orange 560 [A]	535	555
Cy3 [C]	NED [B], Quasar 570 [A], Oyster 556 [D]	550	570
TMR	CAL Fluor Red 590 [A]	555	575
ROX	LC red 610 [E], CAL Fluor Red 610 [A]	575	605
Texas red	LC red 610 [E], CAL Fluor Red 610 [A]	585	605
LC red 640 [E]	CAL Fluor Red 635 [A]	625	640
Cy5 [C]	LC red 670 [E], Quasar 670 [A], Oyster 645 [D]	650	670
LC red 705 [E]	Cy5.5 [C]	680	710

[A)] CAL and Quasar fluorophores are available from Biosearch Technologies; [B)] VIC and NED are available from Applied Biosystems; [C)] Cy dyes are available from Amersham Biosciences; [D)] Oyster fluorophores are available from Integrated DNA Technologies; and [E)] LC (Light Cycler) fluorophores are available from Roche Applied Science.

www.molecular-beacons.org/.../marras,mmb06(335)3.pdf

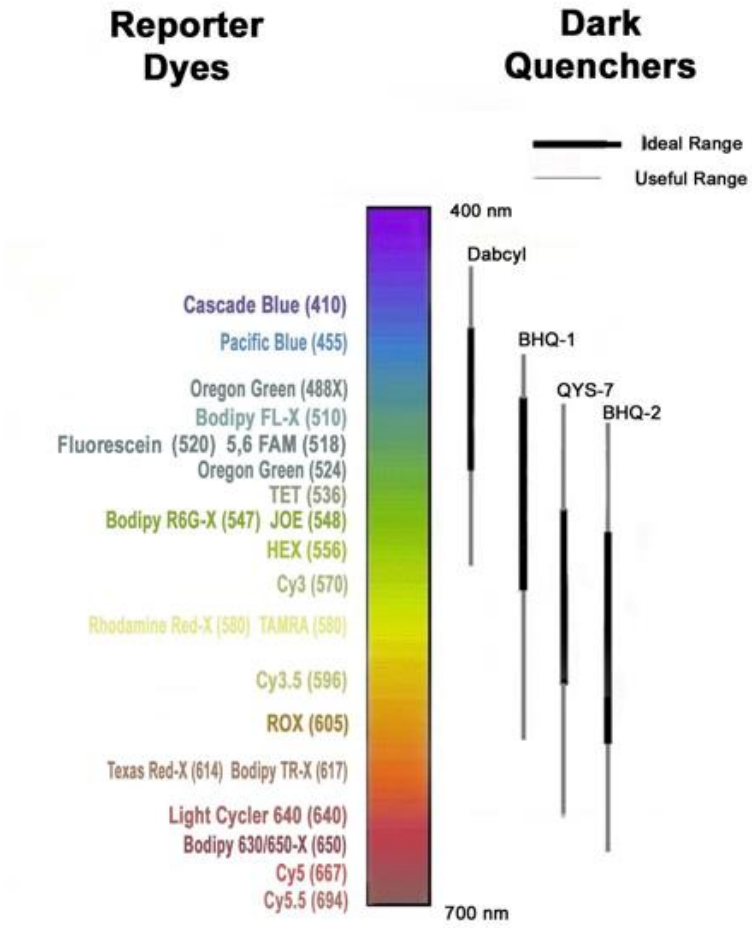
Quenchers

- First generation
  - TAMRA, DABCYL
  - No full quenching, self-emitting
- second generation
  - BHQ1,2,3 (Black Hole Quencher), full quenching



Combinations fluorophores and quenchers

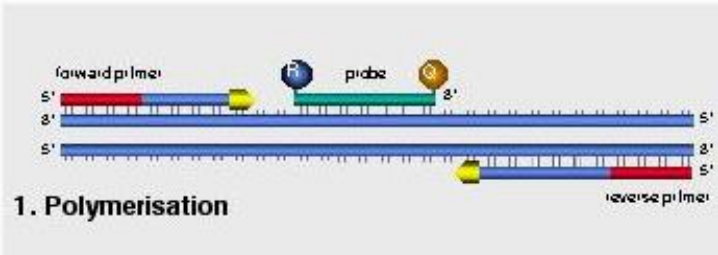
Figure 1: Spectral Overlap of Dark Quenchers and Fluorescent Dyes

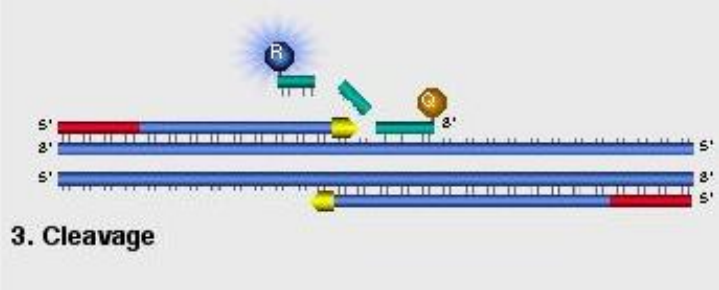


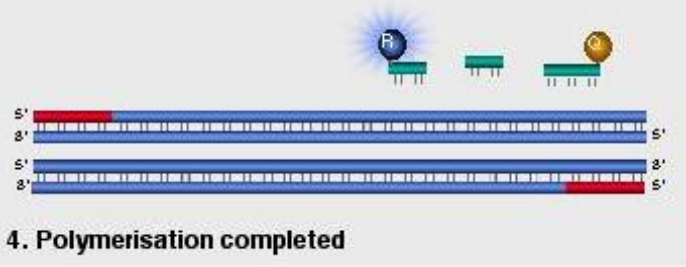
Probe classification

Hydrolysis of dual-labelled probe

- 5’-3’ exonuclease activity of Taq DNA polymerase
- Hydrolysis proceeds from base to base



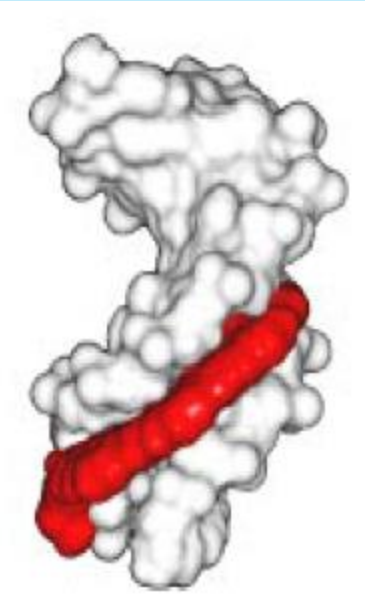




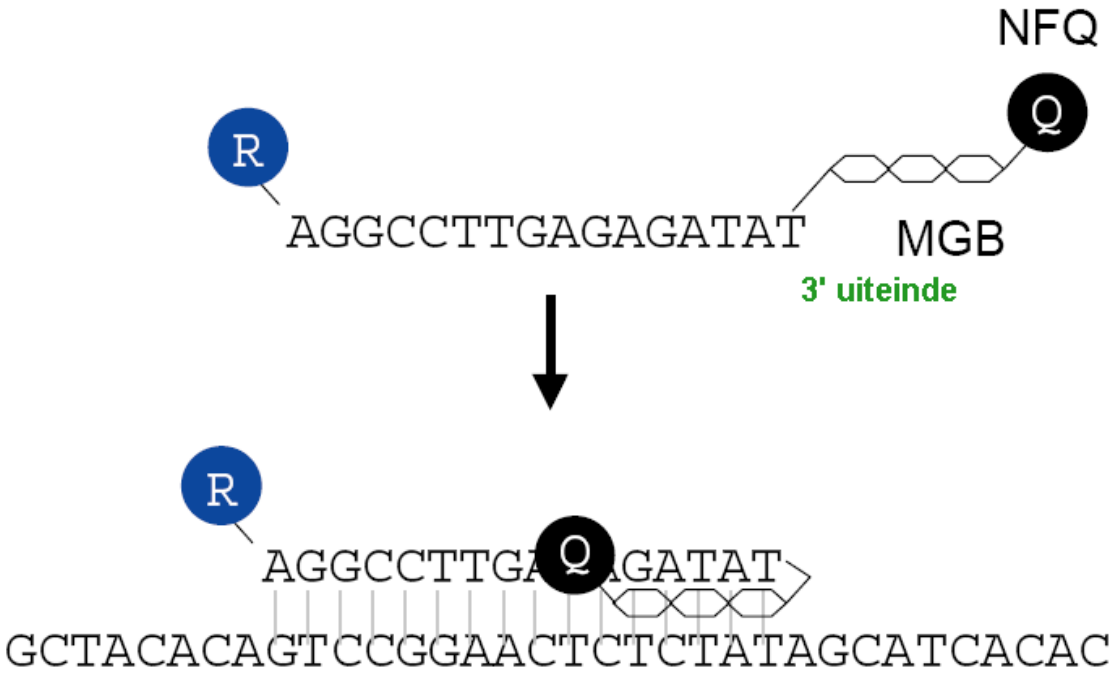
Questions for the audience…

- For what reasonTm probe is 5 degrees higher than Tm primers?
- The length of a probe should not be too long. Why not? The longer, the more specific?

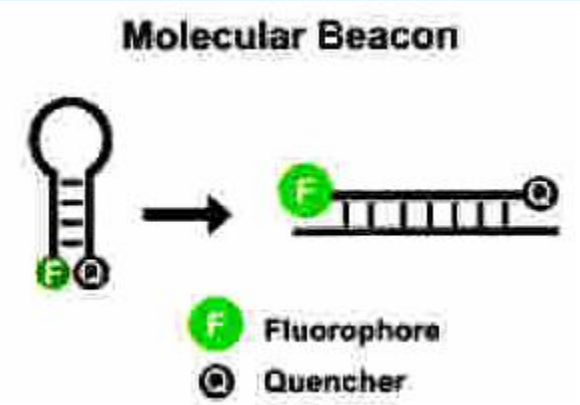
MGB Probe

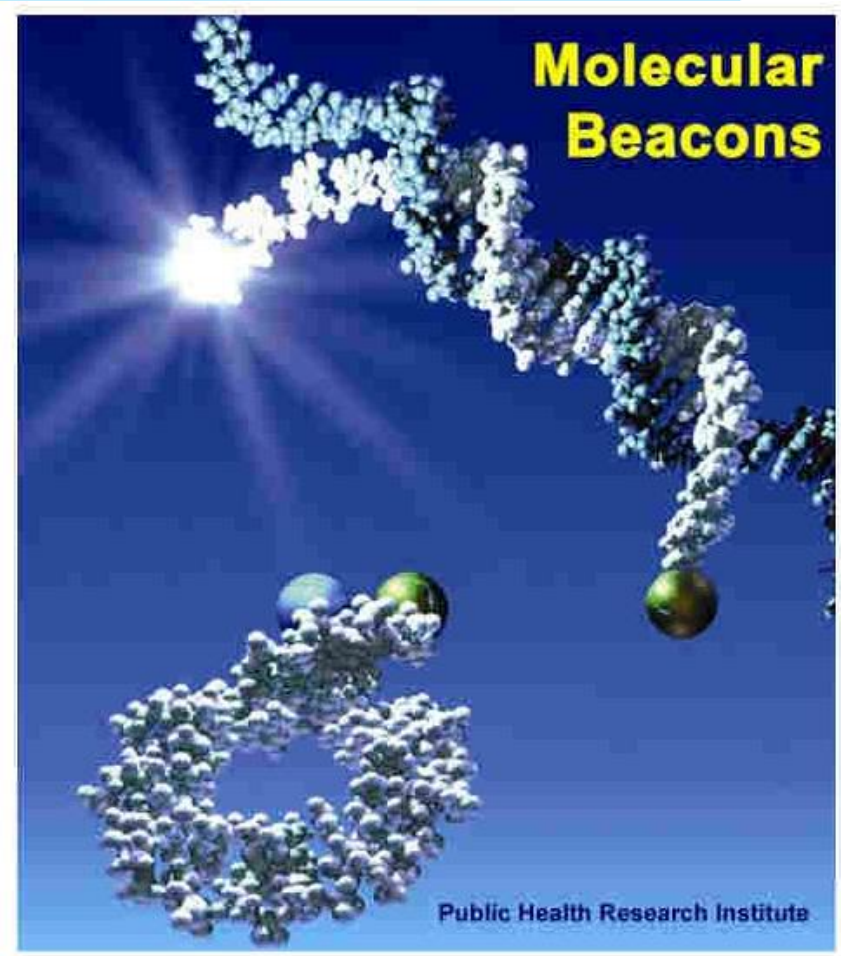


- Stabilizing oligo-target interaction
- Affinity MGB for MG DNA streng
  - increases T_m
- Shorter probes for highly variable targets (virus diagnostics)

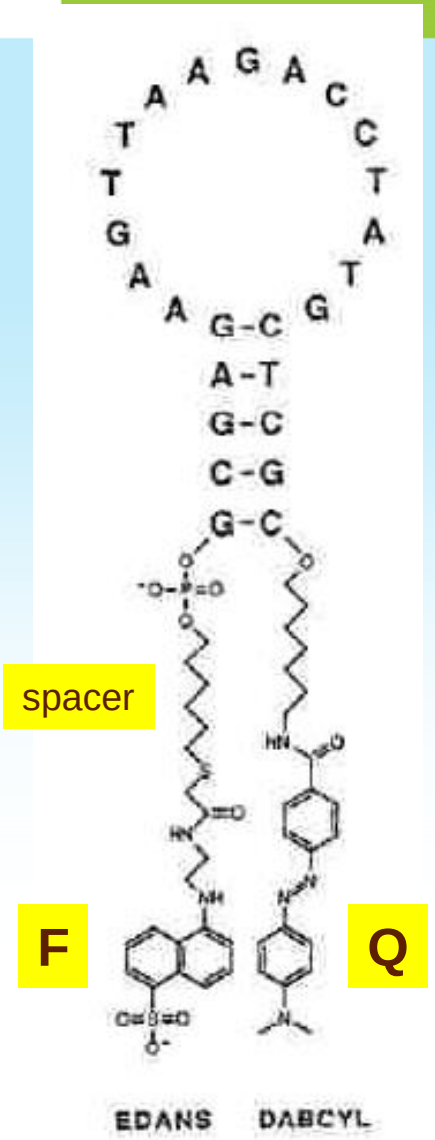


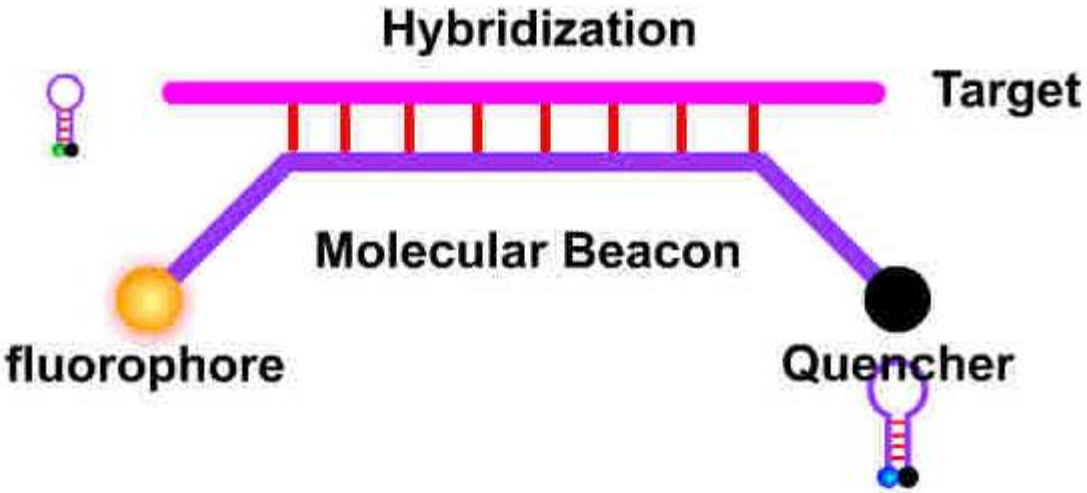
Hairpin probe





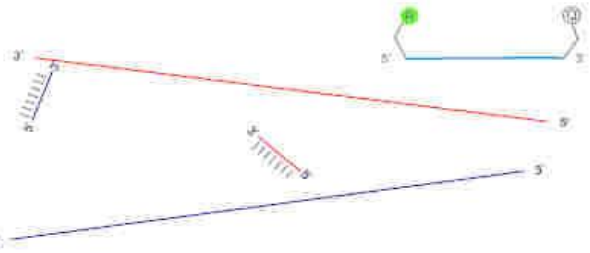
Quenching “molecular beacon”



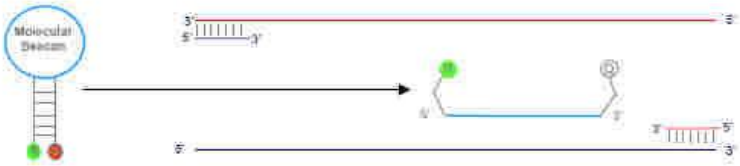


Principle hybridization molecular beacon

1. Denaturation



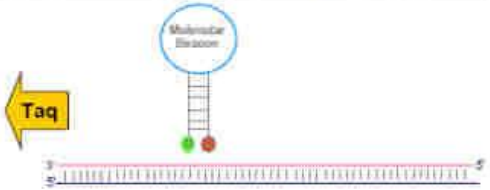
2. Annealing



3. Detection

4. Strand Displacement

5. Polymerization;
Complete, Probe Silent

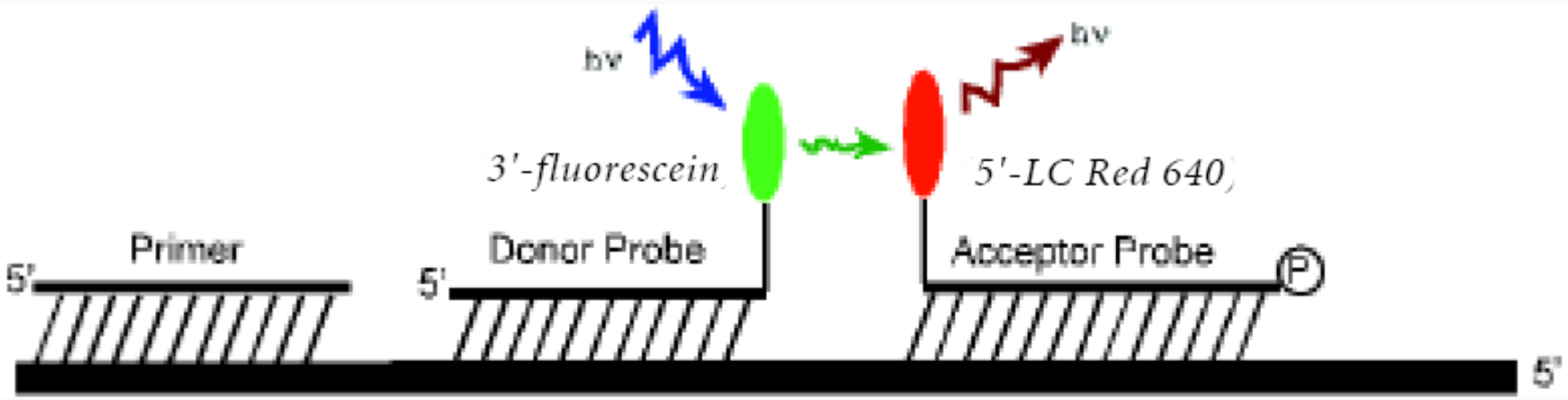


Design of Molecular Beacons

- Tm probe (loop) 5 - 10^{o}C higher than Tm primers

1. 3'and 5': 4 - 6 basen complementary: produces stem-loop
2. Used in a 2-step PCR program. Anneal and elongation at 60-70^{o}C
3. Often used for SNP assay: SNP in middle of probe
4. Dissociation-energy stem < dissociation-energy of highly complementary target,

Hybridization probes



Two reporters (fluorophors) no quencher

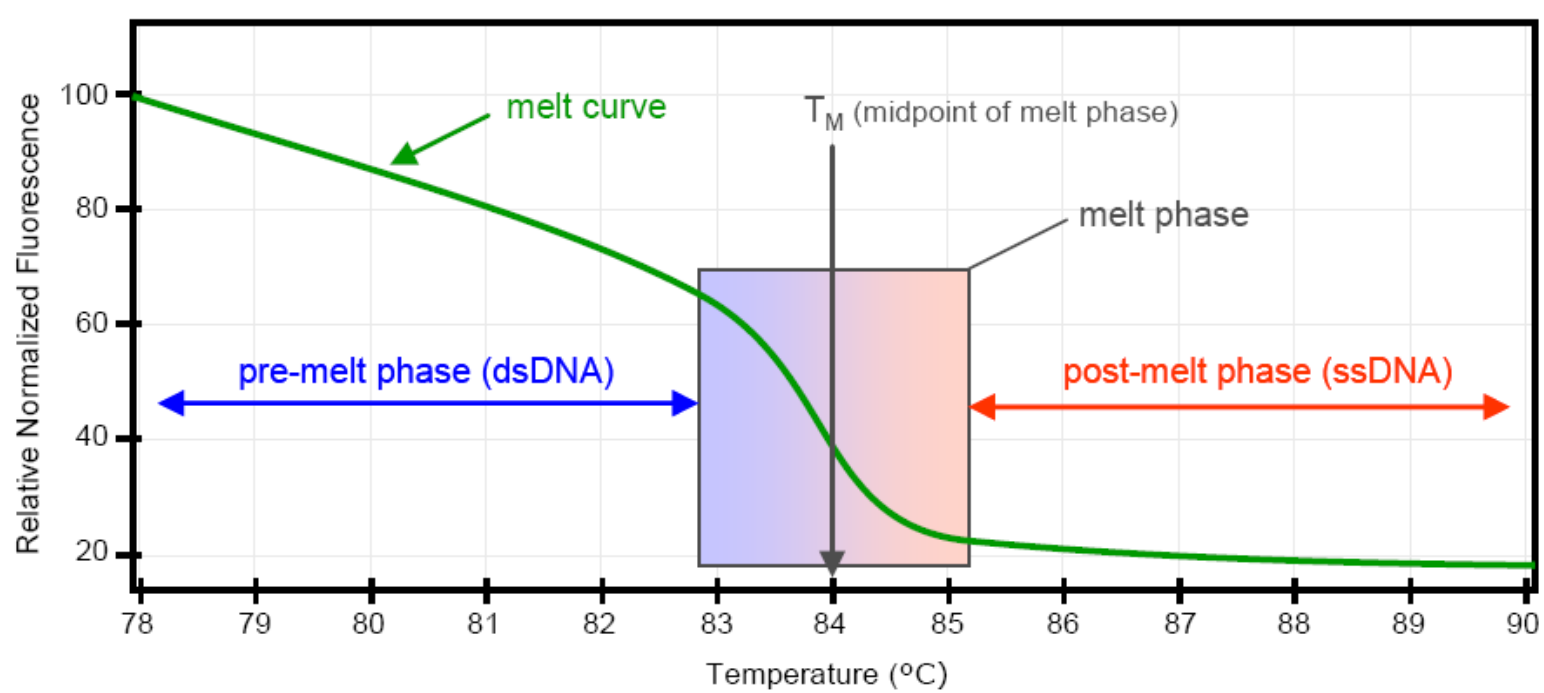
Very useful for genotyping and melting curve analysis

hogeschool Leiden

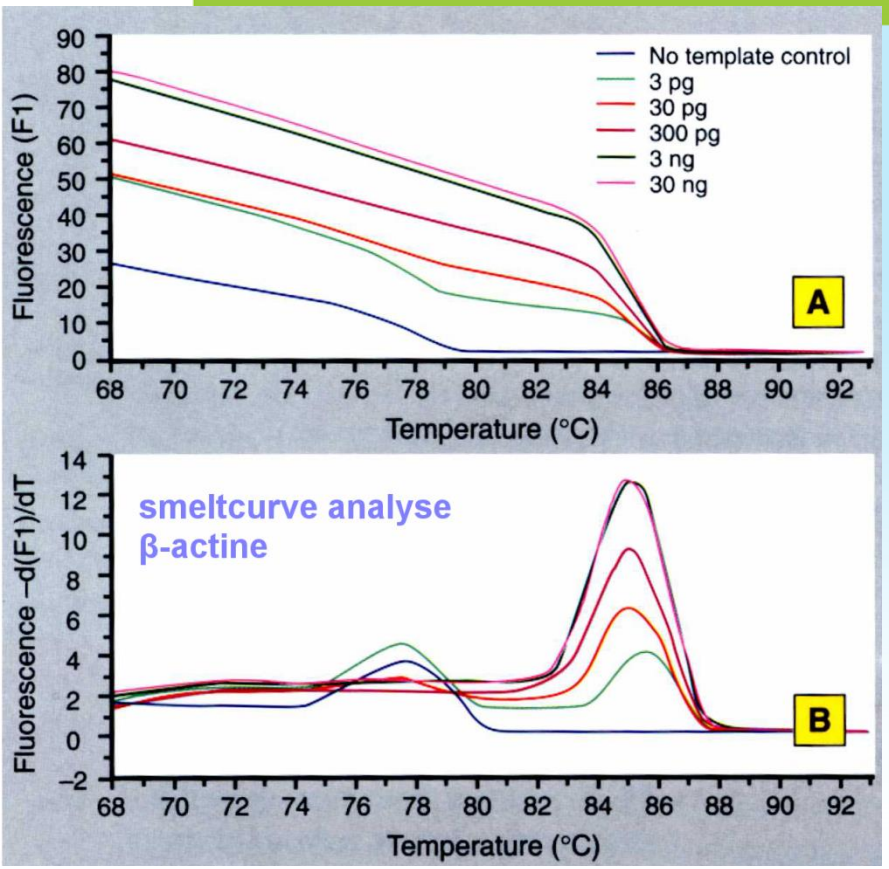
Melting curve analysis

TopLab

Melting curve



Melting curve analyse



- during qPCR
  - SYBR Green
  - intercalting in dsDNA PCR product
  - fluorescence
- after PCR
  - Slow melting PCR product
  - ds DNA (intercal) to ssDNA (no intercal)
- Analysis: decrease fluorescence by increasing temp.

Probe menu for qPCR method

	SYBR® Green 1	Dual-Labeled Probes	Dual-Labeled LNA® Probes	Molecular Beacons*	LightCycler® Probes*	Scorpions™ Probes*
APPLICATION						
Mass screening	••					
Microarray validation	••	•				
Multiple target genes / few samples	•	•				
SNP detection			••	•	•	••
Allelic discrimination			••	•	•	••
Pathogen detection	•	•	••	•	•	••
Multiplex		••	••	••	•	••
Viral load quantification		•	••	•	•	••
Gene expression	•	••	••	••	••	••
Gene copy determination		•	••	•	•	••
End point genotyping				••		••
In vitro quantification or detection			•	••		

Sigma Life Science

To conclude

- Primers
  - Rules for primer design: target sequence/ amplicon (length, GC%, hairpins,..)
- Probes
  - Rules for probe design
  - Fluorescence
  - Fluorophore and quenchers
  - Dual-labeled probes
  - Molcular beacons
  - Hybridisatie probes
- Melting curve analysis