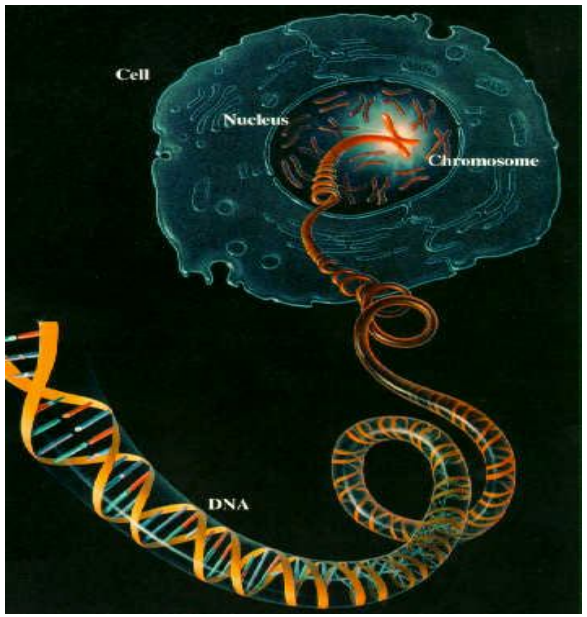
Molekylærbiologiske målemetoder:
Mange metoder og noen anvendelser



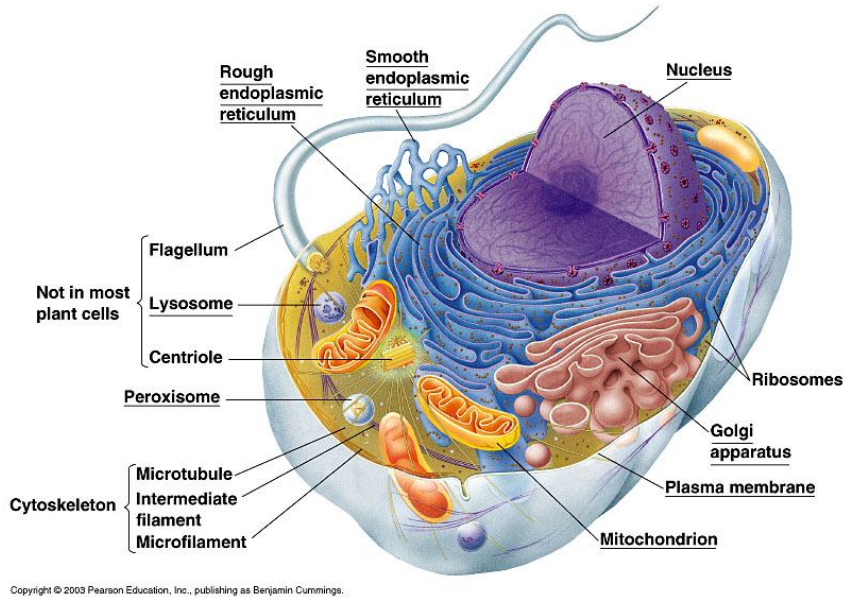
Kari Bente Foss Haug,
Seksjon for forskning
Avdeling for medisinsk biokjemi
Oslo universitetssykehus, Ullevål

Oslo universitetssykehus

HELSE SØR-ØST

Molekylærbiologi:

- Molekylære prosesser som danner basis for biologisk aktivitet
- Overlapper biologi, kjemi, biokjemi og genetikk
- Forstå av interaksjoner mellom de ulike systemer i en Celle
  - DNA
  - RNA
  - Protein



Copyright © 2003 Pearson Education, Inc., publishing as Benjamin Cummings.

Molekylærbiologi:

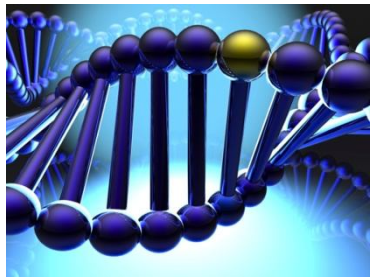
I grenselandet mellom

- Mikrokosmos (atomer og molekyler)
- Makrokosmos (ting vi kan se og ta på)

- Bruk av molekylærbiologiske metoder

→ Revolusjonert moderne biologi og medisin

forts Molekylærbiologi:



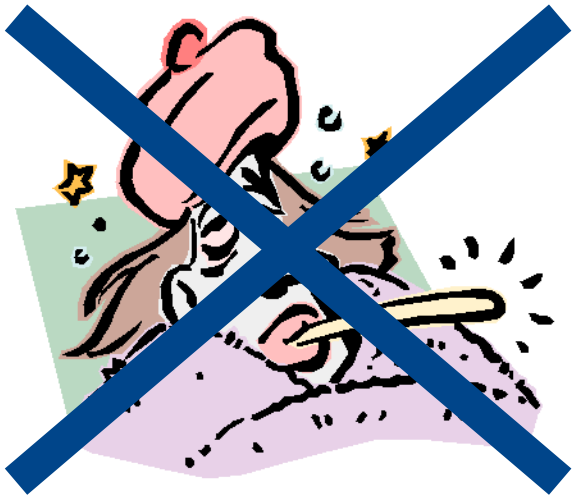
Genteknologi:

- Sett av metoder på DNA/RNA-nivå utviklet fra molekylærbiologi
- Sentralt verktøy i forskning og medisinsk diagnostikk

- Fokus på: GENTEKNOLOGISKE METODER

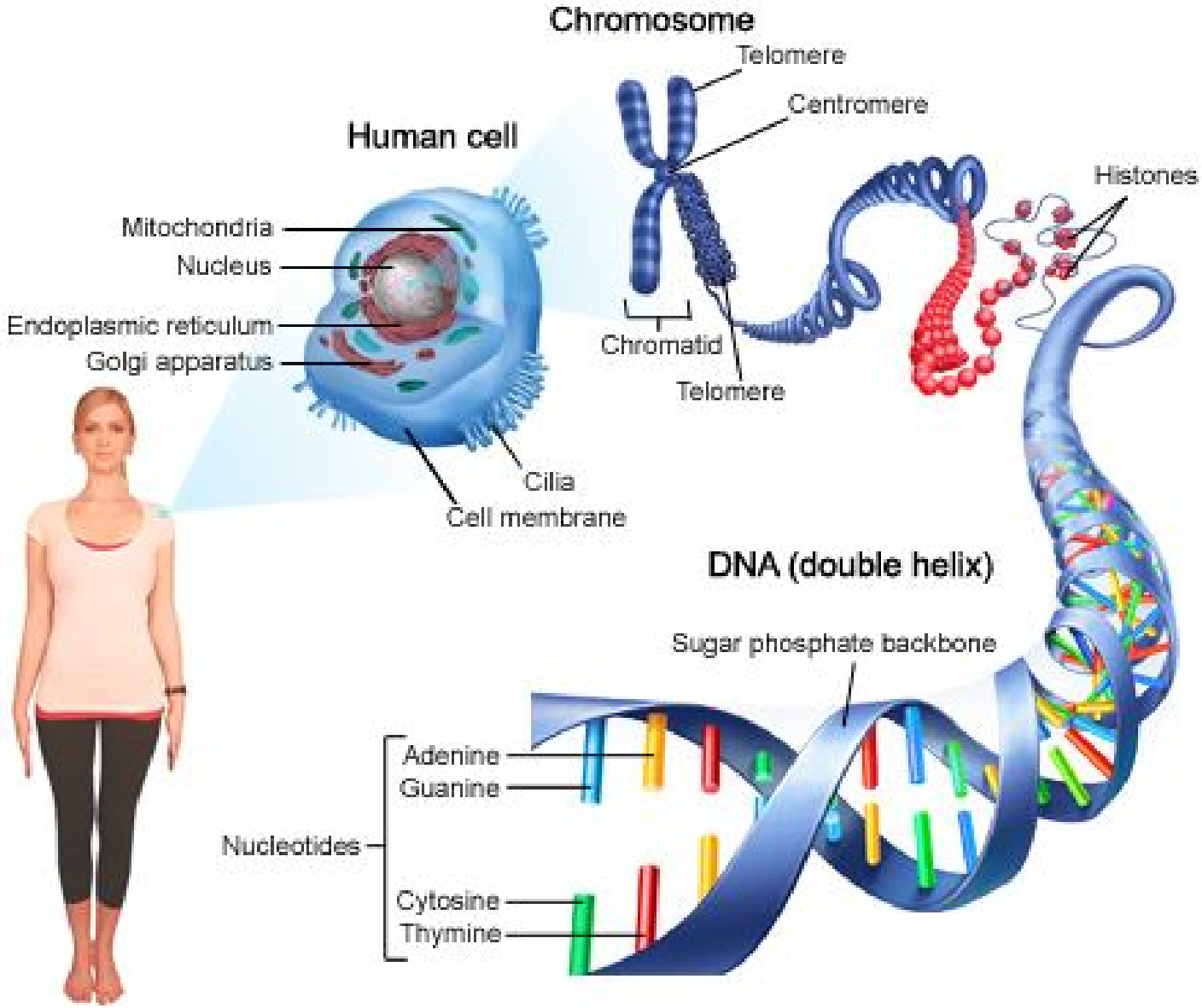
  – Bred anvendelse innen nyere medisinsk forskning og diagnostikk

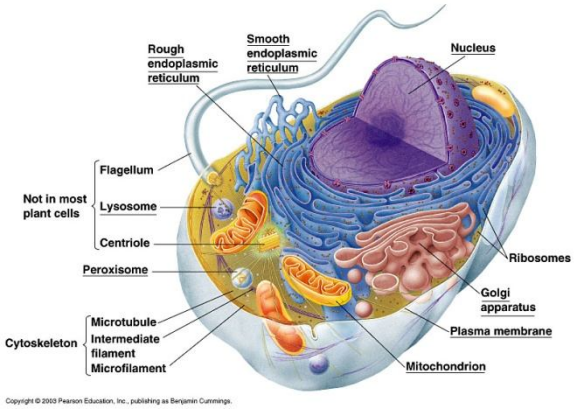
Molekylærbiologiske målemetoder i medisinsk diagnostikk

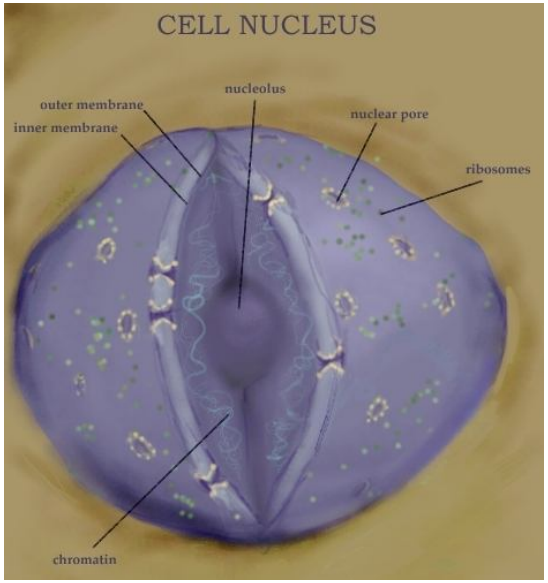


Disposisjon:

- Bakgrunnsinformasjon
  - Arvestoff og Gener
  - DNA og RNA
  - DNA-variasjon
  - Regulering
- Molekylærbiologiske metoder (Gentester)
  - PCR
  - RT-qPCR
  - Genkopitall
  - Sekvensering
    - Pyrosekvensering
    - Sanger sekvensering
    - Helgenom/Dyp sekvensering
    - ”Next generation sequencing”
- Eksempler


Chromosome
Telomere
Centromere
Human cell
Histones
Mitochondria
Nucleus
Endoplasmic reticulum
Golgi apparatus
Chromatid
Telomere
Cilia
Cell membrane
DNA (double helix)
Sugar phosphate backbone
Adenine
Guanine
Nucleotides
Cytosine
Thymine





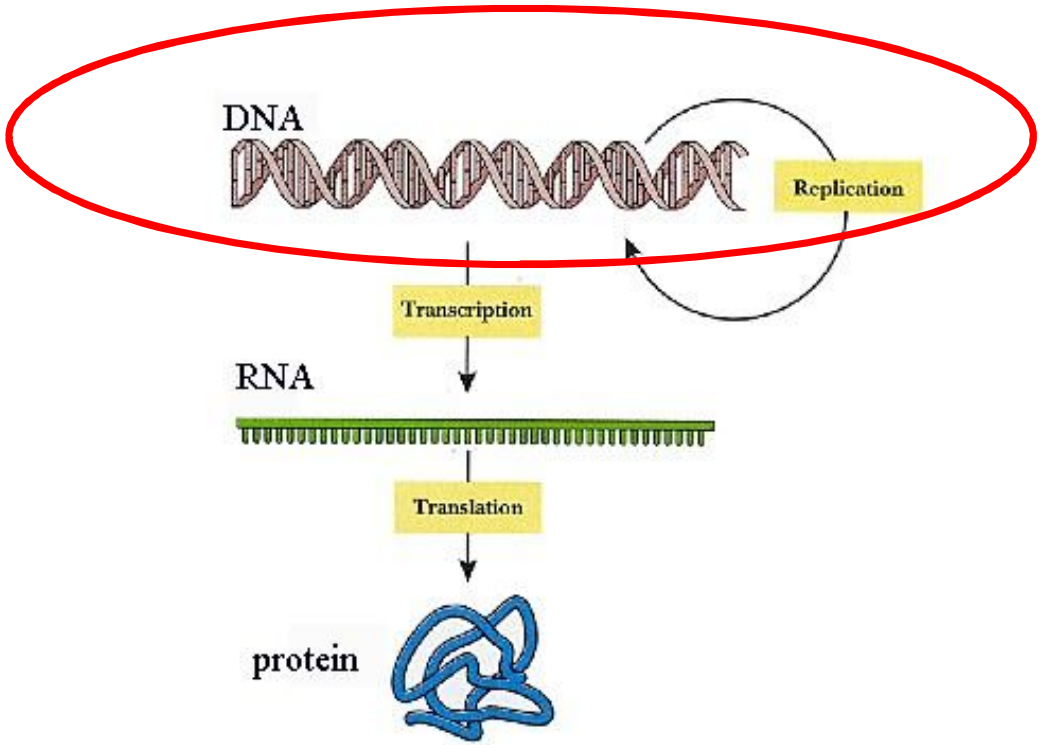
- ”Lang vei å gå”
- Cellens indre
- Cellekjernen
- Vårt innerste indre
- Oppskriften til ”Jeg´et”

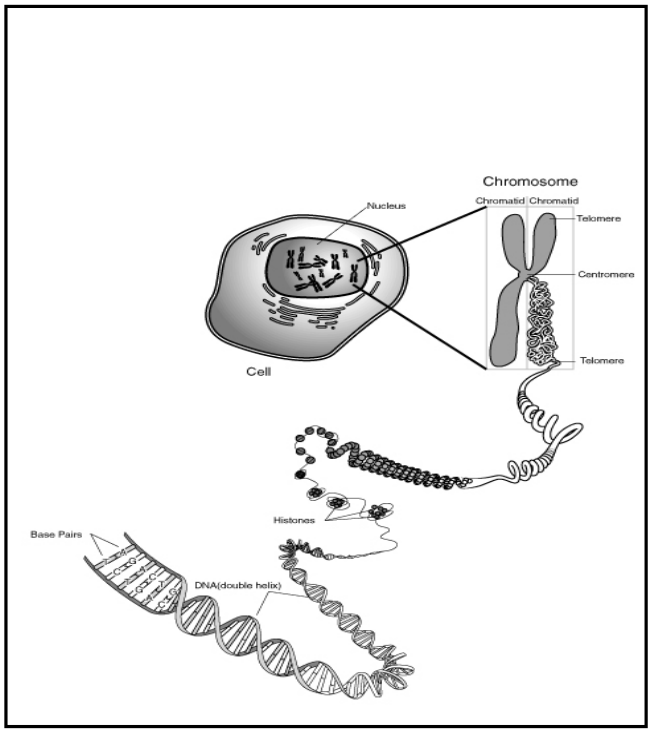
> 1 meter med DNA!!!!

Fascinerende ☺

- Skaperverkets organisering
- Mulig å utnytte kunnskapen om dette

Det sentrale dogmet:

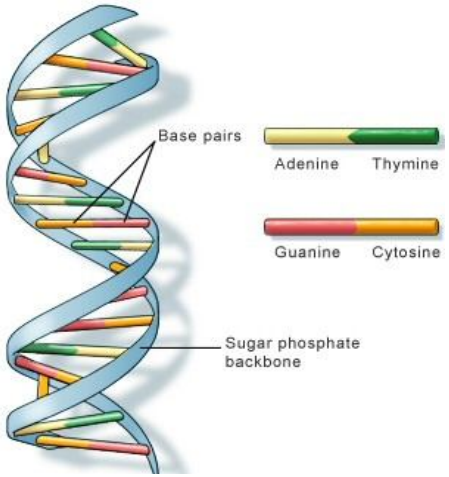


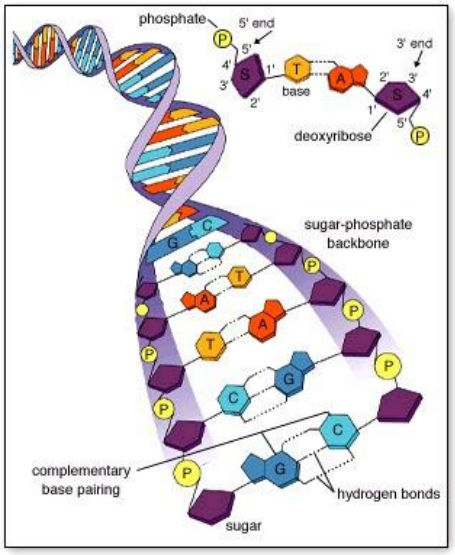


DNA:

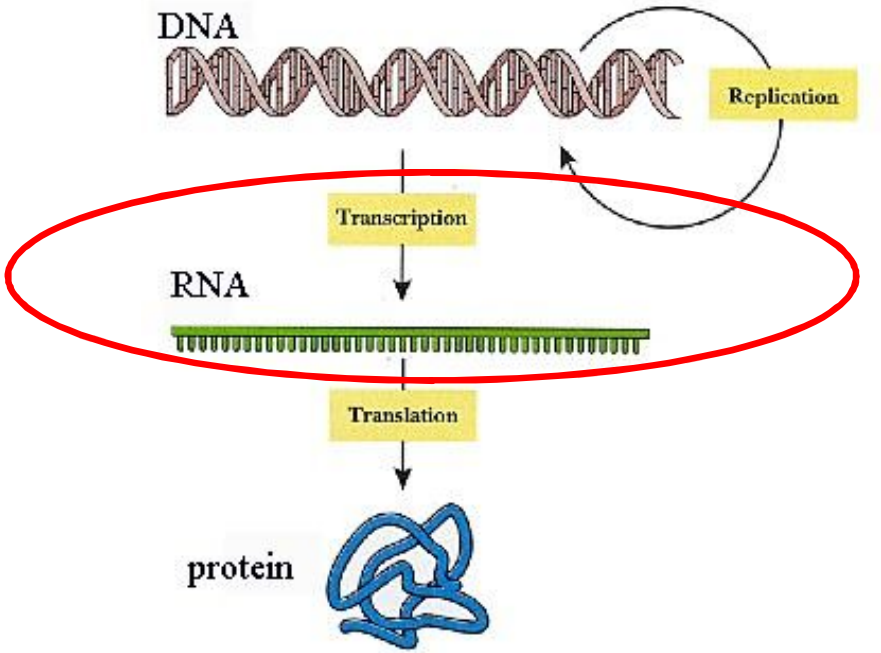
I cellekjernen
Arvematerialet = Alle genene ~ 22 000
Kokebok med oppskrifter for alle proteinene
Pakket på 4 nivåer
"Konstant"

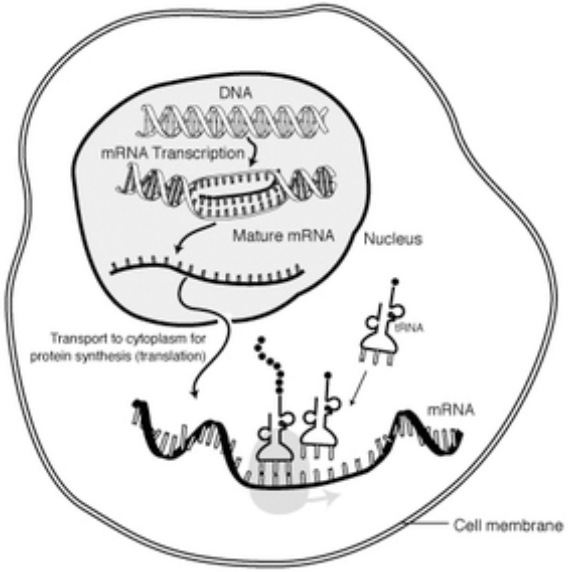
Dobbeltrådet (ds)
Orientert 5´- 3´retning
Antiparallelt
Sukker og fosfatkjeder
4 Ulike Baser (A, T, G, C)
Hydrogenbindinger





Det sentrale dogmet:

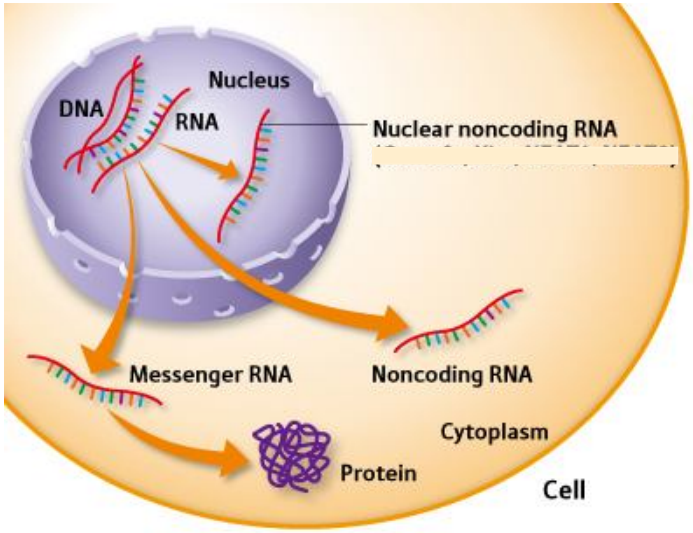




RNA:

I cellekjernen + cytoplasma
Kopi av et aktivt gen
Forløper til et protein
Regulert prosess
mRNA m.fl
(tRNA, rRNA, microRNA, siRNA, snoRNA + +)
Dynamisk RNA-bilde: Stor variasjon

Enkelttrådet (ss)
Orientert 5´- 3´retning
Sukker og fosfatkjeder
4 Ulike Baser (A, U, G, C)



Ny kunnskap om ikke–kodende RNA!

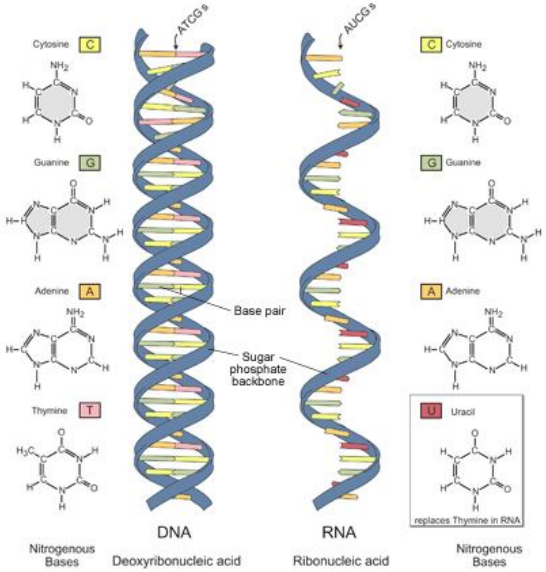
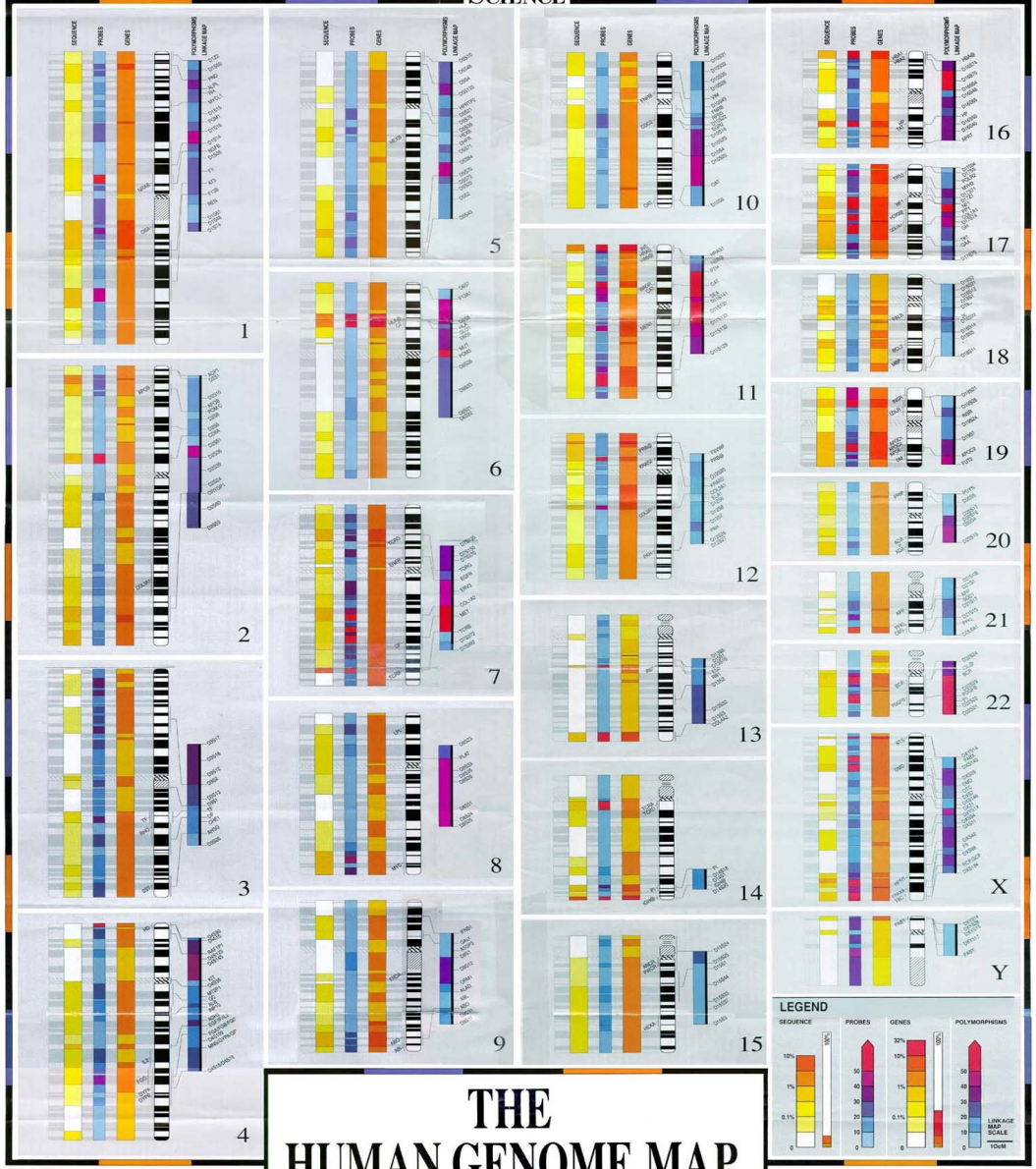


Image adapted from: National Human Genome Research Institute.



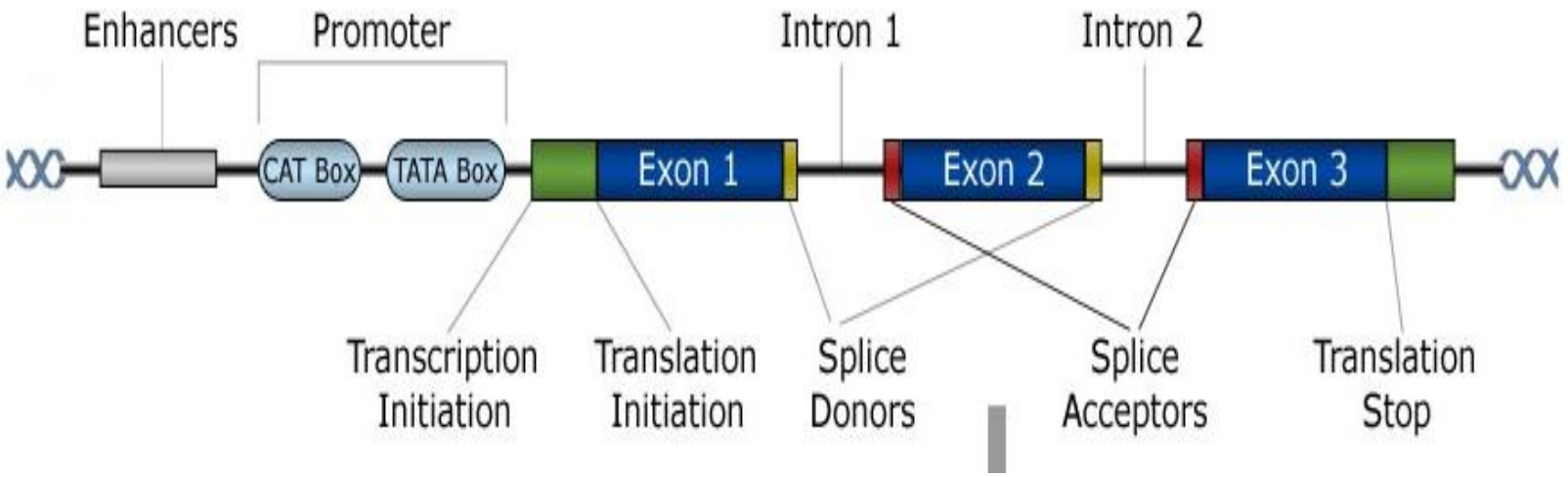
THE HUMAN GENOME MAP 1990

Results of efforts to map and sequence the human genome, either published or in press by July 31, 1990, are depicted in the bars surrounding the ideogram of each of the 24 chromosomes. Each type of bar is labeled at the top of the wall chart. The data are illustrated as colored intervals according to logarithmic or linear scales, as shown in the legend. Further information, including a

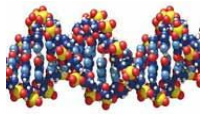
in centimorgans for loci selected from published maps to give an illustrative map with a resolution averaging 10 to 20 cM. The POLYMORPHISMS bar indicates the number of known polymorphic loci allocated to the cytogenetic intervals formed by the linkage map. Lines connect the bar to the ideogram for those loci with a precise cytogenetic map localization.

Representative genes are identified by their alpha-

Oppbygging av et gen



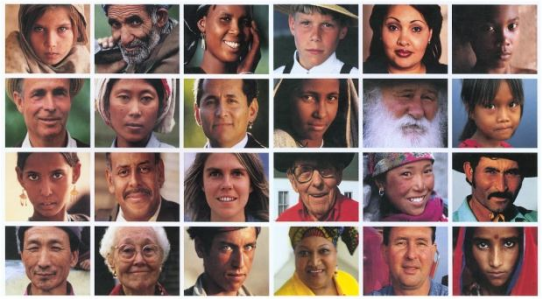
DNA:



DNA: Definerer et unikt menneske
DNA: Inneholder all nedarvet informasjon

Estimat:
99, 9 % av DNA-mengden ser ut til å være lik i alle folkeslag !!!

Kun 0,1 % variasjon på DNA-nivå gir store individuelle, fenotypiske forskjeller



Forandring / Variasjon på DNA-nivå:

Kan påvirke:

Kvalitet / kvantitet av alle virksomme proteiner

→ **Endret proteinstruktur / Endret proteinaktivitet**

→ **Utfall i ulike retninger (⇅)**

Effekt av DNA-variasjon er avhengig av:

- **Type DNA-variasjon**
- **Hvilket gen DNA-variasjonen er lokalisert til**

Balansegang

DNA-variasjon kan føre til:

Normalvariasjon

Sykdom

DNA – variasjon:

- Feil under DNA-replikasjon
- DNA-skade

- Mutasjoner
- Enkeltbasesubstitusjoner (SNP)
- Delesjoner/Insersjoner (korte/lange)
- Kopitall (genkopitall)
- VNTR (Variable number of tandem repeats) - Finnes hos alle
  - Mikrosatelitt (Di, tri, tetra, eller pentanukleotid repeats
  - Minisatelitt (> 5 repeats)
- Alternativ splicing
- Transposone elementer
- Metyleringsgrad
- Tidsavhengige endringer (Telomerase)

DNA-variasjon:

Brukes for å indikere at det eksisterer mer enn én DNA sekvensform i populasjonen

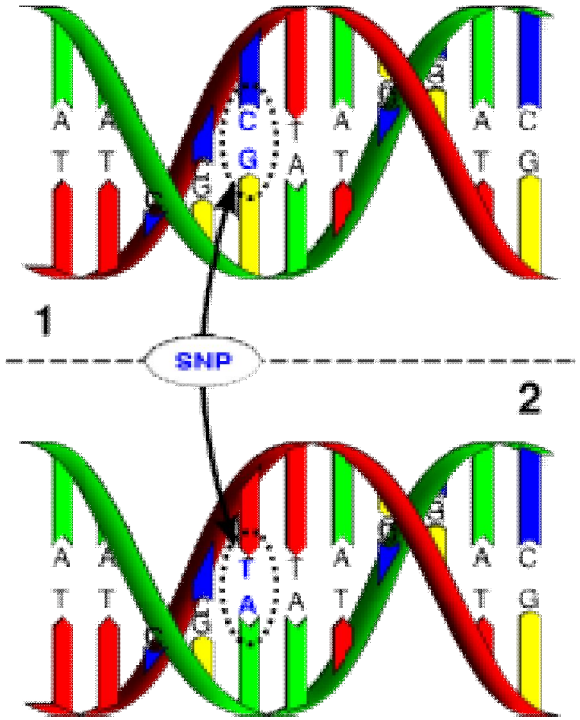
Polymorfisme (variant)

- **Polymorphos = Multiform (gresk)**
- **Oppstått gjennom evolusjonen som konsekvens av mutasjoner**
- **Bærerfrekvens i en populasjonen: > 1%**
- **Brukes for å indikere at det ved en gitt posisjon i DNA eksisterer mer enn en sekvensform i populasjonen**
- **En person vil være bærer av en eller toppen to varianter**
- **En polymorfisme / variant for ca hver 300 basepar**

SNP= Single Nucleotide Polymorphism (”snip”)
SNV= Single Nucleotide Variant

Variasjon i ett enkelt nukleotid – i én base
Kan gi opptil 3 nukleotidvarianter i populasjonen

1. **Homozygot villtype / Homozygot negativ (-/-)**
2. **Homozygot mutant / Homozygot positiv (+/+)**
3. **Heterozygot (+/-)**



SNV-klassifisering:

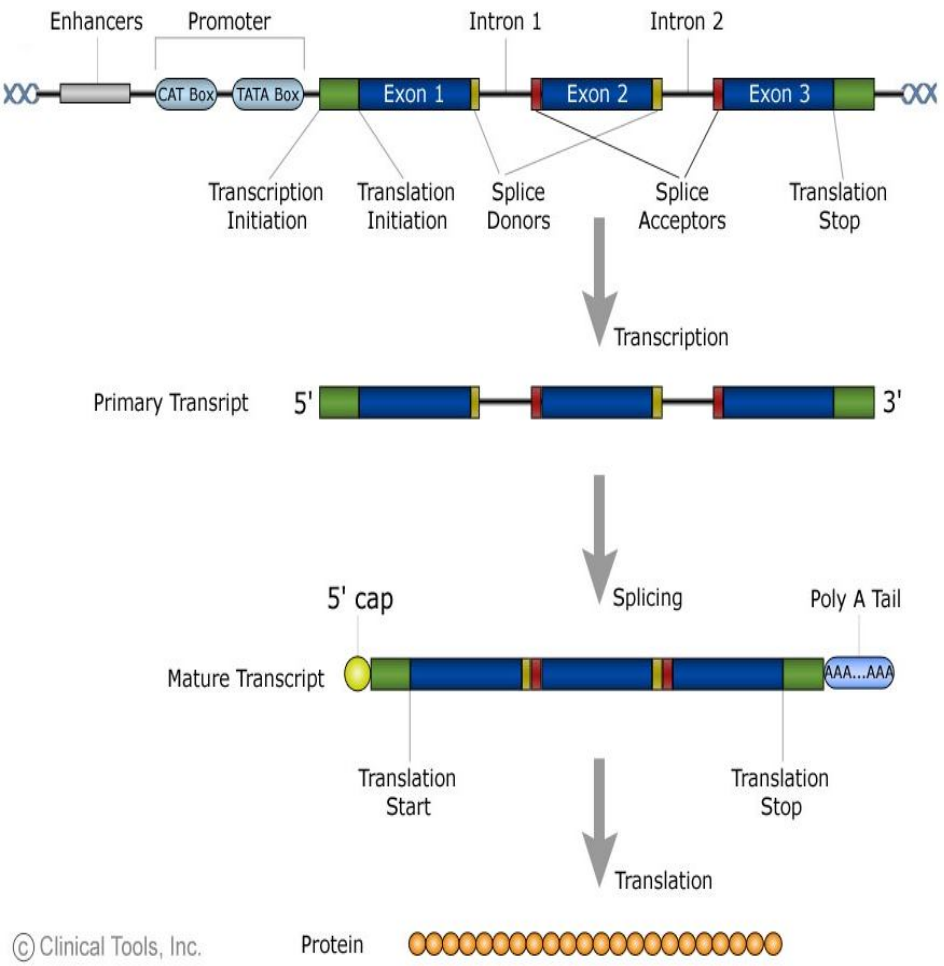
Noncoding SNVs:

Endrer basesekvensen i den delen av DNA som ikke koder for protein (Promoter, intron, 5´og 3´-region, intergenic)

Endring i:

- **Transkripsjonsnivå**
- **Prosessering**
- **Stabilitet**

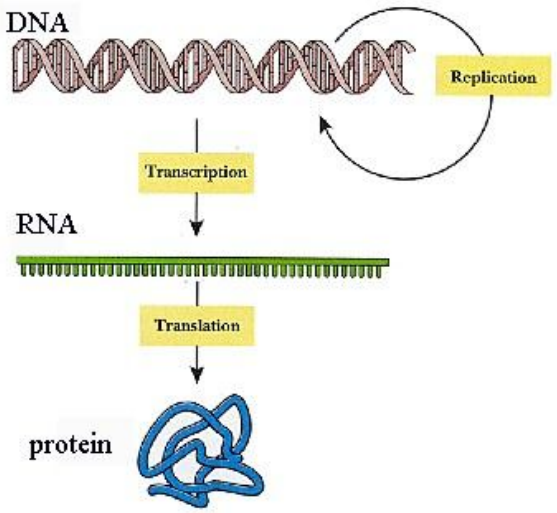
} **Endret RNA nivå (Protein)**



Coding SNVs:

Endrer basesekvens i den delen av DNA som koder for protein

- **Synonym: Endrer ikke Aminosyre, Protein uforandret**
- **Ikke synonym: Endrer Aminosyre, Protein kan endres**



DNA-variasjon

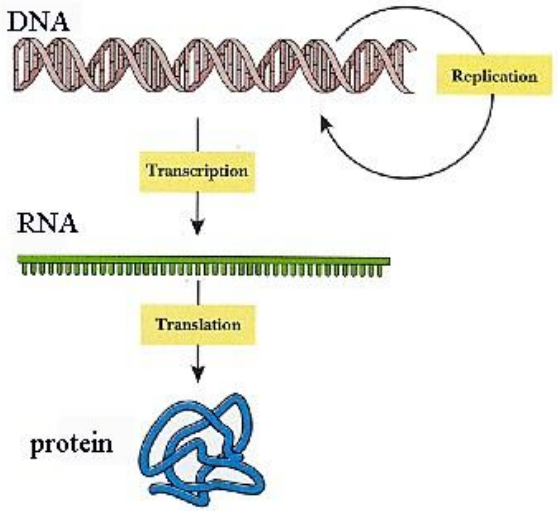
- Sekvensforandringer på DNA-nivå
- Overføres via RNA-nivå
- Effekt på protein-nivå

DNA-variasjon kan føre til:

- Redusert eller økt nivå av ”korrekt” protein
- Endring av proteinstruktur til et ”ikke korrekt” protein

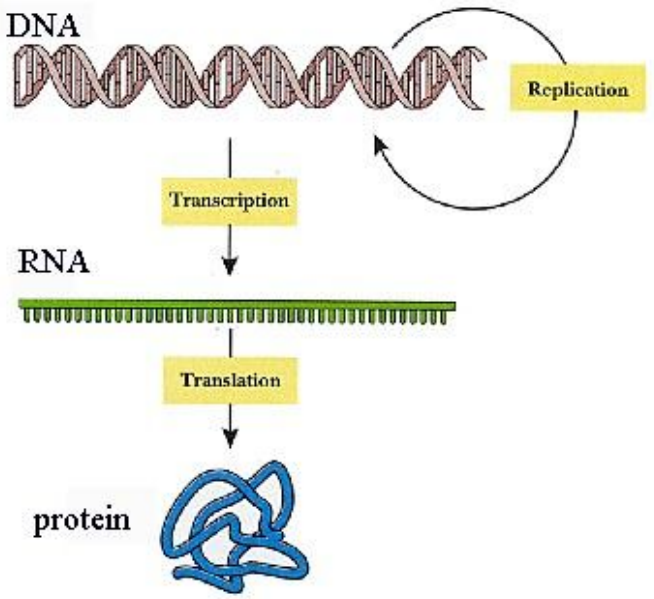
”Tap av funksjon” (Loss of function)
”Funksjonsøkning” (Gain of function)

- Ulike mekanismer for slike prosesser



Forandring på DNA-nivå

- **DNA-analyser**
- **RNA**
- **Protein**

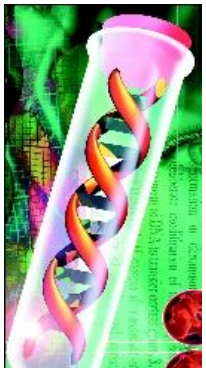


Forandring på DNA-nivå

Genetiske tester:

- **Diagnostiske**
- **Presymptomatiske**
- **Prediktive**

Hvorfor er det ønskelig å ta i bruk en genetisk test?



Mange årsaker:

- Flytte diagnostikk fra protein-nivå til gen-nivå
  - Problemer med opprinnelig analyse
  - Opprinnelig metode er tidkrevende
  - Opprinnelig metode gir ikke full utredning av problemstilling
  - Komplettering av diagnostisk verktøy

- Revolusjonerende medisinsk nyhet publisert i Nature!
  - Sykdom assosiert til DNA-variasjon / funnet mekanismen
  - Mulighet for engangsanalysering

- Monitorere varianter i legemiddelmetaboliserende enzymer (farmakogenetikk)
  - Rapportert om problemer med bruk av en gruppe legemidler
  - Vanskelig å innstille legemiddeldosering
  - Bivirkningsproblematikk

- Forskningsprosjekt

Valg av genetisk locus:

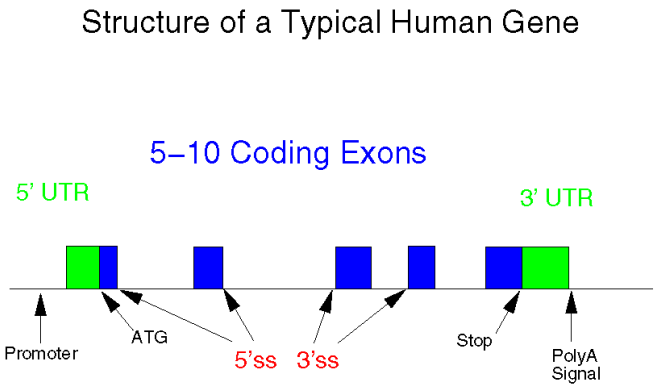
- Sykdom assosiert til
  - Enkeltkomponent
    - Høy penetranse (Duchenne muskelatrofi, Familiær hypercholesterolemi m.fl)
  - Multikomponent
    - Lavere penetranse (Diabetes, hjerte/kar, psykiatriske lidelser m.fl)
- Aktiv komponent i sykdomsprosessen
- Biomarkør

Behov for opplysninger om:

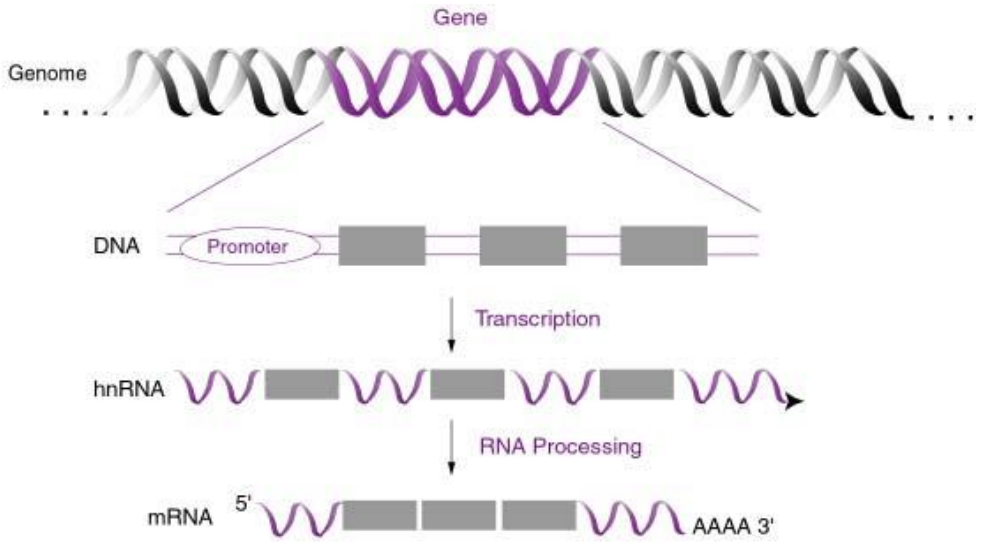
- **Hvilken DNA-variasjon det er snakk om**
  - **SNP, delesjon/insersjon, repeat, genkopitall etc**

- **Informasjon om genet**
  - **Beliggenhet**

- **Hvordan genet er bygget opp**
  - **Promoter, Exon, Intron**

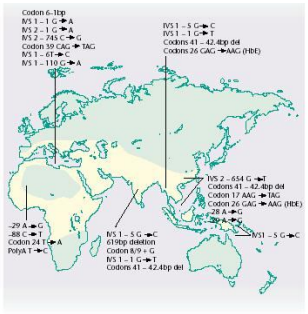


RNA:


Gene
Genome
DNA
Promoter
Transcription
hnRNA
RNA Processing
mRNA
5'
AAAA 3'

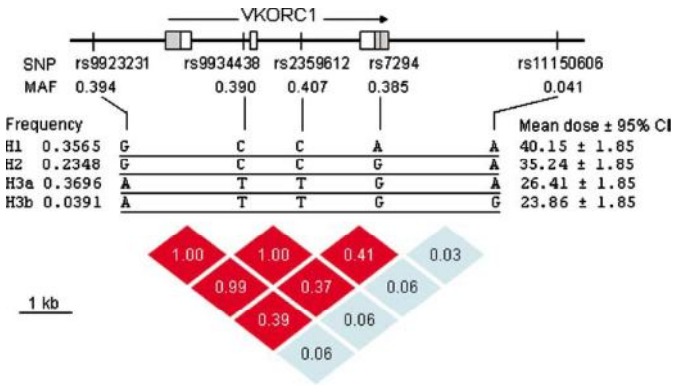
- **Informasjon om subtyper med høy homologi**
- **Informasjon om pseudogener**

- **Informasjon om bærerfrekvens i ulike populasjoner**



World distribution of the β thalassaemias. Each population has a different set of mutations. These are described either by the nucleotide base position in introns (IVS 1 or 2) or in the particular codons in exons. Mutations that are given the prefix are those in the 5′ noncoding regions of the β globin genes. Those marked polyA are mutations in the 3′ noncoding regions. bp, Base pair.

- **Finnes det haplotyper?**



- **Hvilken penetranse (gjennomslagsskraft) har mutasjonen?**

- **Er det publisert metodeartikler?**

- **Er publiserte metoder kompatible med eget utstyr i lab´en?**

Trinn i genetisk test:

Oftest på DNA-nivå

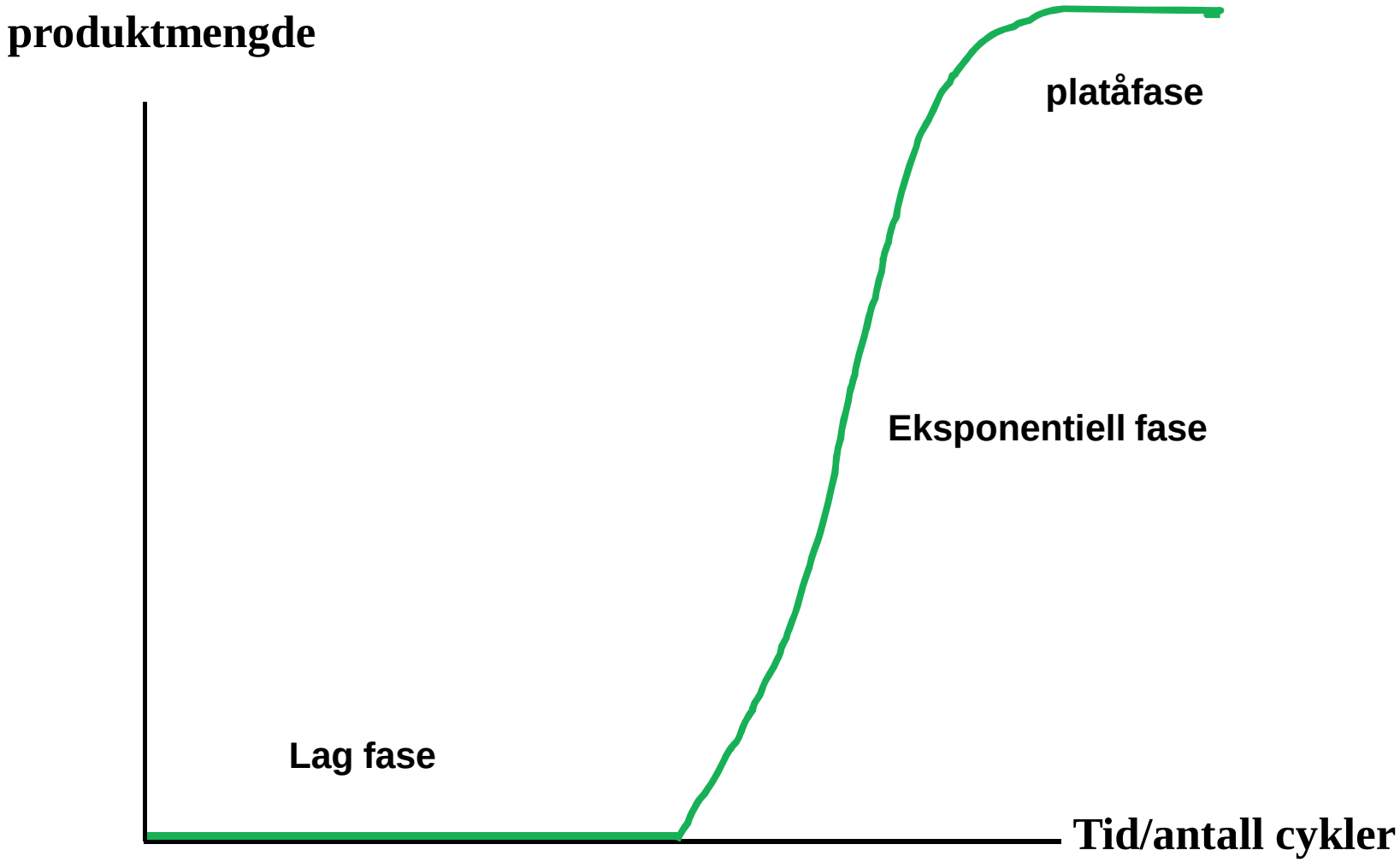
- **Lokalisere DNA-område med kjent variasjon**
- **Isolere DNA / RNA (Fullblod, Vev, Ulike Kroppsvæsker m.m.)**
- **Mangfoldiggjøre et kortere DNA-fragment rundt kjent variasjon (PCR) / RT- qPCR**
- **Bestemme genotype**
  - **Avlese genetisk variant i et gitt locus vha ulike metoder**
  - **Kvantitere RNA-nivå / RNA fusjonsmolekyler**

Polymerase Chain Reaction (PCR):

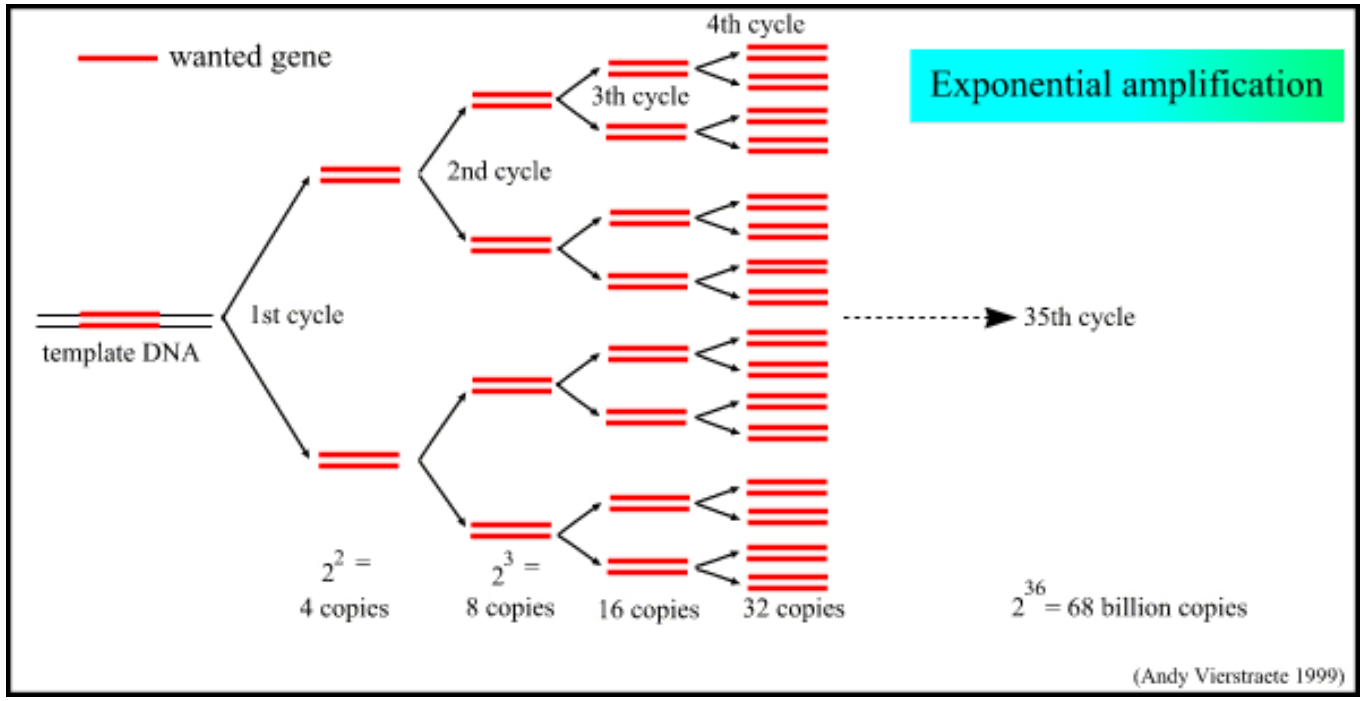
- **PCR**
  - **Kvalitativ** (Avleser DNA-variant - bestemmer Genotype)
  - **Kvantitativ** (qPCR - Avleser mengde PCR-produkt)
    - Større krav til PCR-metodikken
    - Referansegener
    - Kalibratorgener
    - Langsgående kontroll

- **RT-PCR** (Reverse transkriptase-PCR)
  - Benytter RNA som templat
  - Omdanner RNA til cDNA før PCR
  - Kvantitativ: RNA-ekspresjon av et normalt RNA/ RNA-fusjonsprodukter (Bcr-Abl)
  - Større krav til PCR-metodikken

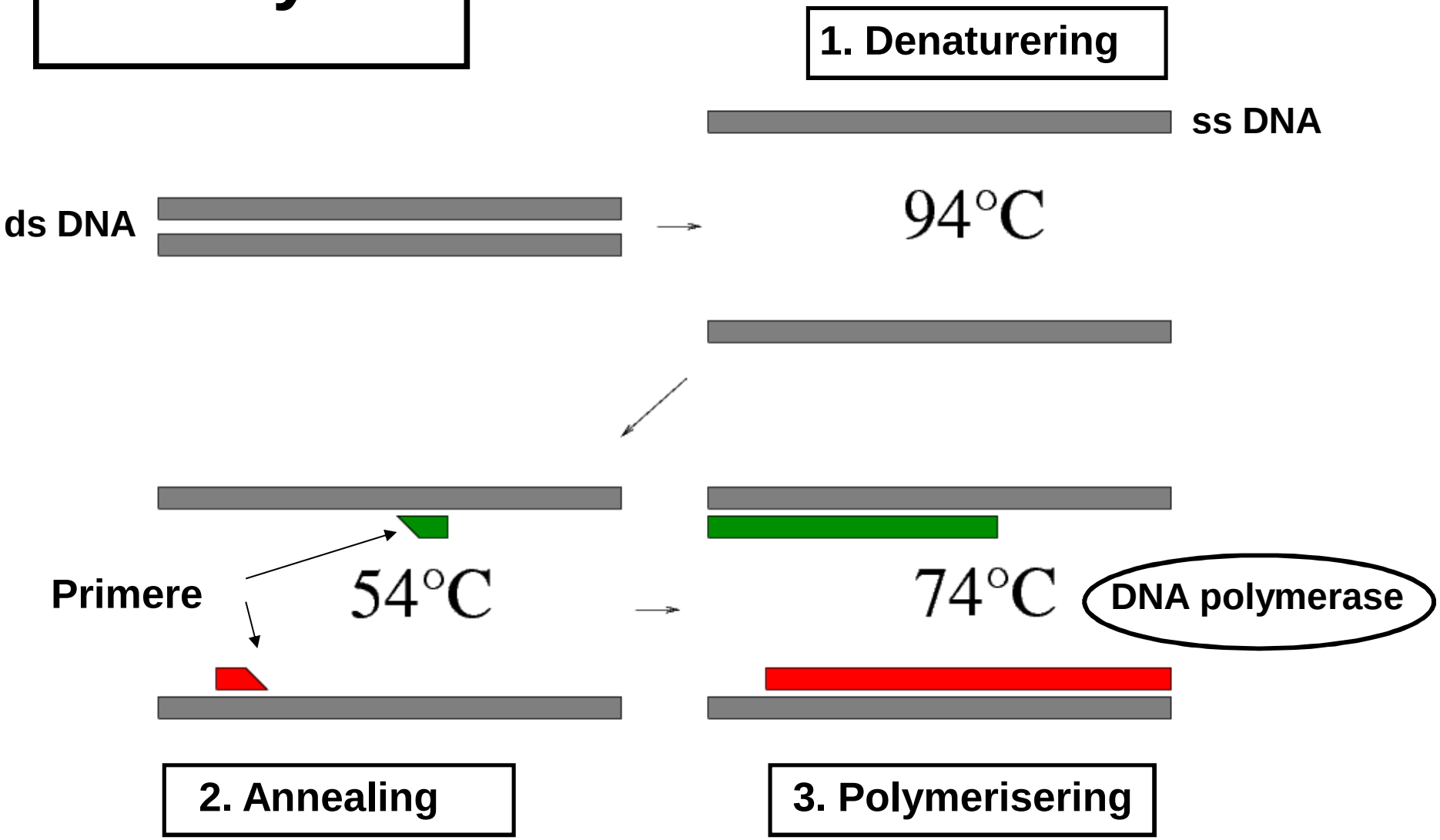
PCR-kurve


produktmengde
platåfase
Eksponentiell fase
Lag fase
Tid/antall cykler

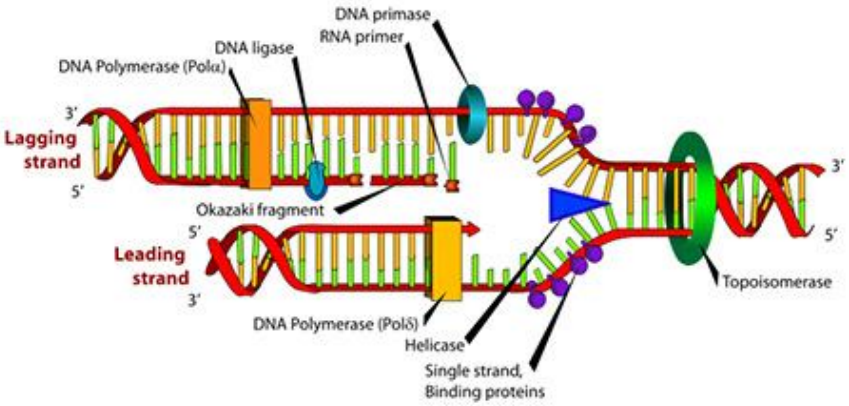
PCR:


wanted gene
4th cycle
Exponential amplification
3th cycle
2nd cycle
1st cycle
template DNA
35th cycle
2^2 =
4 copies
2^3 =
8 copies
16 copies
32 copies
2^{36} = 68 billion copies
(Andy Vierstraete 1999)

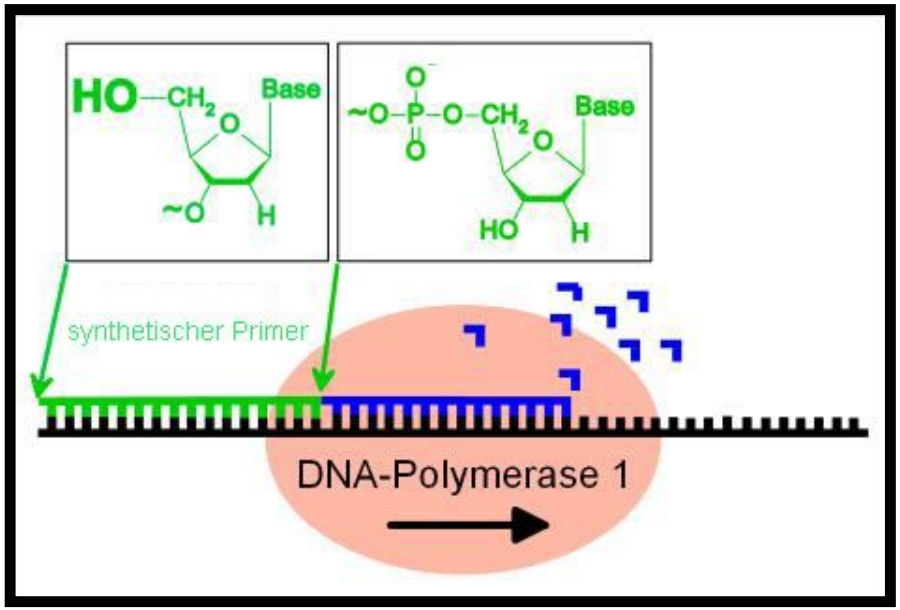
PCR-cycle:


1. Denaturering
ss DNA
ds DNA
94°C
Primere
54°C
74°C
DNA polymerase
2. Annealing
3. Polymerisering

DNA-polymerisering (in vivo):



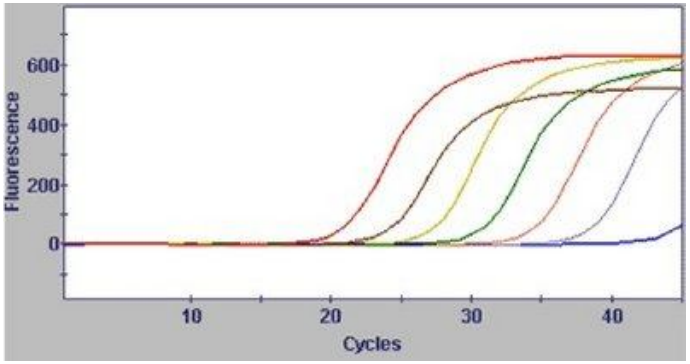
PCR-polymerisering:



Hva trengs for å kjøre en PCR?

- DNA
- Primere
- DNA-polymerase
- Nukleotider (dATP, dTTP, dGTP, dCTP)
- Ulike salter (Mg $^{2+}$ m.m.)
- PCR-instrument med temperatur-regulator

- Teknikk for å avlese resultat



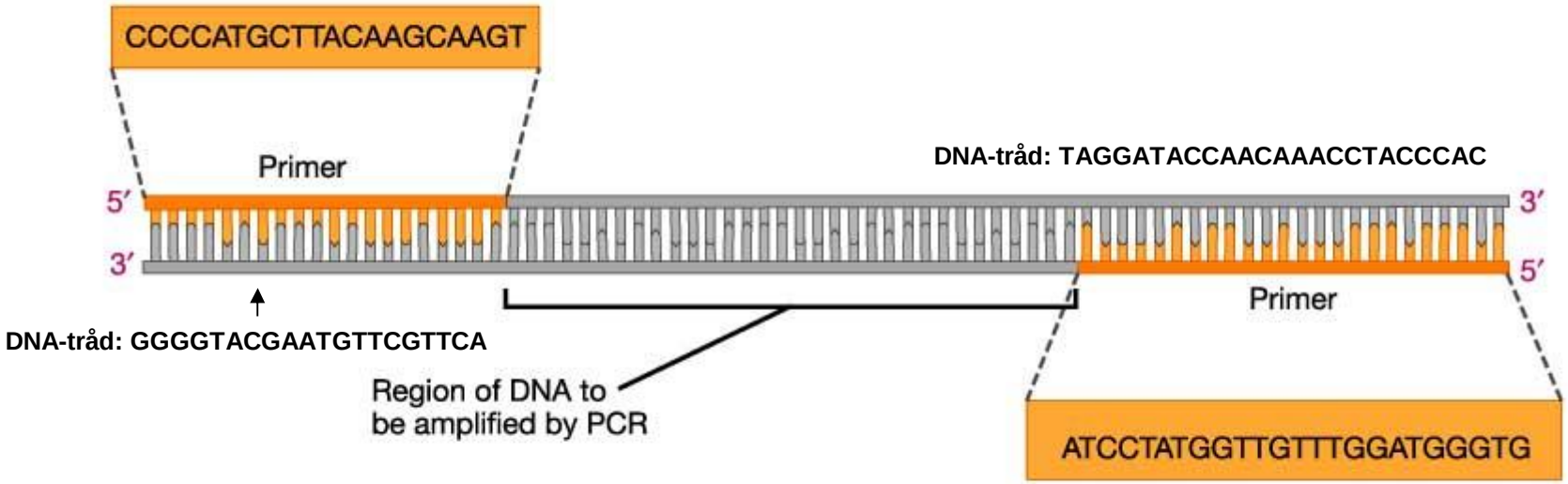
- **Antall PCR-cycler varierer (40 – 55)**

- **Før cyclene starter:**
  - **Enzymaktiveringsstep**

- **Etter cyklene er avsluttet: Ulike tilleggsaktiviter**
  - **Kjøling**
  - **Smeltekurver**

Primere:

- **Små (15-25 bp)**
- **Syntetiske DNA-biter**
- **Komplementære til en gitt DNA-sekvens**



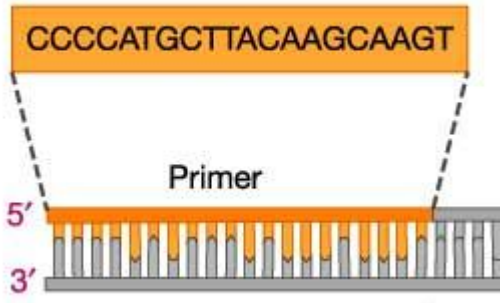
1. **Forward primer / Left primer (parallell)**
2. **Reverse primer / Right primer (antiparallell)**

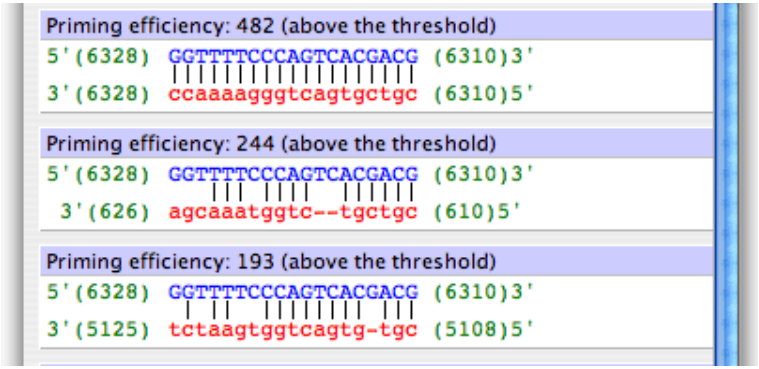
➡ **Søker etter komplementær DNA-strekk å binde seg til**

Binding av primer:

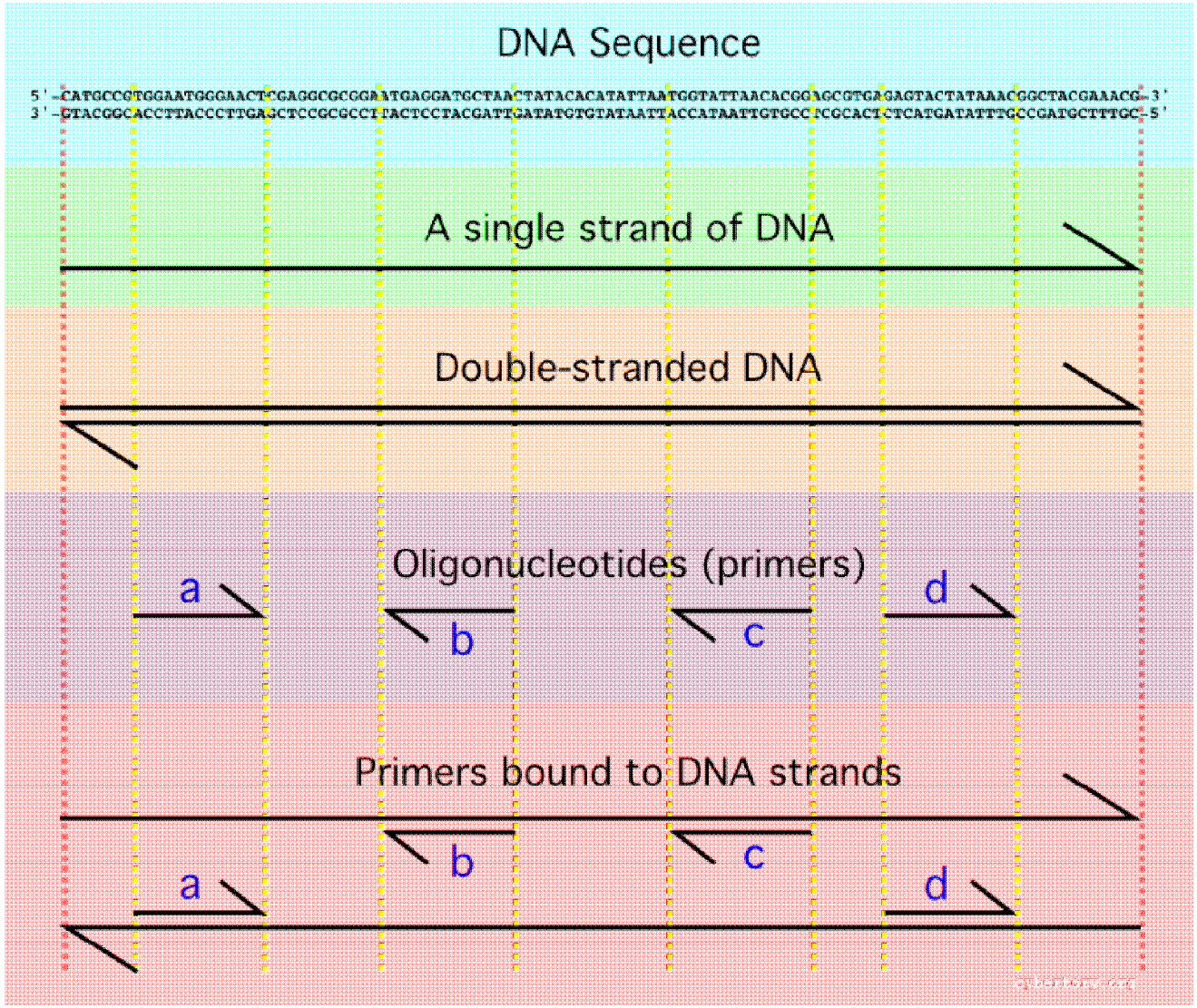
- **Temperatur**
- **[Salt]**
- **[Primer]**

- **Kan ”styre” primerens evne til å binde DNA (Stringens)**
  - **kun de sekvenser som er 100 % homologe**
  - **sekvenser med bare delvis homolog sekvens**



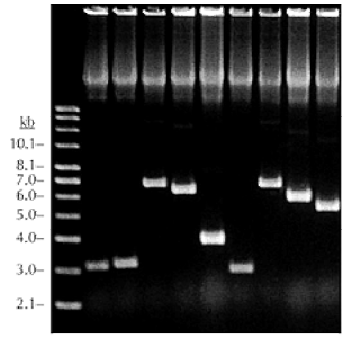


Arrow Notation for DNA

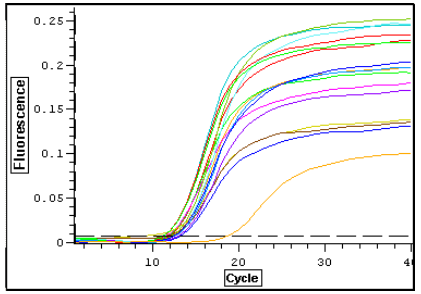


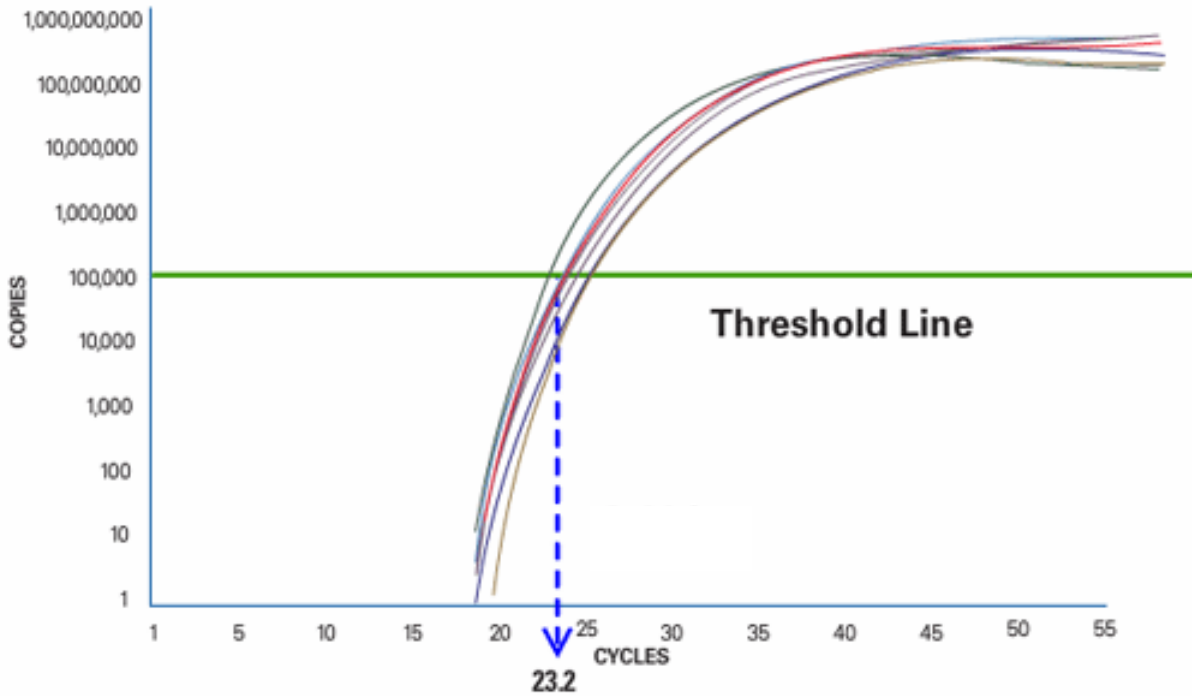
Teknikk for å lese av resultat (usynlig DNA)– Ulike typer PCR:

- **Konvensjonell PCR (Blokkcycler)**
  - Merker PCR-fragmentene etter avsluttet PCR-reaksjon (fluorescence etc)
  - Visualiserer på gel
  - Uspesifikk farging



- **Real time PCR (sanntids PCR)**
  - Merker PCR-produktet underveis i PCR-reaksjonen (fluorescence etc)
  - Visualiserer PCR-kurven på PC-skjermen u/ kjøring
  - DNA-spesifikk / uspesifikk innfarging





C_t / C_p /C_q Value

Primerdesign:

- Viktigste jobben!!!
- Vanskeligste jobben!!!
- En god PCR = avhengig av gode primere

- Prøve primere som tidligere er publisert
  - Heldig/ikke heldig....

- Designe egne primere – Bruke nettbaserte ”primer-design” programvare
  - Heldig/ikke heldig

- Kjøpe ferdigproduserte primere /kit

- I stor grad avhengig av DNA-sekvensen som skal amplifiseres!
  - GC-rike områder vanskelige
  - Størrelsesrestriksjoner på PCR-produkt/setespesifikk mutasjon

- Stor grad avhengig av empiri – prøve og feile....
- Prøve flere sett med primere
- Optimalisere – visualisere uten sekvensspesifikk deteksjon (SYBRGreen)
- Avhengig av type gen – noen gener enklere enn andre
- Må sikre et rent PCR-produkt
  - NB! Diagnostikk

Valg av metode:

Utifra tilgjengelig labutstyr:

PCR-plattform:

- Real time
- Konvensjonell Blokkcykler
- Sekvensator (minisekvensering, pyrosekvensering)
- Kapillærelektroforese

Type deteksjon:

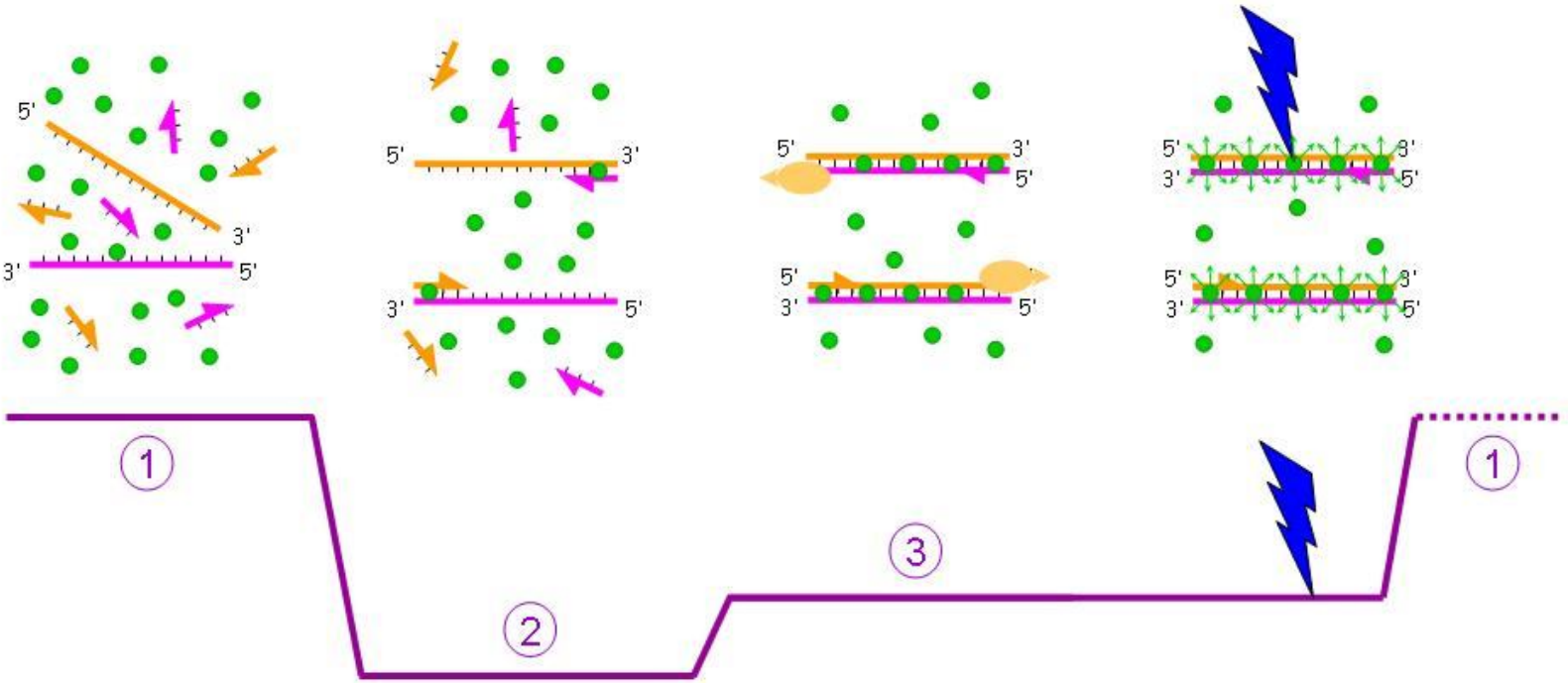
- Fluorescence:
  - Uspesifikk merking (SYBR-Green)
  - Sekvensspesifikke Prober (Taq-Man, FRET mm)
- Restriksjonsenzymer - Fragmentanalyse

Deteksjon av PCR-produkter:

- **Uspesifikk innmerking:**
  - Alt DNA tilstede merkes
  - ds/ss DNA, minor/major groove
- **Spesifikk innmerking:**
  - Bruk av sekvensspesifikke prober
  - Likner en primer som er påhektet en fargekomponent

Uspesifikk merking av PCR-produktet (real time)

Eks. SYBRGreen

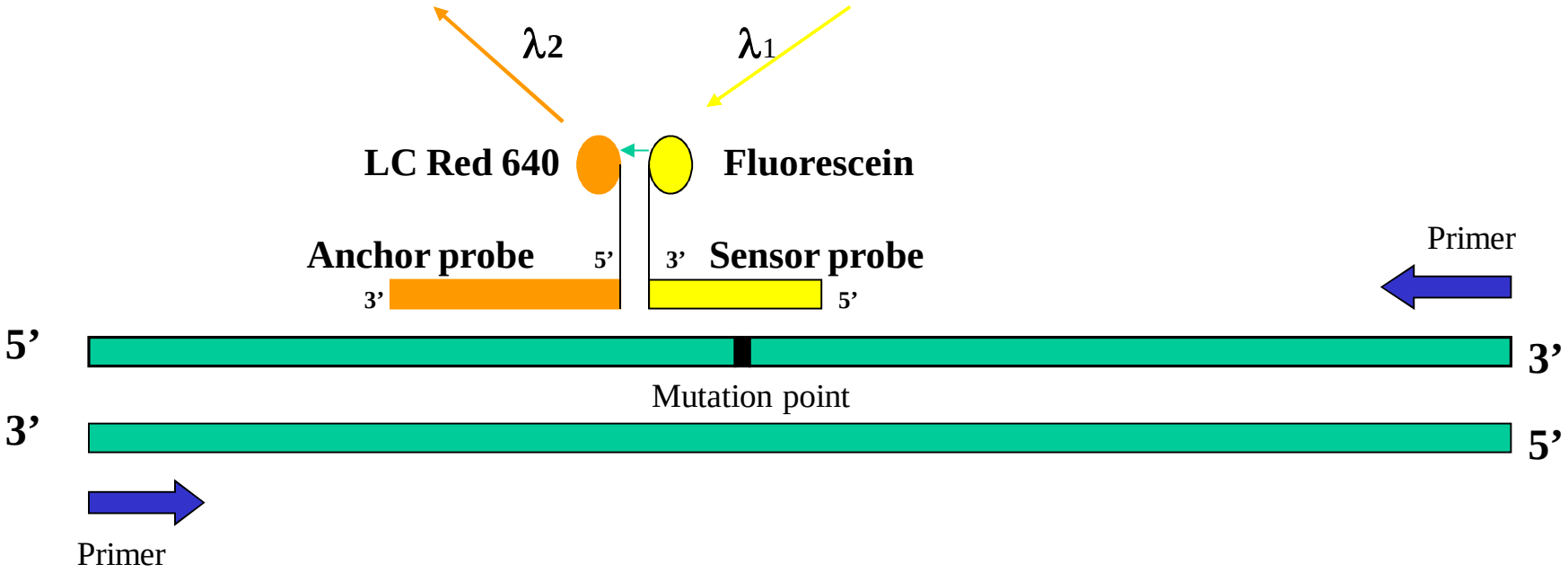


PCR-produktet kan også merkes uspesifikt etter avsluttet PCR – feks ved at agarosegelen farges

Valg av prober:

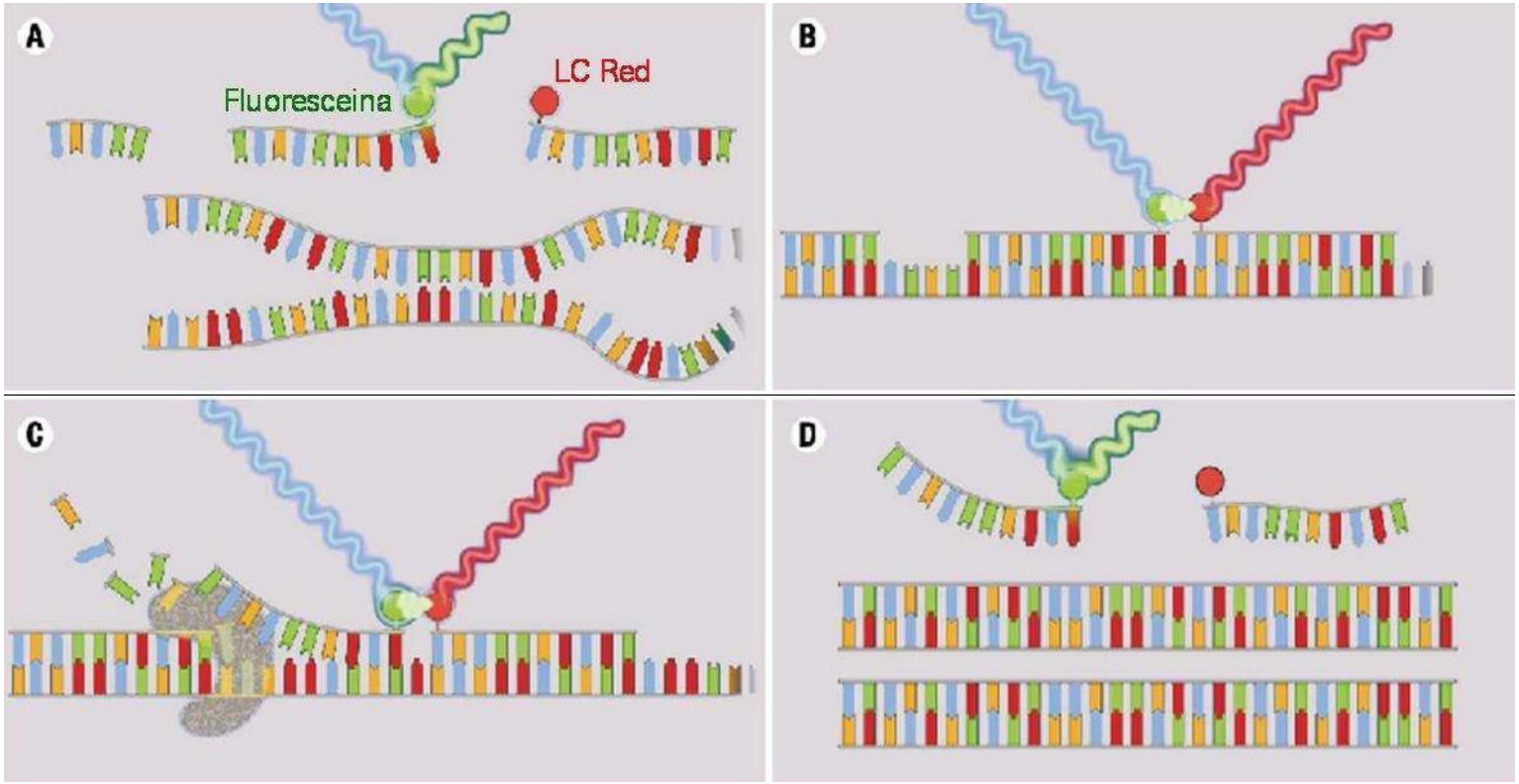
- Mange type prober, ulike prinsipper
- Fordel med prober: Økt spesifisitet inn i systemet!
- Avhengig av hvilken type metode som velges
  - Multiplex: Flere probesett med ulike fluoroforer
- Begrensning: Mutasjonen ligger der den ligger….
- Må ofte optimalisere på nytt etter probetilsetning pga bruk av andre reaksjonsmixer

THE FRET PRINCIPLE

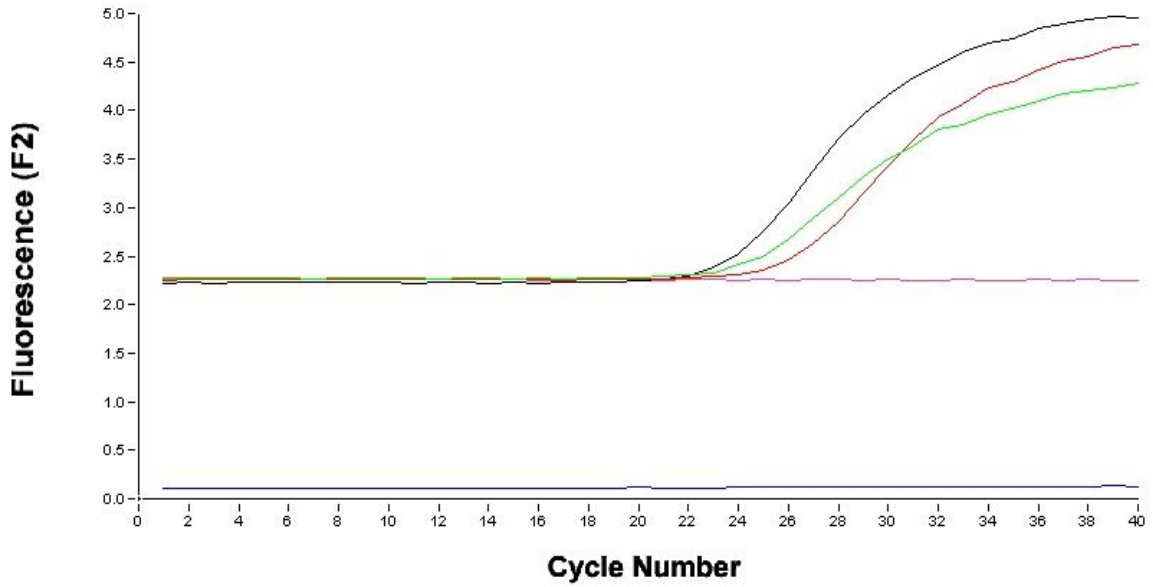


Bruk av FRET-prober

(Fluorescence Resonance Electron Transfer):



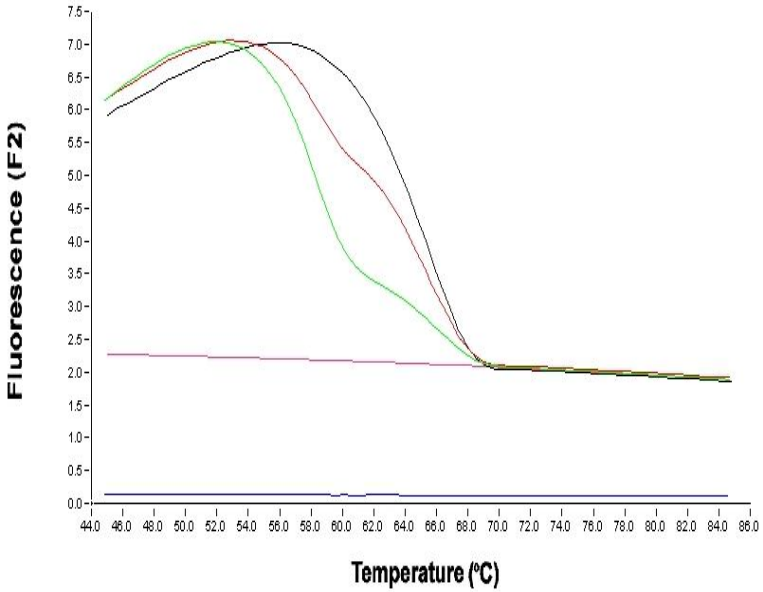
Real-time PCR amplifisering (LightCycler):



Start cursor: 0:10:19.8781 Stop cursor: 0:39:38.9917

Color Compensation: Off

Smeltekurver – generering av fall i fluoresence:



FRET probene - 100 % homologi

- **Villtype sekvens**
- **Mutant sekvens**

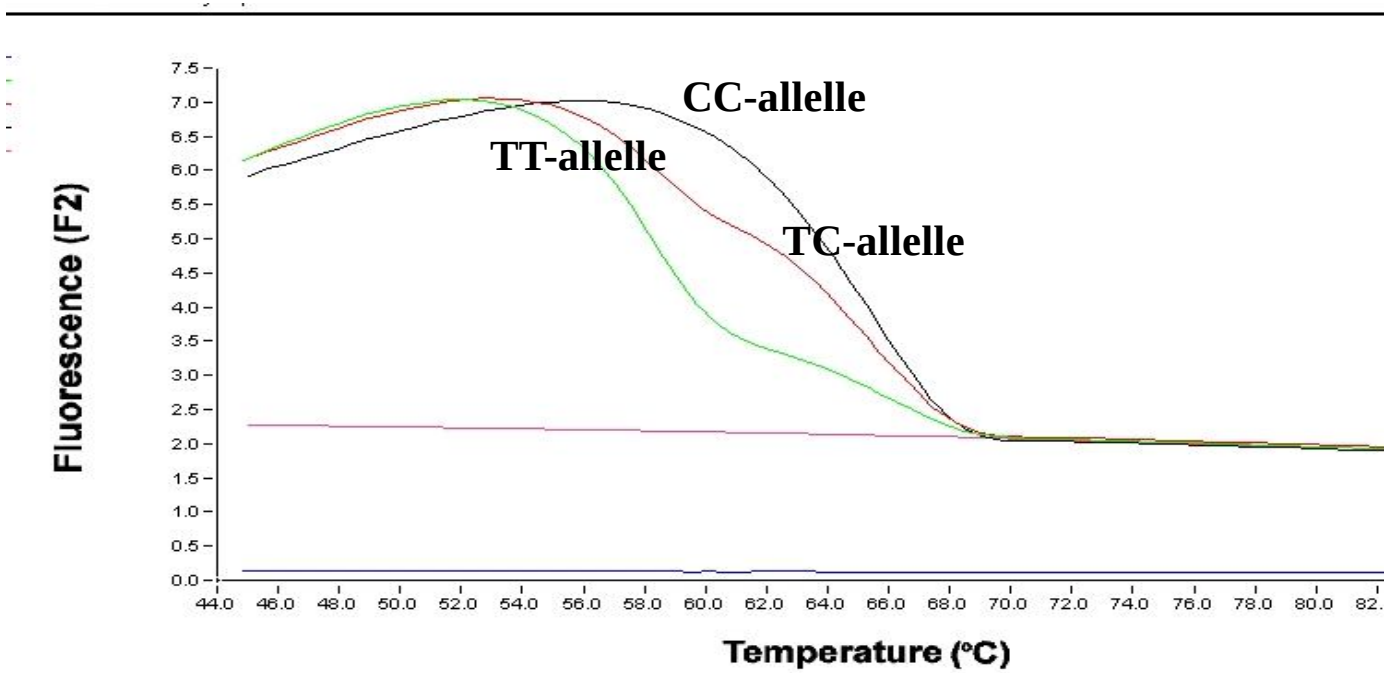
- **100 % Homologi:**

Mest stabile DNA-kompleks – denaturerer ved høy temperatur

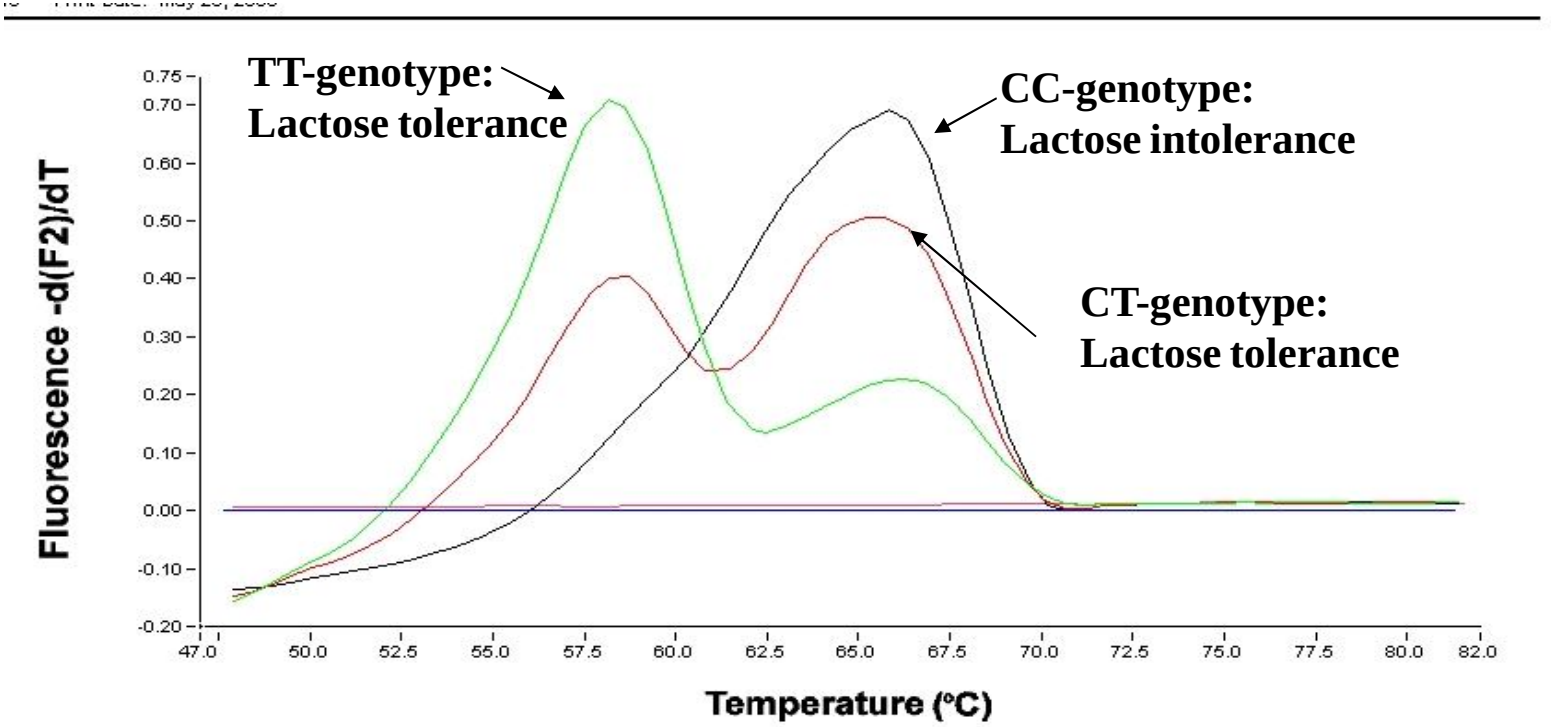
- **Mismatch:**

Minst stabile DNA-kompleks - denaturerer ved lavere temperatur

Smeltekurver:



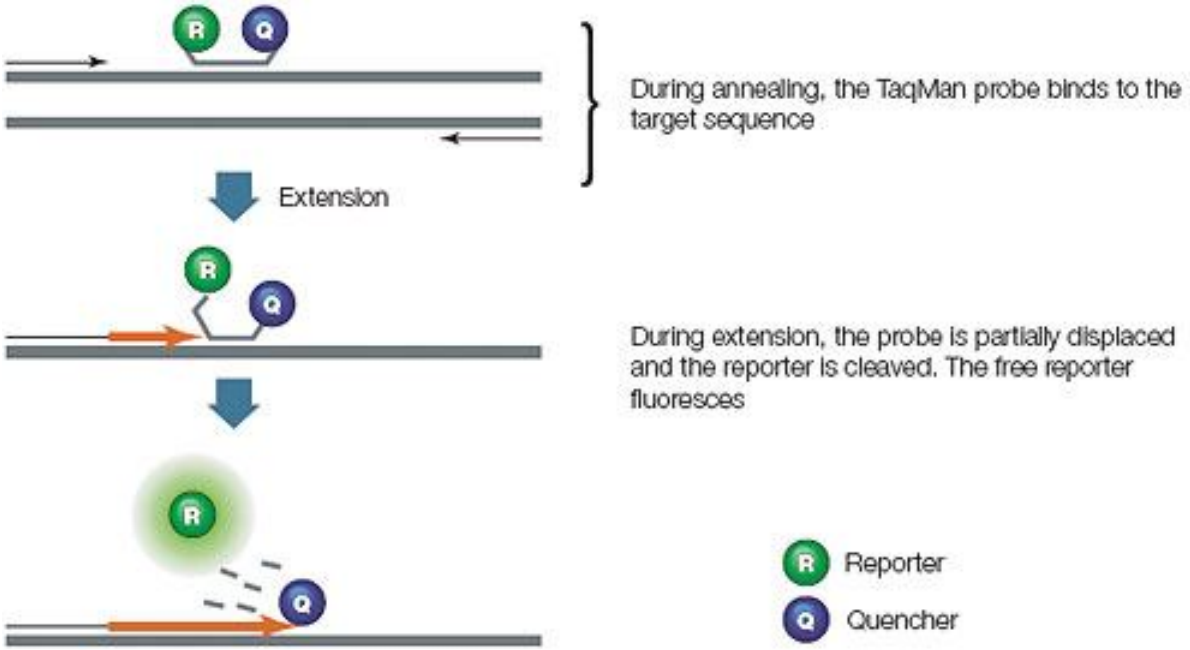
Smeltetopper:



TaqMan-prober:

Benytter 2 ulike DNA-prober til ett gitt locus

1. Mutasjon-probe
2. Villtype-probe



PCR-plattformer:

- **LightCycler**

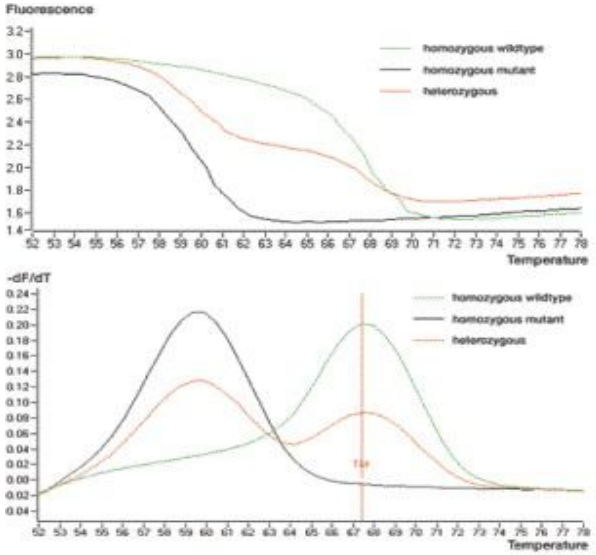
- **Uno Cycler**

- **ABI-Prism**

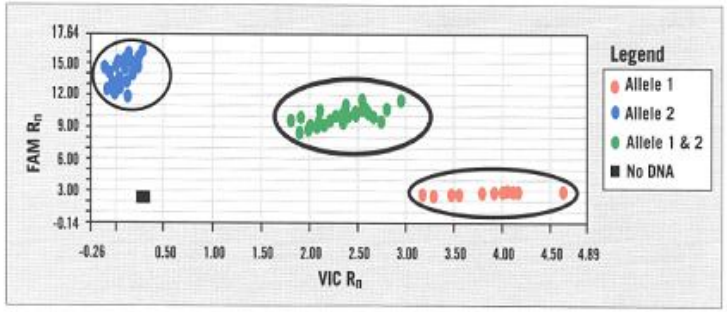
- **Thermo Cycler**

- **Blokkcycler**

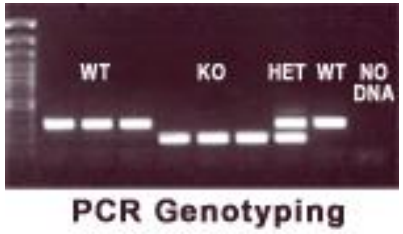
- **Ulike formater** (24, 32, 96, 380, Array)
- **Automasjon**
- **Multiplex**
- **IVD- sporbarhet**



Avlesning av genotype vha Smeltepunktsanalyse.
Baserer seg på at mutant og villtype-sekvens har ulikt smeltepunkt (Tm) når de er bundet til en sekvensspesifikk probe.



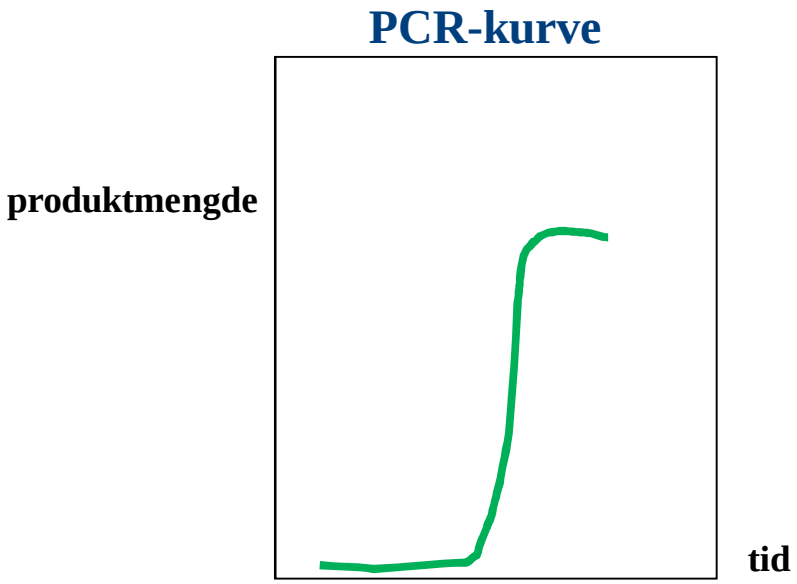
Avlesning av genotype i et koordinatsystem vha generert fluorescence.
Baserer seg på sekvenshomologi mellom to allel-spesifikke prober (mutant/villtype-sekvens).



Avlesning av genotype vha separering på agarose/PAGE gelelektroforese.
Baserer seg på observasjon av fragmentstørrelser m/u restriksjonsenzymer.

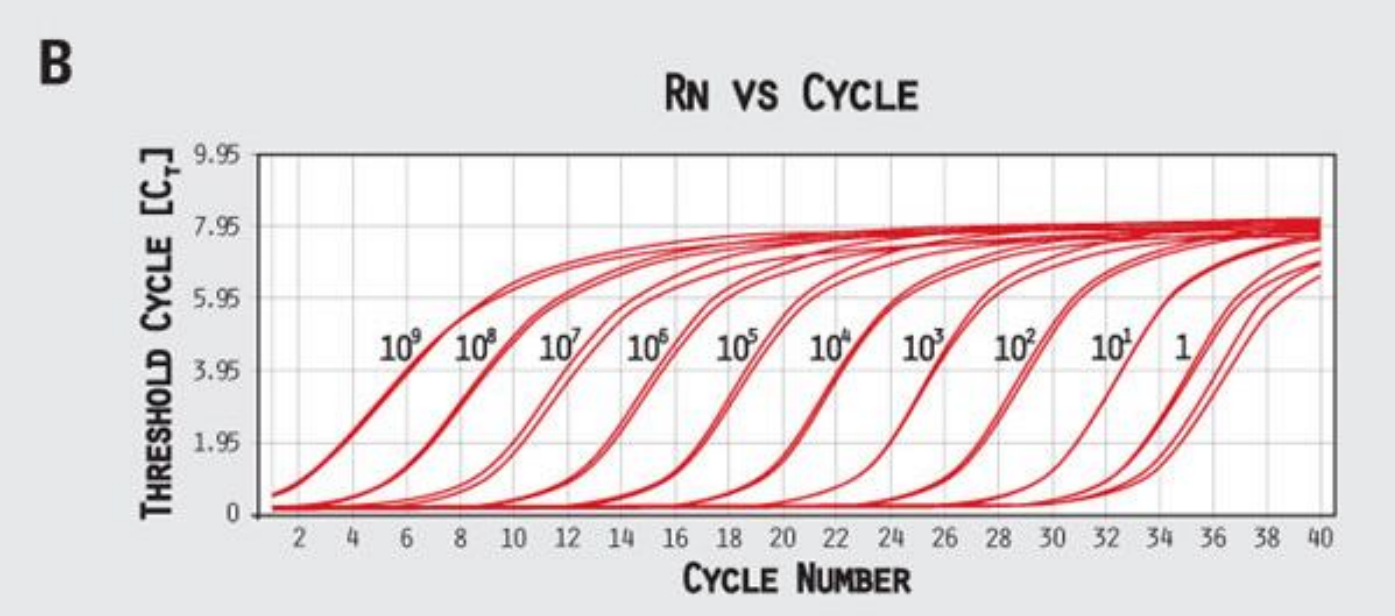
Optimalisere PCR-reaksjonen:

- PCR-effektivitet
- Temperatur
- Mg^{+}-konsentrasjon
- Primere (design/kons.)
- Prober (design/kons.)
- Andre tilsetningsstoffer for å bedre betingelser



PCR-effektivitet:

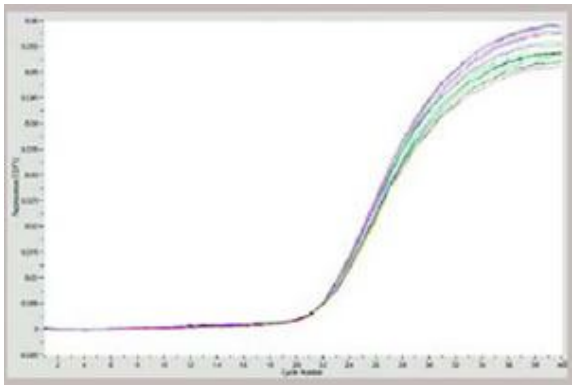
Fortynningsrekker



Validere:

- Alternativ analyseplattform
- Sekvensere PCR-produkt =Gullestandard
- Kontroller
  - Positive / negative mutasjonskontroller
  - Vannkontroll (No template)

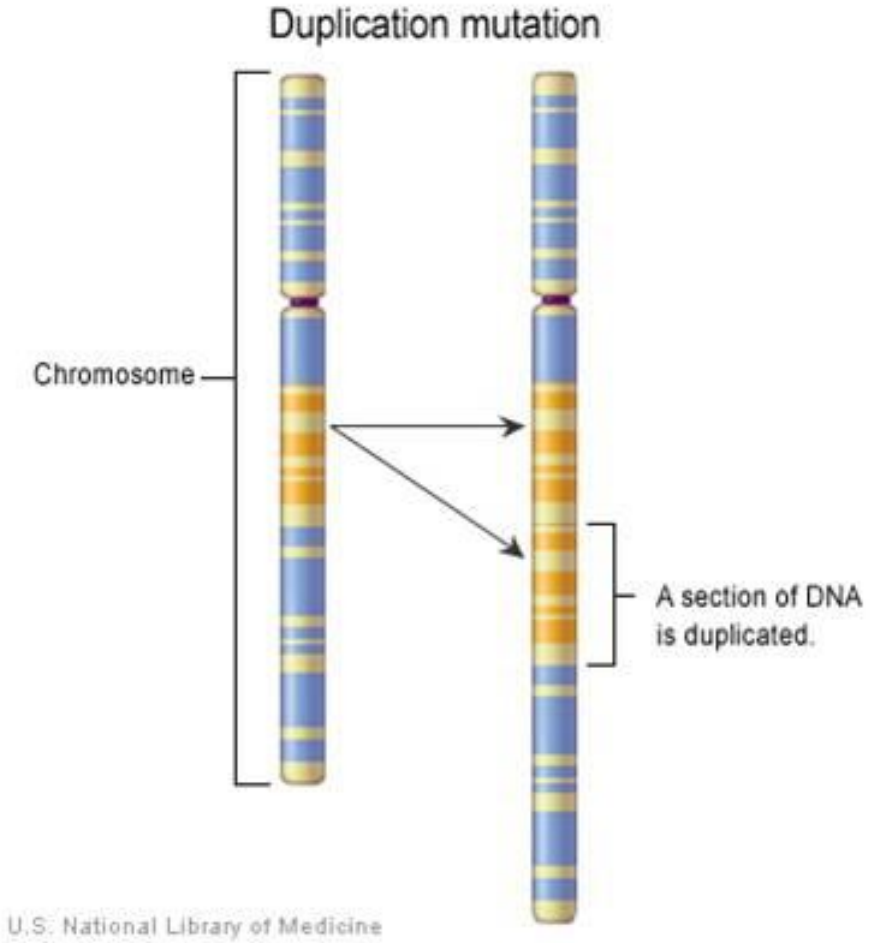
- Teste: DNA-pool



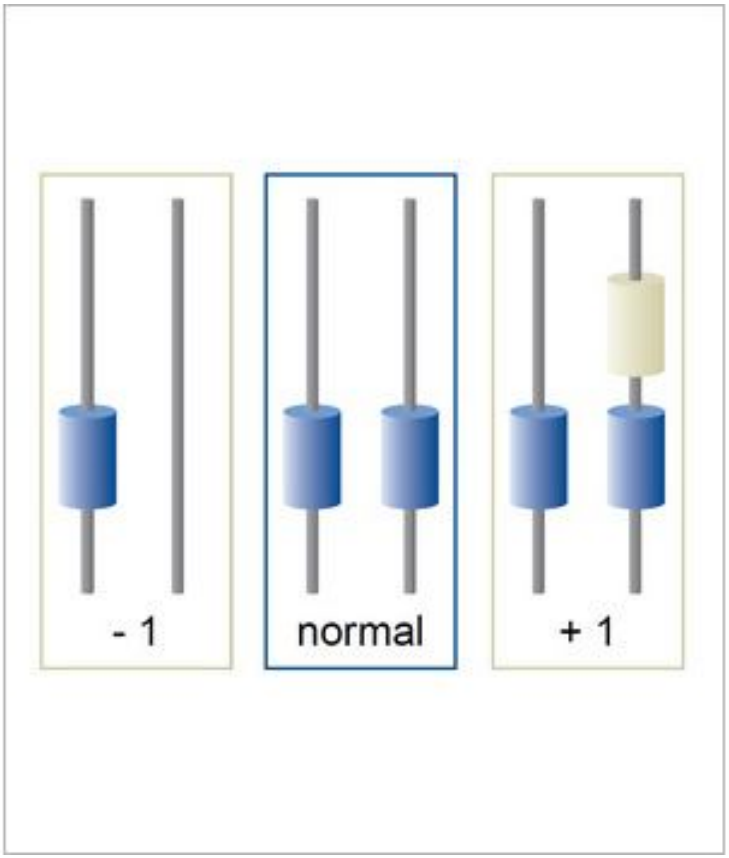
Generelt Gentester:

- **NB! FORURENSING!!!**

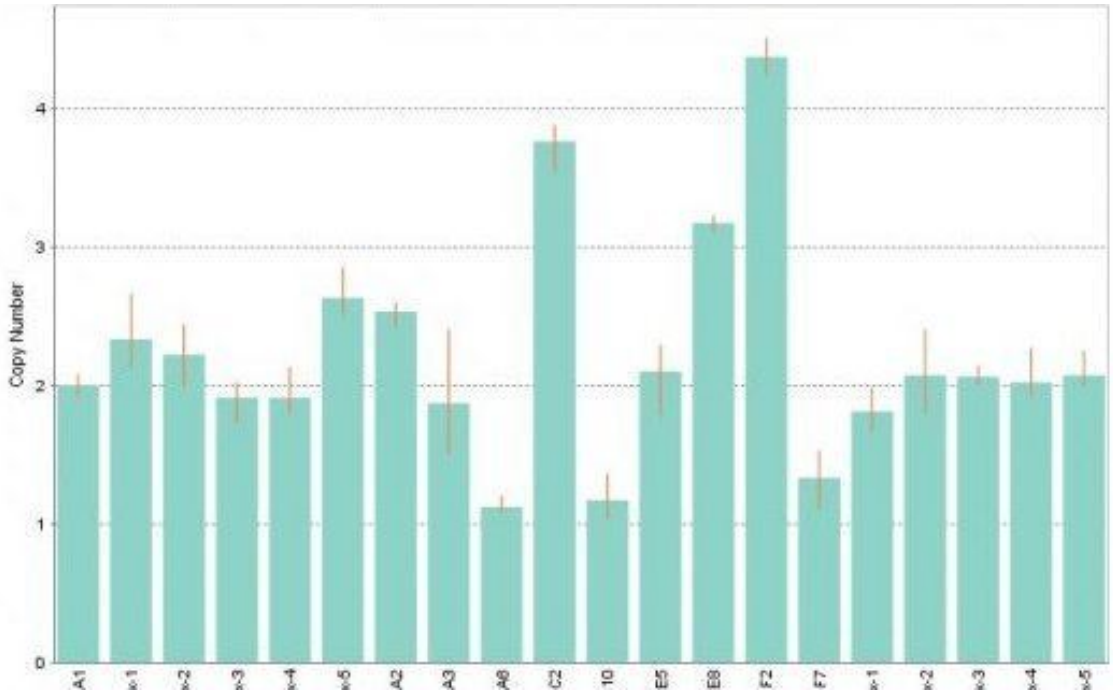
Kopitallvariasjon (Copy number variation=CNV)



CNV:



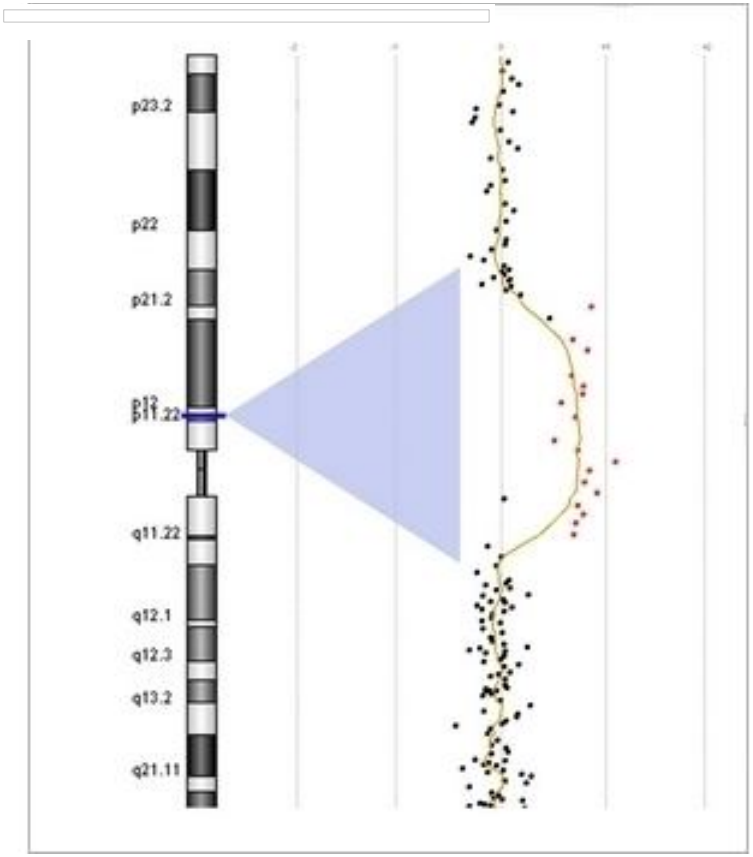
Kopitallanalyse:



Kvantitativ PCR: Mange muligheter
Hvis mistanke om større Delesjoner/Insersjoner

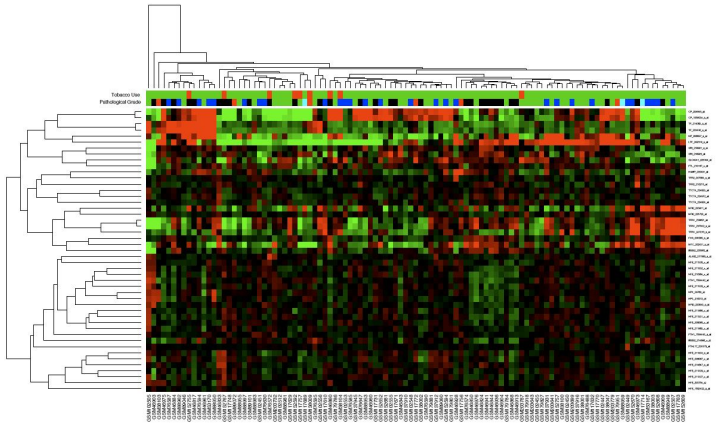
Viktig med primerdesign
Få et visst overblikk over stort DNA-område
Ofte etterfulgt av sekvensering

CGH (comparative genomic hybridization)



”High resolution whole-genome screening data for genomic aberrations”

Microarray /chips:



- Analyser DNA / RNA
- Global DNA-variasjon
- Global RNA-ekspresjon
- Alternativ splicing , microRNA m.m.
- Sykdoms-"pattern" (DNA-variasjon / RNA-ekspresjon

- Foreløpig mest forskning
- Aktuelt innen medisinsk diagnostikk

DNA-Sekvensering = Gullstandarden

- Sekvensering : Bestemmer basenes interne rekkefølge i DNA
  - Pyrosekvensering
  - Sanger sekvensering
  - Helgenom/Dyp sekvensering
  - ”Next generation sequencing”

Pyrosekvensering:

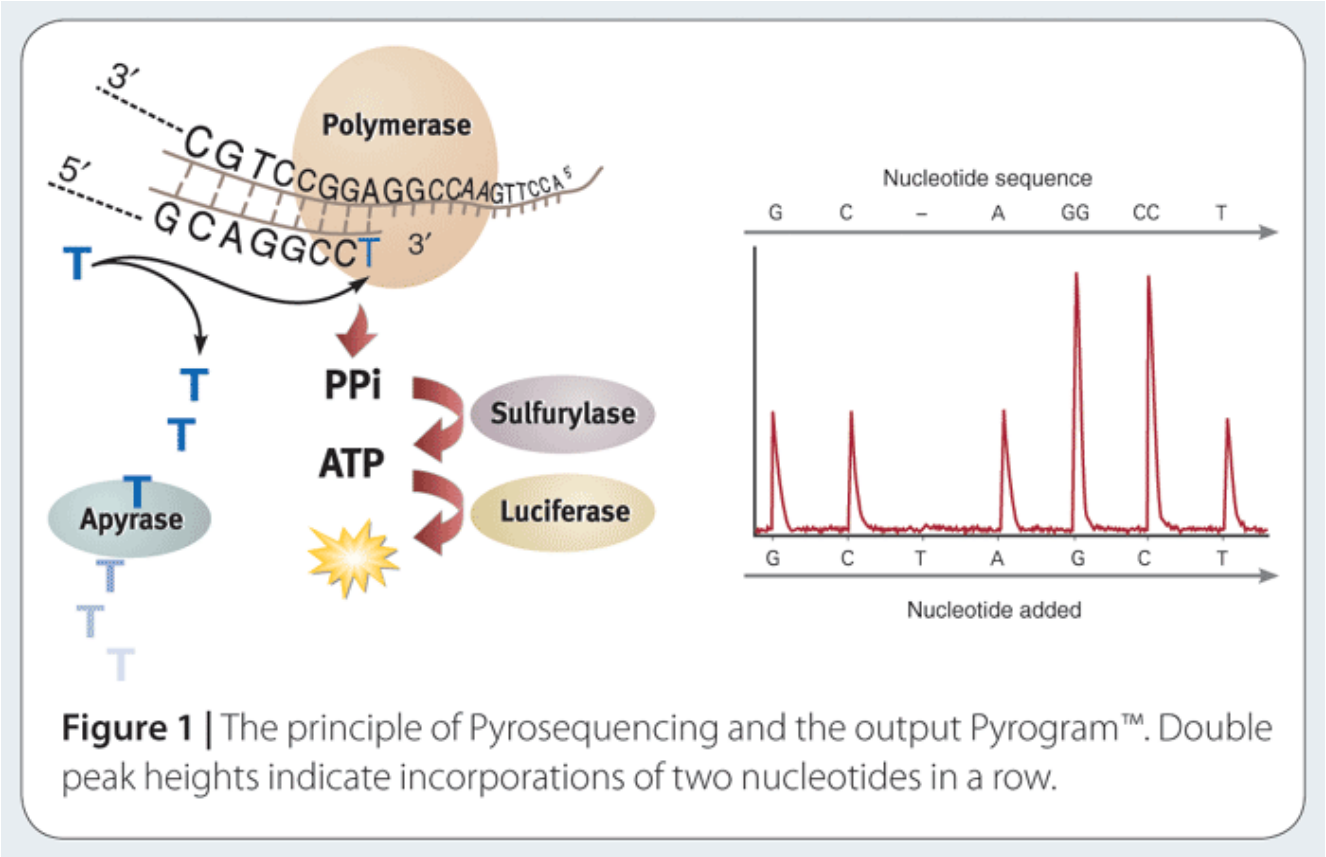
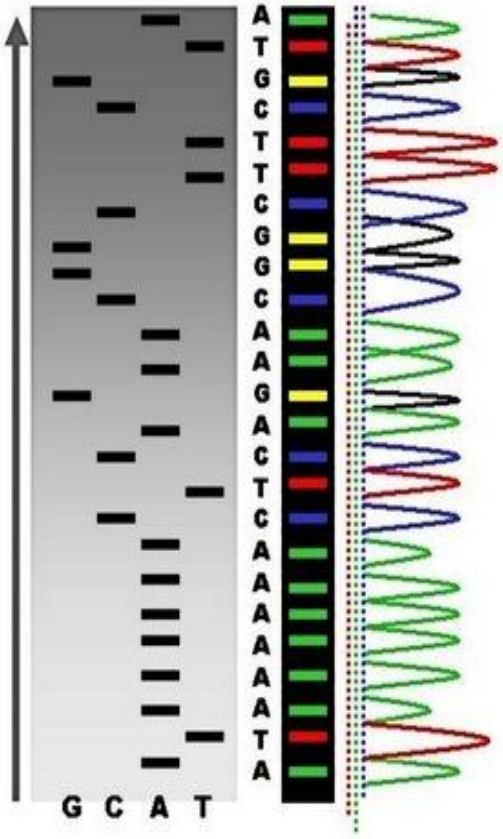
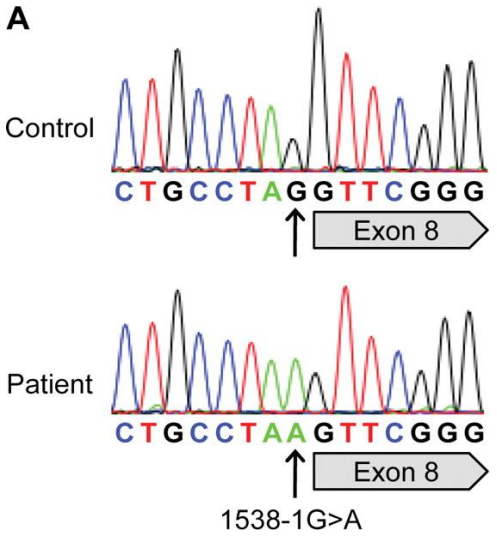


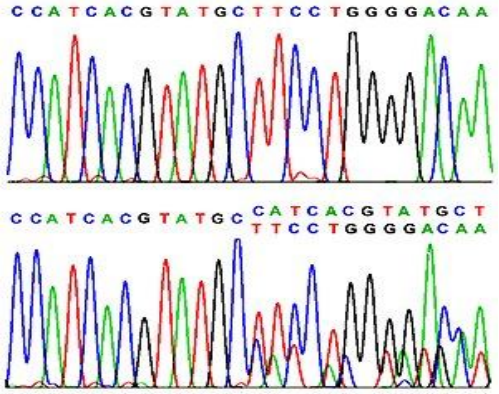
Figure 1 | The principle of Pyrosequencing and the output Pyrogram™. Double peak heights indicate incorporations of two nucleotides in a row.

NB! Sekvenserer kun kortere biter av DNA

Sanger sekvensering (tradisjonell sekvensering):



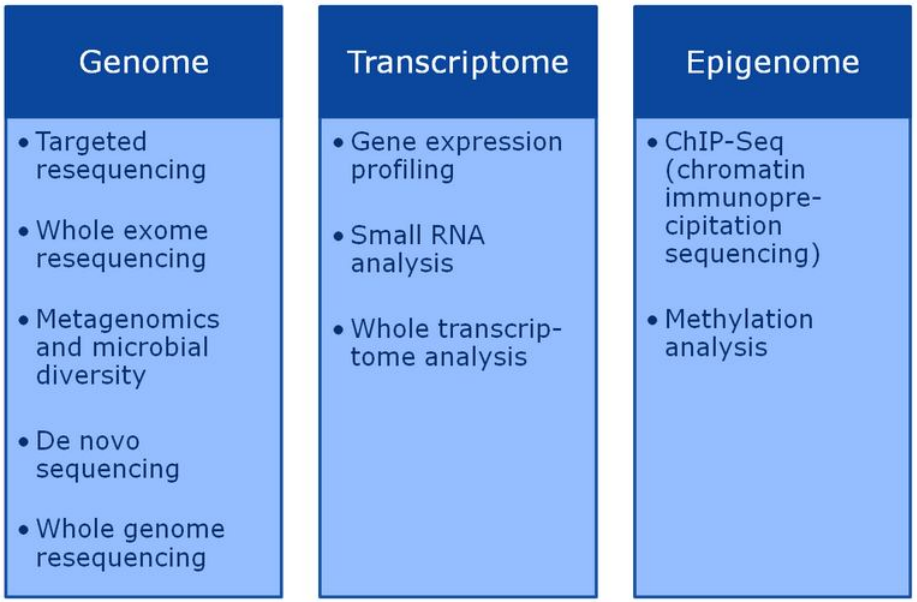




Helgenom / Dyp sekvensering / ”Next generation sequencing”

Nye teknikker:

- **Økt sekvenseringshastighet**
- **Reduserte kostnader**
- **Økt brukspotensiale**
- **Økt innsyn i genmaterialet**
- **MANGE MULIGHETER + MANGE ETISKE BETENKELIGHETER !!!**



Økt sekvenserinshastighet:

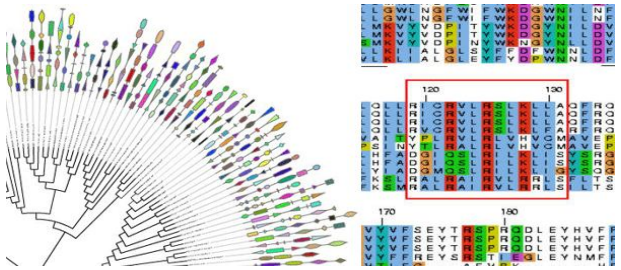
Enorme mengder genetisk informasjon fra kartlegging av en rekke genomer

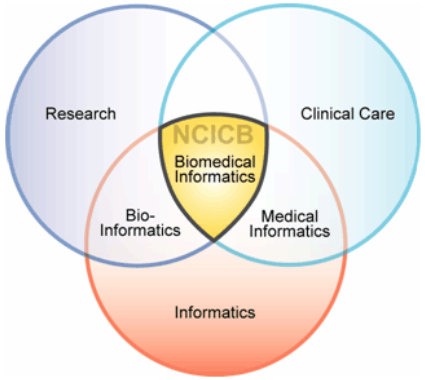
IT: Sentral rolle i genteknologien / molekylærbiologien

- **Bioinformatikk = ny disiplin**
  Samarbeid mellom molekylærbiologer / informatikere

Metoder for:

Systematisering
Tolkning





Eksempel:
Diagnostikk av Thalassemi

- Gruppe av arvelige sykdommer
- ⇓ Hb
- Anemi
- Alvorlighetsgrad/type anemi varierer

- **Thalassemier: Redusert mengde / bortfall av Hb-kjeder som inngår i Hb molekylet**
- **Defekt syntese av normalt Hb protein**
- **Syntese av defekte proteiner**

- **<u>På verdensbasis:</u> Defekt Hb syntese en av de mest vanlige medfødte sykdommene – beregnet at det finnes ca. 150 mill. genbærere**

- **”Thalassemi beltet”: Fra Middelhavslandene over Midt-Østen, via India og Syd-Øst Asia til Indonesia + Afrika**

- **Lav forekomst blant Nord-Europeiske populasjoner**

- **I Norge: Økende forekomst som følge av økt innvandring**

  Blant innvandrere opp mot 3 – 4%

Hemoglobin varianter:

Finnes 4 globinkjeder: α, β, δ, γ

Hemoglobin: **Bygget opp av 4 globinkjeder**
2 par kjeder (2x2)
Aldersbundet Hb mønster.

2 α + 2 β	**HbA1, dominerer**	**Normalt**
2 α + 2 δ	**HbA2, små mengder**	
2 α + 2 γ	**HbF, føtalt**	

4 β	**HbH**	**Patologisk**
4 γ	**Hb Barth**	

Diagnostikk basert på:

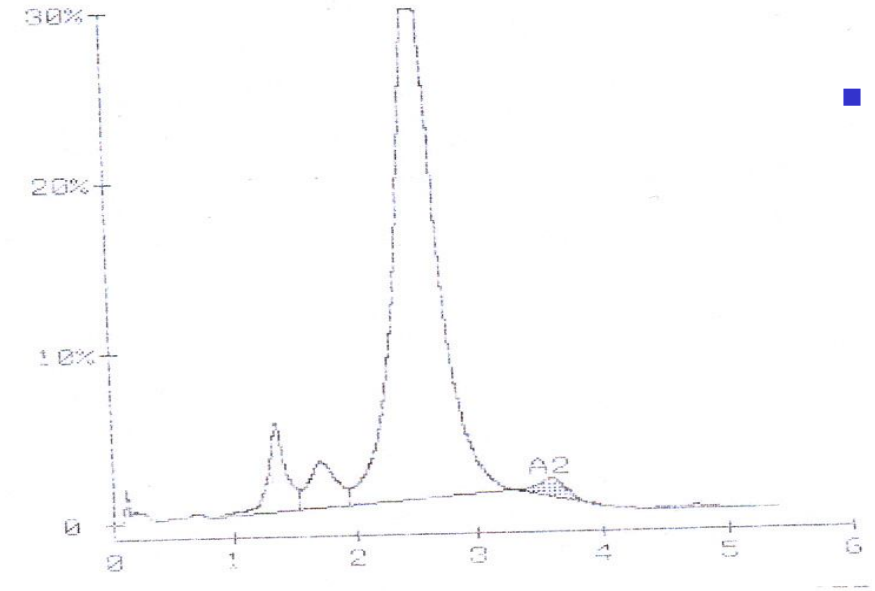
- Anemi som ikke responderer på jerntilskudd
- MCV < 75
- MCH, Ret-Hb
- Hb undersøkelse
- Utredet mhp jernmangel
- Geografisk opprinnelse
- Familiebakgrunn

Laboratoriediagnostikk

- Hematologiske undersøkelser
- Hb HPLC / (elektroforese)
- Molekylærbiologiske metoder

Hemoglobin subtyping (HPLC):

ANALYTE ID	%	TIME	AREA
>2	6.2	1.34	103917
>3	5.0	1.72	84322
40	86.4	2.53	1445984
42	2.3	3.59	34068
	TOTAL AREA		1668291
F	0.0%	A2	2.3%

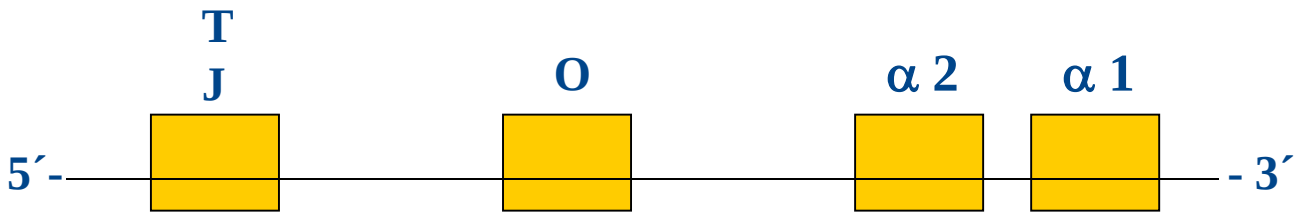


- **Normalt Hemoglobinmønster**

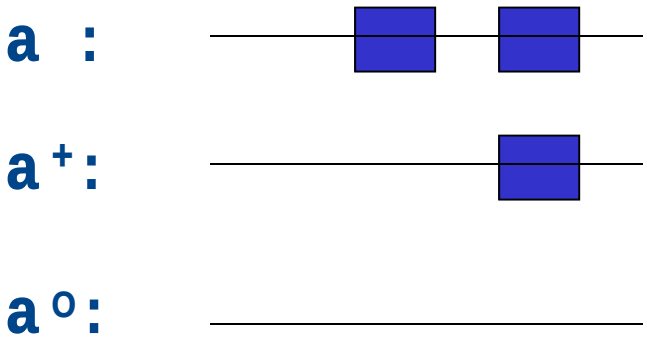
Genetikk – α-thalassemi

- **α-globin protein produksjon:**
- **Styres av 4 α-gener**
- **Kromosom 16**
- **2 gener på hvert kromosom (α1 og α2)**

α-gen kompleks:



- **Vesentlig delesjoner årsak til α-globin svikt (finnes også en del mutasjoner)**
- **Single/Dobbel delesjon av α-gener (på hvert kromosom)**
- **<u>Allelvarianter:</u> cis/trans**



- **Tap av 1 α-kjede (aa^+):** Silent carrier
  - Normalt ingen kliniske symptomer

- **Tap av 2 α-kjeder (a^+a^+):** Mild a-thalassemi
  - Mild anemi, mikrocytose, lav Hb

- **Tap av 3 α-kjeder (a^+a^o):** HbH
  - Alvorlig anemi, transfusjonsavhengig

- **Tap av alle 4 α-kjeder (a^oa^o):** Hydrops fetalis
  - Uforenlig med liv (bare g-kjeder, Hb Barts)

→ **Sykdomsmanifestasjon korrelerer med antall defekte gener!**

α-thalassemi - delesjons varianter

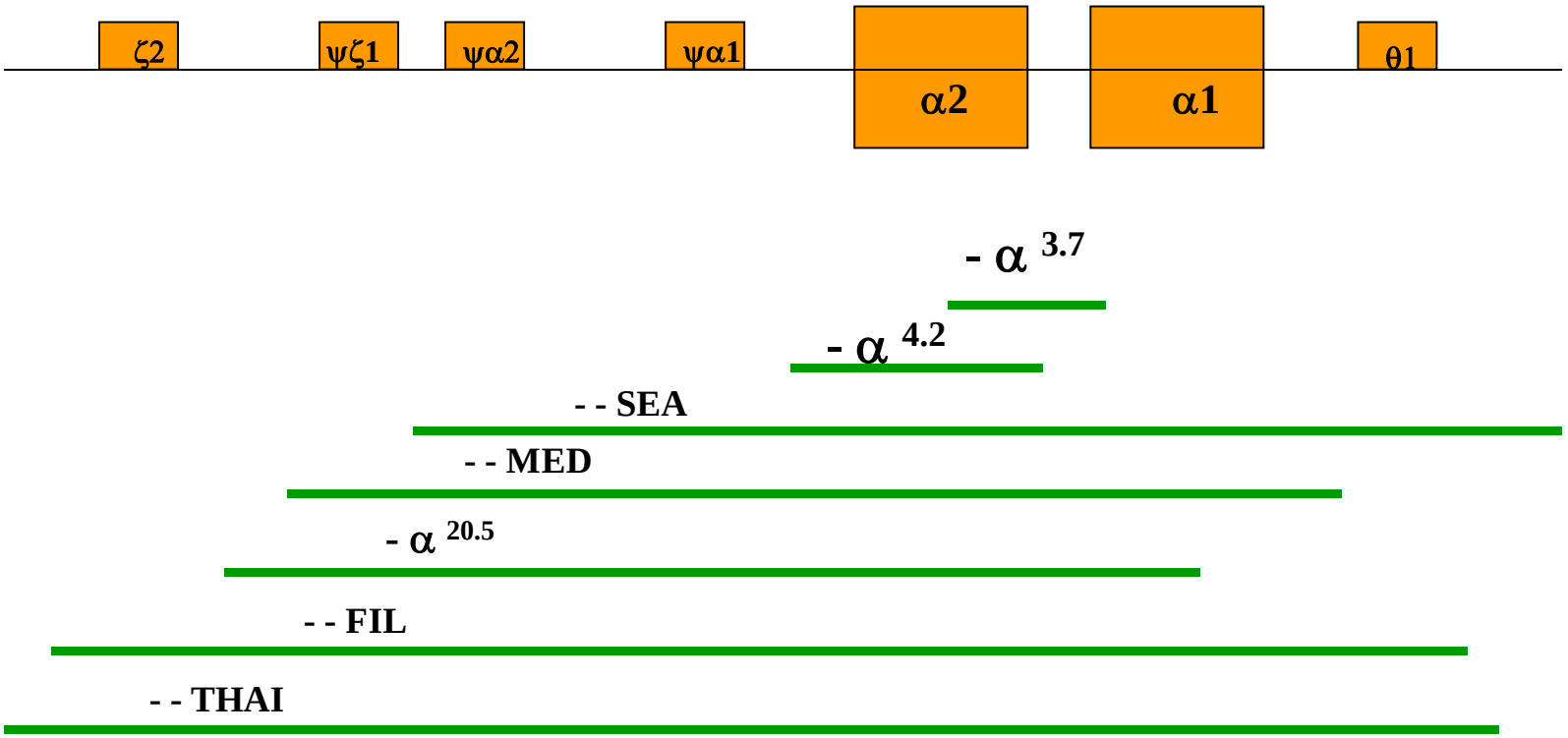
- **Delesjoner hvor ett α-globin gen er slukket:**

  - $-\alpha^{4.2}$: Fjerner α_2 genet

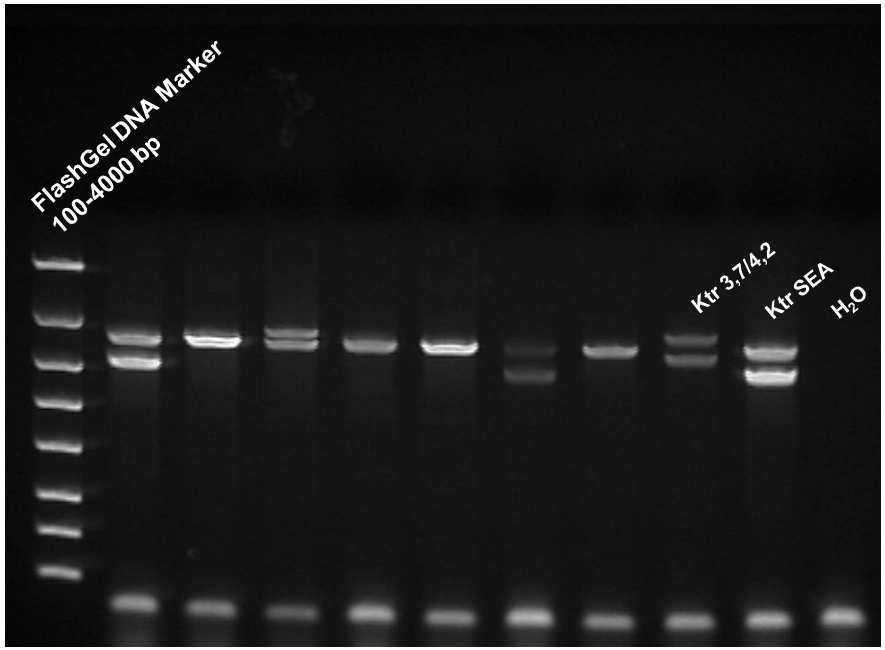
  - $-\alpha^{3.7}$: Fjerner deler av α_1 og α_2 genet – danner fusjonsgen

- **Delesjoner hvor begge α-globin gener er slukket:**
  - $-\alpha^{20,5}$, $-\alpha^{FIL}$, $-\alpha^{SEA}$, $-\alpha^{THAI}$, $-\alpha^{MED}$

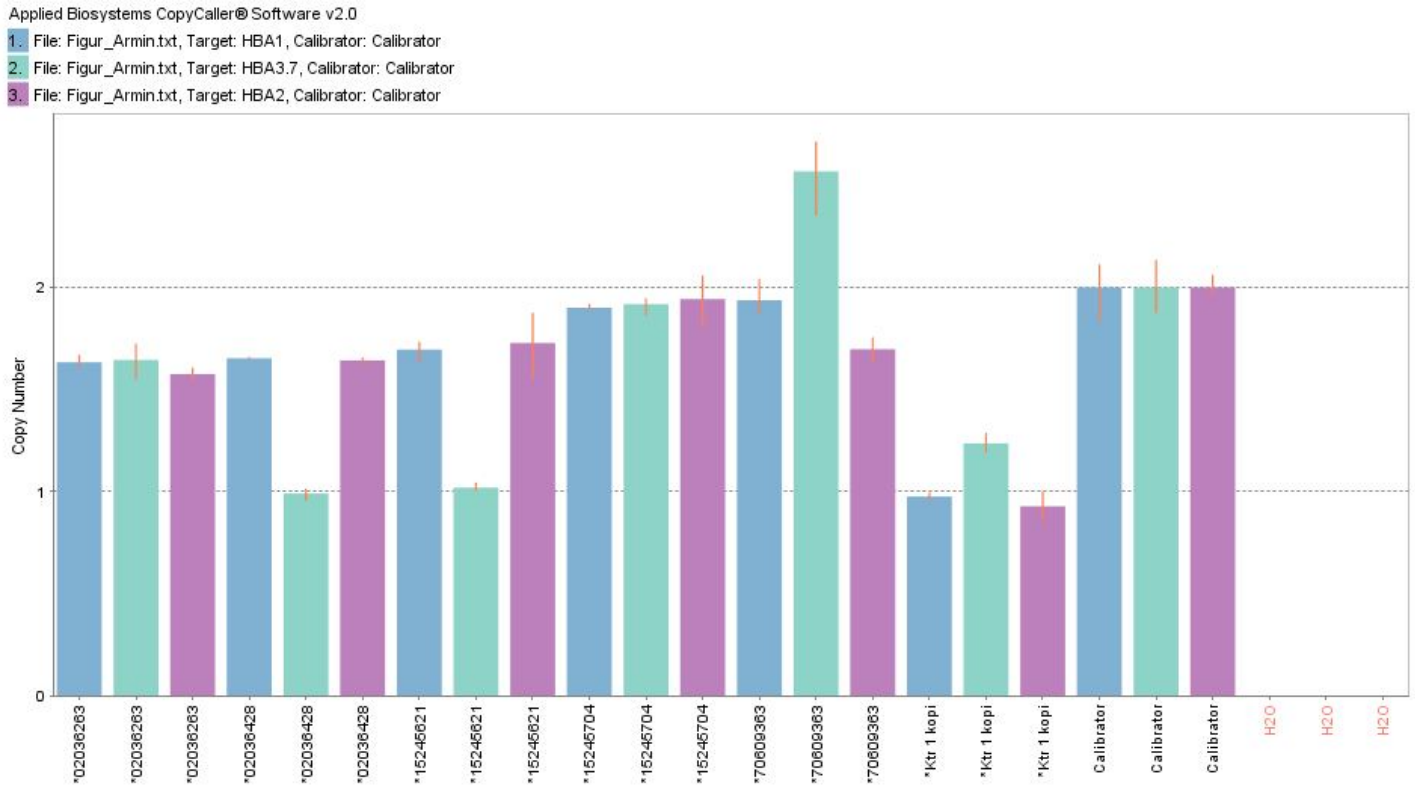
De mest vanlige α-thalassemi delesjonene:



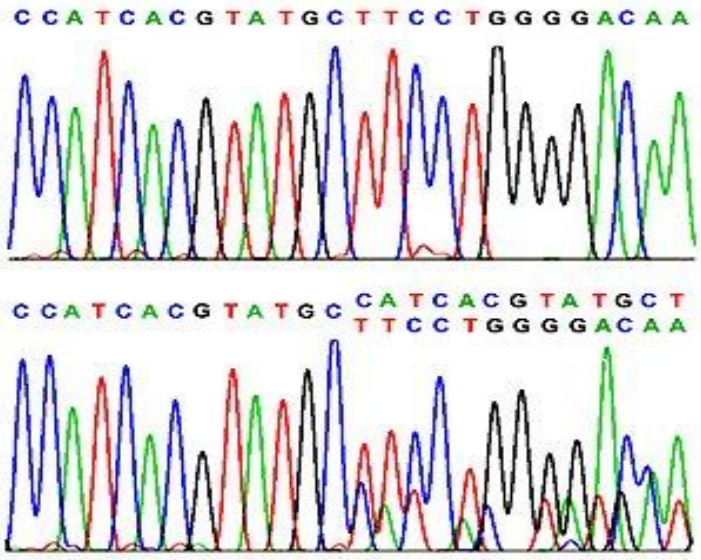
Agarose gelelektoforese α-thalassemi Multiplex-PCR (gap-PCR) :



Kopitallvariasjon Hb α-genene:



Sekvensering av α-genene



Oppsummering α-thalassemi:

- Ved mistanke om a-thalassemi:
  - Anemi – utelukke jernmangel
  - MCV, MCH, Ret-Hb
  - Etnisk bakgrunn
- Hemoglobin HPLC
- Molekylærbiologisk utredning
  - Multiplex PCR med primere for de vanligste α-globin delesjonene
  - Kopitallvariasjon
  - Sekvensering av α-globin genene

Take home message:

Molekylærbiologi / Genteknologi:

- **Spennende ”fagfelt”**
- **Passer godt i et medisinsk biokjemisk fagfelt**
- **Mange muligheter i medisinsk diagnostikk**
- **Krever spesialkompetanse for å jobbe med**
- **Ikke lenger ”spesialanalyser”**
- **VERKTØY til å komme til bunns i medisinsk diagnostikk**