Roche

Diagnostics

Real-time PCR, a personal perspective

Russ Higuchi

qPCR 2005 Symposium

Technische Universität München

September 5, 2005

Part I

Me, DNA and real-time PCR – a brief history

Part II – Two things about real-time PCR

- Why microarray hybridization cannot be as sensitive or have the dynamic range of PCR
- Why thermocycling is an advantage over isothermal amplification

Application of real-time PCR to quantitative genetics

Proc. Natl. Acad. Sci. USA
Vol. 73, No. 9, pp. 3146–3150, September 1976
Biochemistry

A general method for cloning eukaryotic structural gene sequences

(complementary DNA/messenger RNA sequencing/gene isolation/terminal transferase)

RUSS HIGUCHI*, GARY V. PADDOCK*, RANDOLPH WALL†, AND WINSTON SALSER*

* Department of Biology and Molecular Biology Institute, and † Department of Microbiology and Immunology, School of Medicine, and Molecular Biology Institute, University of California, Los Angeles, Calif. 90024

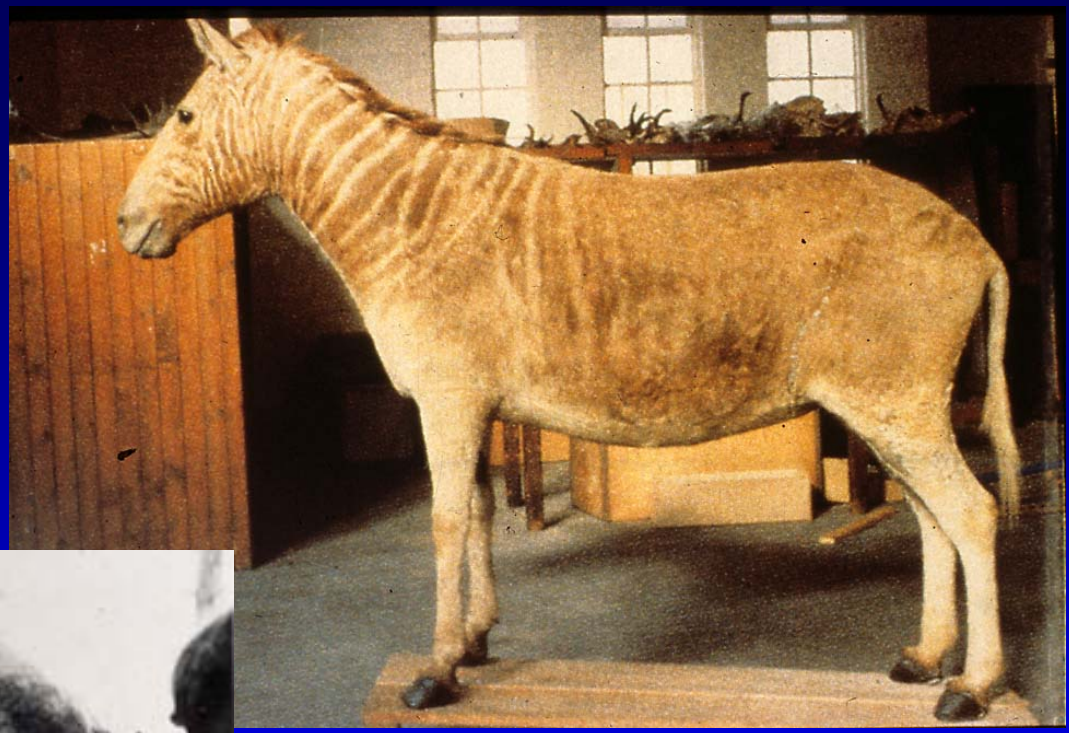
Communicated by Paul D. Boyer, July 7, 1976

ABSTRACT Complementary DNA, transcribed *in vitro* from purified rabbit globin messenger RNA and made double-stranded, has been inserted into *Escherichia coli* plasmids pSC101 and pMB9 by the poly(dT)/poly(dA) "tailing" and annealing technique. *E. coli* transformants given by this DNA preparation have been shown to contain globin sequences by the hybridization of globin RNA to DNA from clones grown and lysed *in situ* on nitrocellulose filters. An estimate of the amount of inserted globin sequences has been provided by fingerprint analysis of globin mRNA sequences hybridized to the purified plasmid chimeras. Inserted sequences so far subjected to detailed analysis have been ascribed to the rabbit beta globin chain. The susceptibility of inserted beta globin sequences to the restriction endonuclease *Eco*RI confirms the existence of a site already found through previous nucleotide sequence analysis.

Graduate school, UCLA
Molecular Biology Institute
1974 -1980

Advisor: Winston Salser

Ancient DNA?



110 year-old quagga skin



20,000 year-old frozen wooly mammoth

post-doctoral fellowships
UC Berkeley
Biochemistry Dept.
1980 – 1986

Advisor: Allan Wilson

	Quagga	Burchell	Mt. Zebra	Horse	Cow
Quagga	–	2*	12	14	41
Burchell		–	10	12	39
Mt. Zebra			–	16	39
Horse				–	38
Cow					–

*possible post-mortem change

Matrix of nucleotide differences

Reprinted from Nature, Vol. 312, No. 5991, pp. 282-284, 15 November 1984
© *Macmillan Journals Ltd., 1984*

DNA sequences from the quagga, an extinct member of the horse family

Russell Higuchi*, Barbara Bowman*, Mary Freiberger*, Oliver A. Ryder† & Allan C. Wilson*

* Department of Biochemistry, University of California, Berkeley, California 94720, USA
† Research Department, San Diego Zoo, San Diego, California 92103, USA

To determine whether DNA survives and can be recovered from the remains of extinct creatures, we have examined dried muscle from a museum specimen of the quagga, a zebra-like species (*Equus quagga*) that became extinct in 1883 (ref. 1). We report that DNA was extracted from this tissue in amounts approaching 1% of that expected from fresh muscle, and that the DNA was of relatively low molecular weight. Among the many clones obtained from the quagga DNA, two containing pieces of mitochondrial DNA (mtDNA) were sequenced. These sequences, comprising 229 nucleotide pairs, differ by 12 base substitutions from the corresponding sequences of mtDNA from a mountain zebra, an extant member of the genus *Equus*. The number, nature and locations of the substitutions imply that there has been little or no postmortem modification of the quagga DNA sequences, and that the two species had a common ancestor 3–4 Myr ago, consistent with fossil evidence concerning the age of the genus *Equus*[2].

Mad scientists are cloning dinosaurs as weapons of the future

EGGHEAD SCIENTISTS are secretly cloning dinosaurs — and terrified humans may soon be fleeing for their lives from gigantic monsters belonging to the prehistoric past.

Realizing that a nuclear war could end the world, both U.S. and Soviet scientists plan to conquer the world with invasions of destructive monsters.

This is the incredible analysis of alarmed scientific and religious leaders, as well as environmentalists.

Monsters would be last line of defense in nuclear war

By LEWIS CLIFTON

According to a respected source close to the Washington scientific establishment, genetic material from long-dead dinosaurs found deep-frozen in Siberia, China and the Antarctic is already being reproduced in clandestine laboratories.

Word of the dangerous experiments on dinosaurs being conducted in super secrecy began to leak after disclosure that scientists in California have reproduced gene fragments from an extinct relative of the horse and zebra.

Said the Washington insider: "These experiments — on the Berkeley campus of the University of California — are, of course, perfectly legitimate, and I can understand the excitement over the possibilities of someday reproducing a long-extinct species.

"But it doesn't take much imagination to conceive of the horror and potential destruction to man and environment if this kind of knowledge is used to reproduce gigantic, flesh-eating, reptilian monsters of 20 to 40 tons.

"And we know that a certain radical clique of scientists is doing exactly that in secret laboratories in the USSR, the American Southwest — and possibly in China and Japan."

In Russia, Dr. Sverbighooze Yasmilov of the University of Irkutz has already implanted Indian elephants with frozen mammoth egg cells found in Siberia, and produced eight living mammoths covered in yellow-brown hair.

Nature

The outrageous tampering with nature has also drawn dire warnings from the Reverend James Turner, a Biblical scholar, former church pastor and current traveling evangelist throughout the Southwest.

Turner first heard more than 10 years ago of the wacko experiments on gene fragments extracted from recovered dinosaur remains.

"We are moving here from the province of science to the province of God, and it could mean the ultimate destruction of mankind," he warned.

"These hideous lizards disappeared from this Earth more than 200 million years ago for a good reason, and no matter how far advanced scientifically we may have been allowed to become — it's not our business to try to bring them back."

Environmentalists are also known to be shaken by the prospect of the huge beasts lumbering through our remaining forests, swamplands — and cities.

Said Jack Turner, a trapper and hunting guide in the Pacific Northwest, who is no relation to the evangelist:

"I can't tell you how much truth there might be to this story, but I've heard it. And there's no way I want to share the outdoors with those things. They're unnatural."

The critics of the plan to bring back the dinosaur say the Berkeley experiments provide chilling proof that the time when dinosaurs again walk the Earth may be fast approaching.

The California scientists reported they were working with DNA, the building-block chemical of life, obtained from an animal called a quagga.

The quagga was extermination during the late 19th century by South African farmers who wanted the grazing land for crops.

Amazingly, the flesh was preserved 140 years ago and kept in a museum in West Germany before a tiny portion was sacrificed for the laboratory experiments.



FLESH-EATING BEASTS disappeared from Earth over 200 million years ago.

Report

Significantly, the Berkeley scientists said in a report to the American Society of Biological Chemists on the startling results with the quagga gene fragments:

"The initial success reported today may open the way to recover intact genes from the muscle of frozen mammoths, as well as from bones and teeth of species that died out millions of years ago."

The report was presented by Dr. Russell Higuchi, a leading authority in analysis and laboratory reproduction of DNA.

Advertisement



the hunters ear

Advertisement

Reprinted from Nature, Vol. 332, No. 6164, pp. 543-546, 7 April 1988
© *Macmillan Magazines Ltd., 1988*

DNA typing from single hairs

Russell Higuchi, Cecilia H. von Beroldingen*, George F. Sensabaugh* & Henry A. Erlich

Department of Human Genetics, Cetus Corporation, Emeryville, California 94608, USA
* Forensic Sciences Program, University of California, Berkeley, California 94720, USA

The characterization of genetic variation at the DNA level has generated significant advances in gene and disease mapping[1], and in the forensic identification of individuals[2–6]. The most common method of DNA analysis, that of restriction fragment length polymorphism (RFLP), requires microgram amounts of relatively undegraded DNA for multi-locus typing, and hundreds of nanograms for single-locus comparisons[7]. Such DNA frequently cannot be obtained from forensic samples such as single hairs and blood stains, or from anthropological, genetic or zoological samples collected in the field. To detect polymorphic DNA sequences from single human hairs, we have used the polymerase chain reaction (PCR), in which specific short regions of a gene can be greatly amplified *in vitro*[8–10] from as little as a single molecule of DNA[10]. We have detected genetically variable mitochondrial and nuclear DNA sequences from the root region of shed, as well as freshly-plucked, single hairs; mitochondrial DNA (mtDNA) sequences have been detected in a sample from a single hair shaft. We have used three different means of DNA typing on these samples: the determination of amplified DNA fragment length differences, hybridization with allele-specific oligonucleotide probes, and direct DNA sequencing.

Cetus Corp.
1986 -1991
Human Genetics Dept.

Dept. Director: Henry Erlich

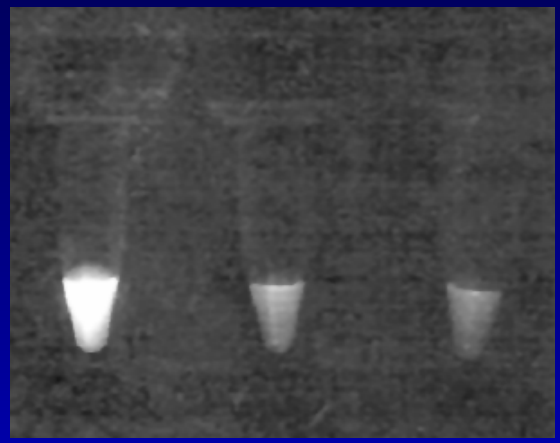
First PCR forensic DNA typing kit

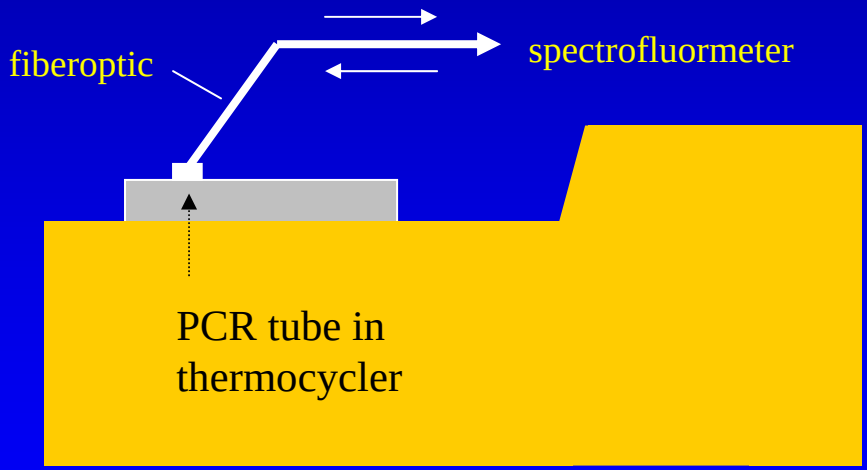
Perkin-Elmer Cetus Amplitype DQα

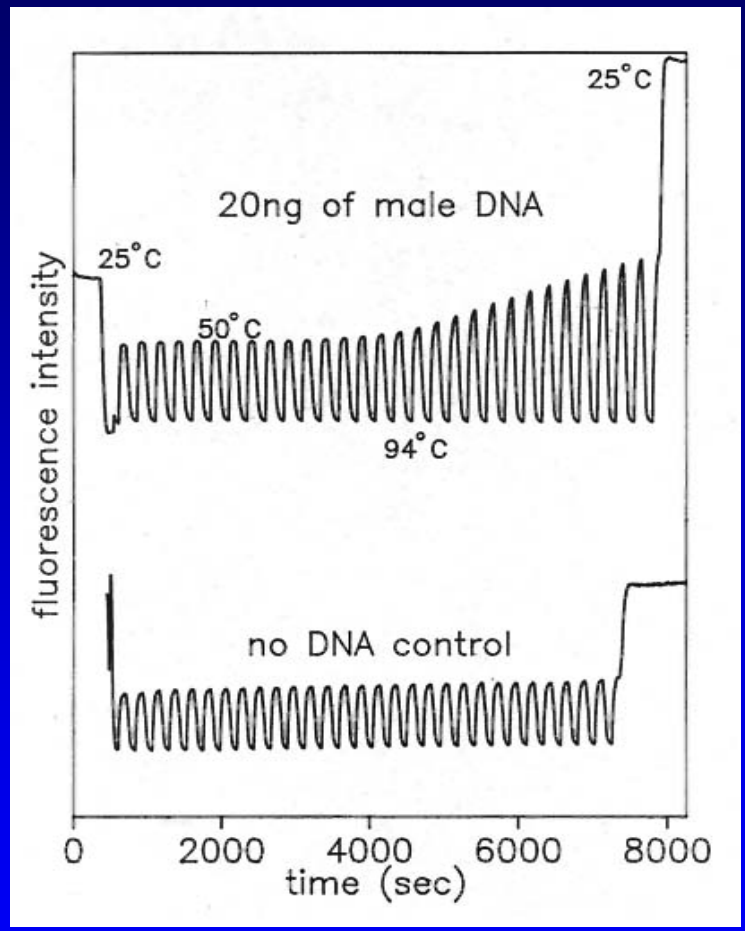
(circa 1990)



First real-time PCR
1991





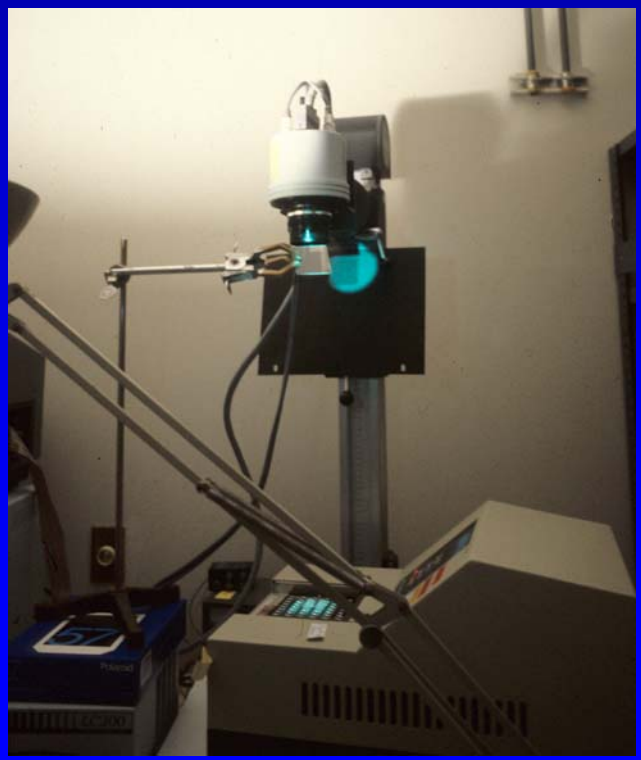


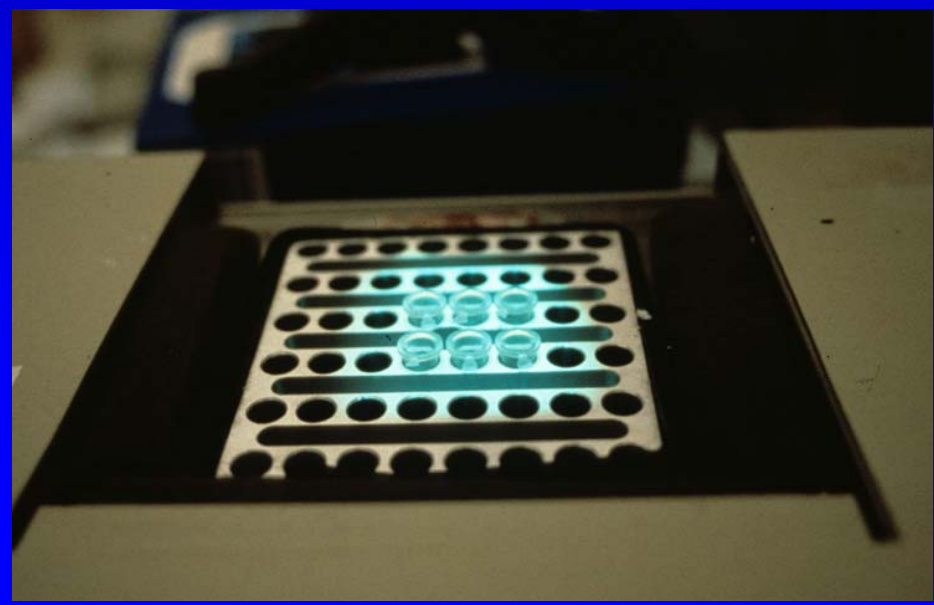
Fifty Years of Molecular (DNA/RNA) Diagnostics
Clin Chem. 2005 Mar;51(3):661-71 (C.Wittwer, ed.)

CCD camera approach 1991/92

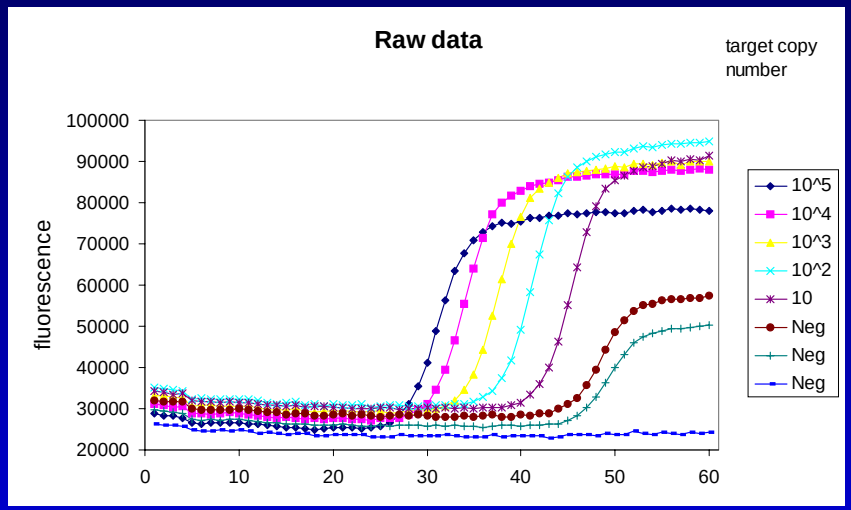


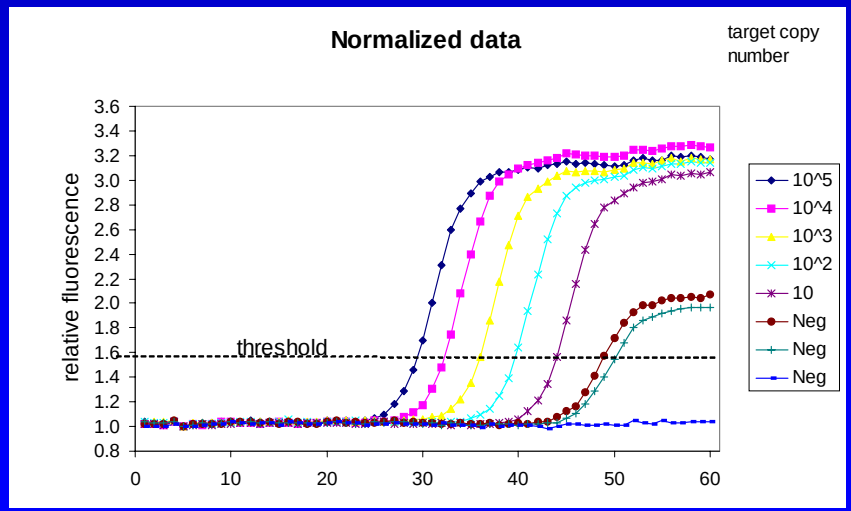


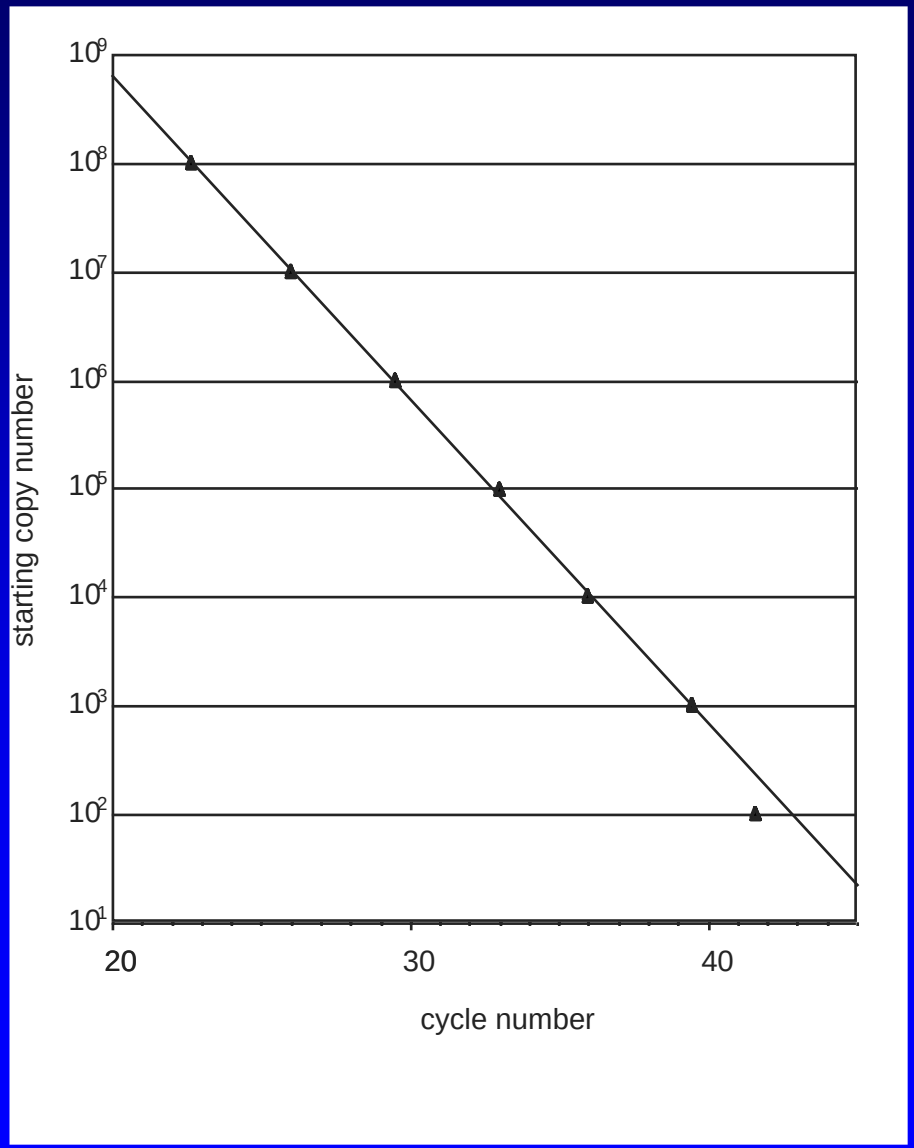




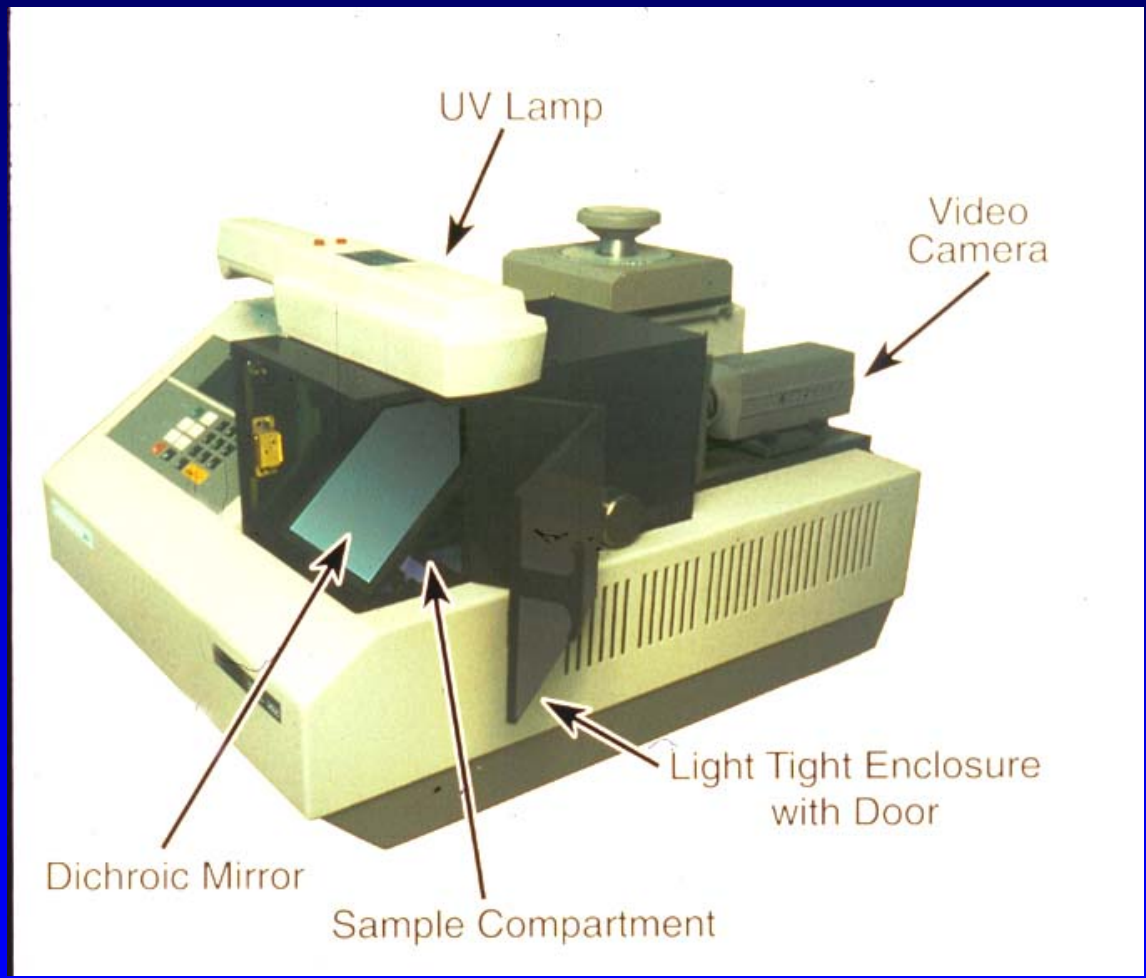
Normalized growth curves from CCD camera approach

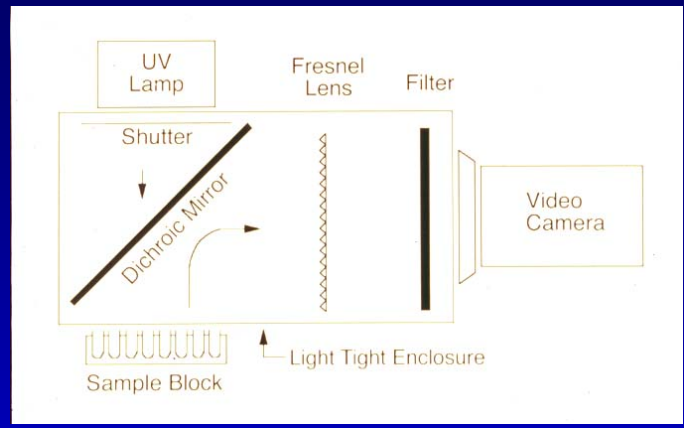






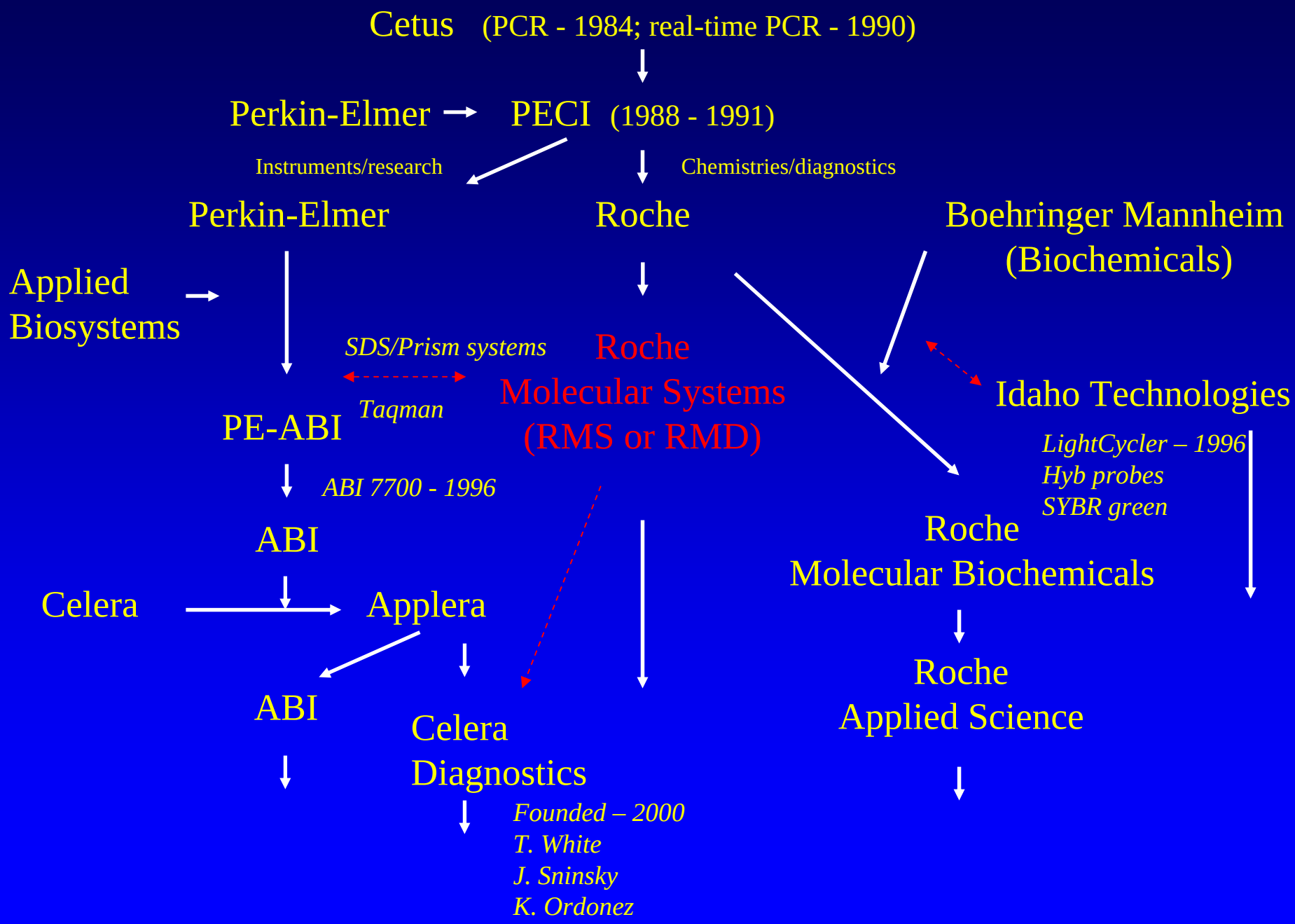
Integrated CCD camera approach - 1995







PE-ABI 5700 - 1997


Cetus (PCR - 1984; real-time PCR - 1990)
Perkin-Elmer
PECI (1988 - 1991)
Instruments/research
Chemistries/diagnostics
Perkin-Elmer
Roche
Boehringer Mannheim (Biochemicals)
Applied Biosystems
SDS/Prism systems
Taqman
Roche Molecular Systems (RMS or RMD)
Idaho Technologies
LightCycler – 1996
Hyb probes
SYBR green
PE-ABI
ABI 7700 - 1996
ABI
Celera
Applera
Roche Molecular Biochemicals
ABI
Celera Diagnostics
Roche Applied Science
Founded – 2000
T. White
J. Sninsky
K. Ordonez

Part II – Two things about real-time PCR

- Why microarray hybridization cannot be as sensitive or have the dynamic range of PCR
- Why thermocycling is an advantage over isothermal amplification

Application of real-time PCR to quantitative genetics

Accelerated Publication

THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 277, No. 17, Issue of April 26, pp. 14363–14366, 2002
© 2002 by The American Society for Biochemistry and Molecular Biology, Inc.
Printed in U.S.A.

Transcript Abundance in Yeast Varies over Six Orders of Magnitude*

Received for publication, February 15, 2002,
and in revised form, March 4, 2002
Published, JBC Papers in Press, March 6, 2002,
DOI 10.1074/jbc.C200101200

Michael J. Holland‡

From the Department of Biological Chemistry, School of Medicine, University of California, Davis, California 95616

In the current era of functional genomics, it is remarkable that the intracellular range of transcript abundance is largely unknown. For the yeast *Saccharomyces cerevisiae*, hybridization-based complexity analysis and SAGE analysis showed that the majority of yeast mRNAs are present at one or fewer copies per cell; however, neither method provides an accurate estimate of the full range of low abundance transcripts. Here we examine the range of intracellular transcript abundance in yeast using kinetically monitored, reverse transcriptase-initiated PCR (kRT-PCR). Steady-state transcript levels encoded by all 65 genes on the left arm of chromosome III and 185 transcription factor genes are quantitated. Abundant transcripts encoded by glycolytic genes, previously quantitated by kRT-PCR, are present at a few hundred copies per cell whereas genes encoding physiologically important transcription factors are expressed at levels as low as one-thousandth transcript per cell. Of the genes assessed, only the silent mating type loci, *HML* and *HMR*, are transcriptionally silent. The results show that transcript abundance in yeast varies over six orders of magnitude. Finally, kRT-PCR, cDNA microarray, and high density oligonucleotide array assays are compared for their ability to detect and quantitate the complete yeast transcriptome.

Mike Holland
Dept. of Biological Chemistry
UC Davis

Overview – Thing I

- Kinetics of DNA hybridization
- Kinetics of kinetic (real-time) PCR
- Why you can’t get there (low copy detection) from here (hybridization on microarrays)

DNA hybridization kinetics in solution

$$\frac{d[\text{Crick}]}{d\text{T}} = -k[\text{Watson}]\,[\text{Crick}]$$

The rate of disappearance of Crick (appearance of dsDNA) at any time (T) is proportional to the concentration of Crick and of Watson

Hybridization kinetics when [Watson] >> [Crick]

$$\frac{d[\text{Crick}]}{d\text{T}} = -k[\text{Watson}_\text{o}][\text{Crick}]$$

[Watson] essentially does not diminish and can be substituted for by the constant initial concentration [Watson_o]. The reaction profile over time is determined by the diminishing [Crick]

Wetmur, JG, *Critical Reviews in Biochemistry and Molecular Biology*, **26**:227-259 (1991)

Hybridization of cDNA to immobilized probe approximates conditions where [Watson] >> [Crick] where Watson = cDNA and Crick = probe

Caveats:

- Works if [cDNA] is high enough
- If not, hybridization is also diffusion-limited and the rate is slower
- This is why we gently rock our hybridizations

Immobilized probe-based equation

$$\frac{d[\text{ssProbe}]}{d\text{T}} = -k[\text{cDNA}_\text{o}][\text{ssProbe}]$$

The rate of disappearance of single-strand probe (appearance of bound target sequence) at any time (T) is proportional to the remaining ssProbe

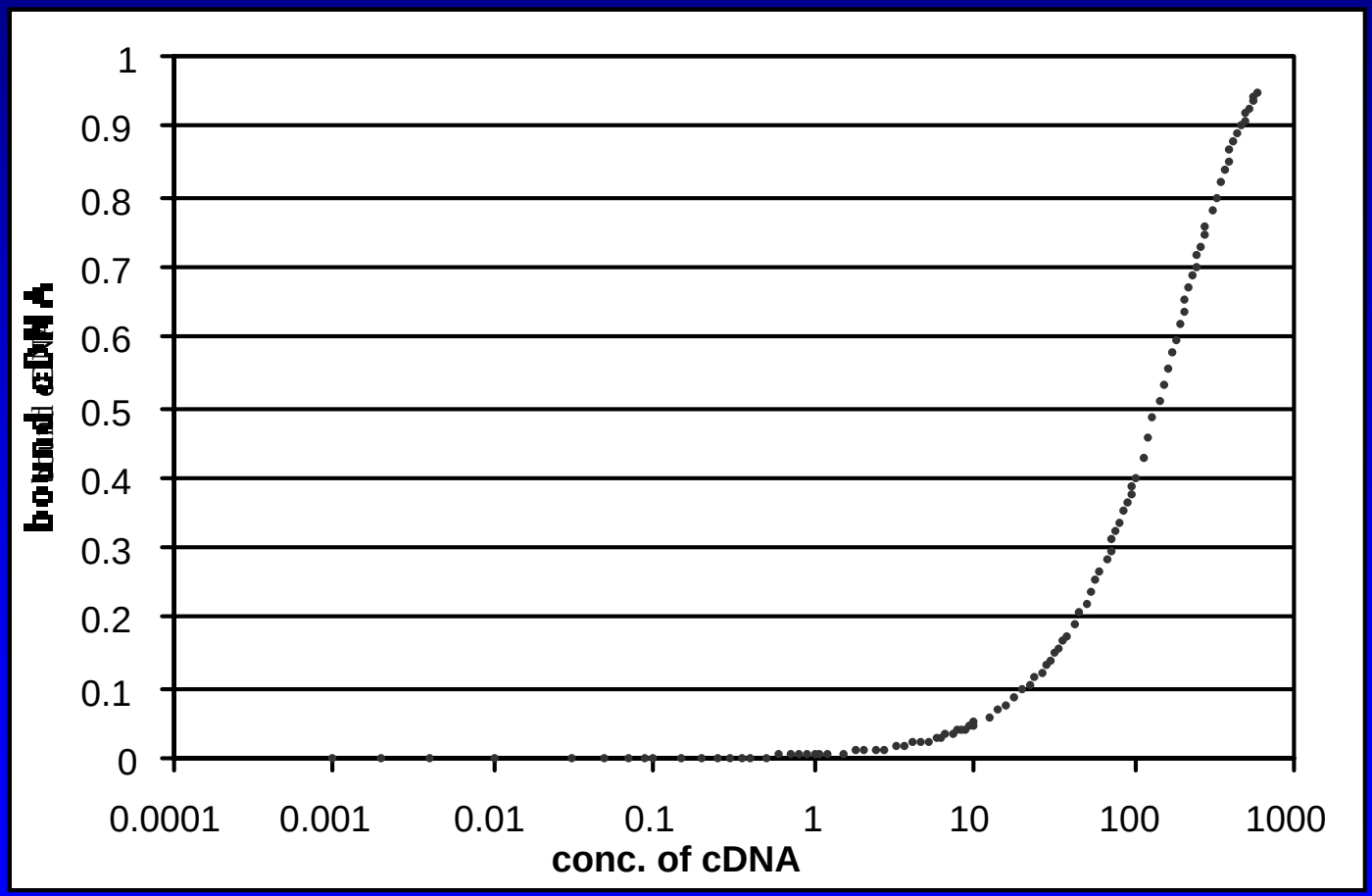
Integrate the equation from T_o to T

$$\frac{[\text{ssProbe}]}{[\text{ssProbe}_o]} = e^{-kT[\text{cDNA}_o]}$$

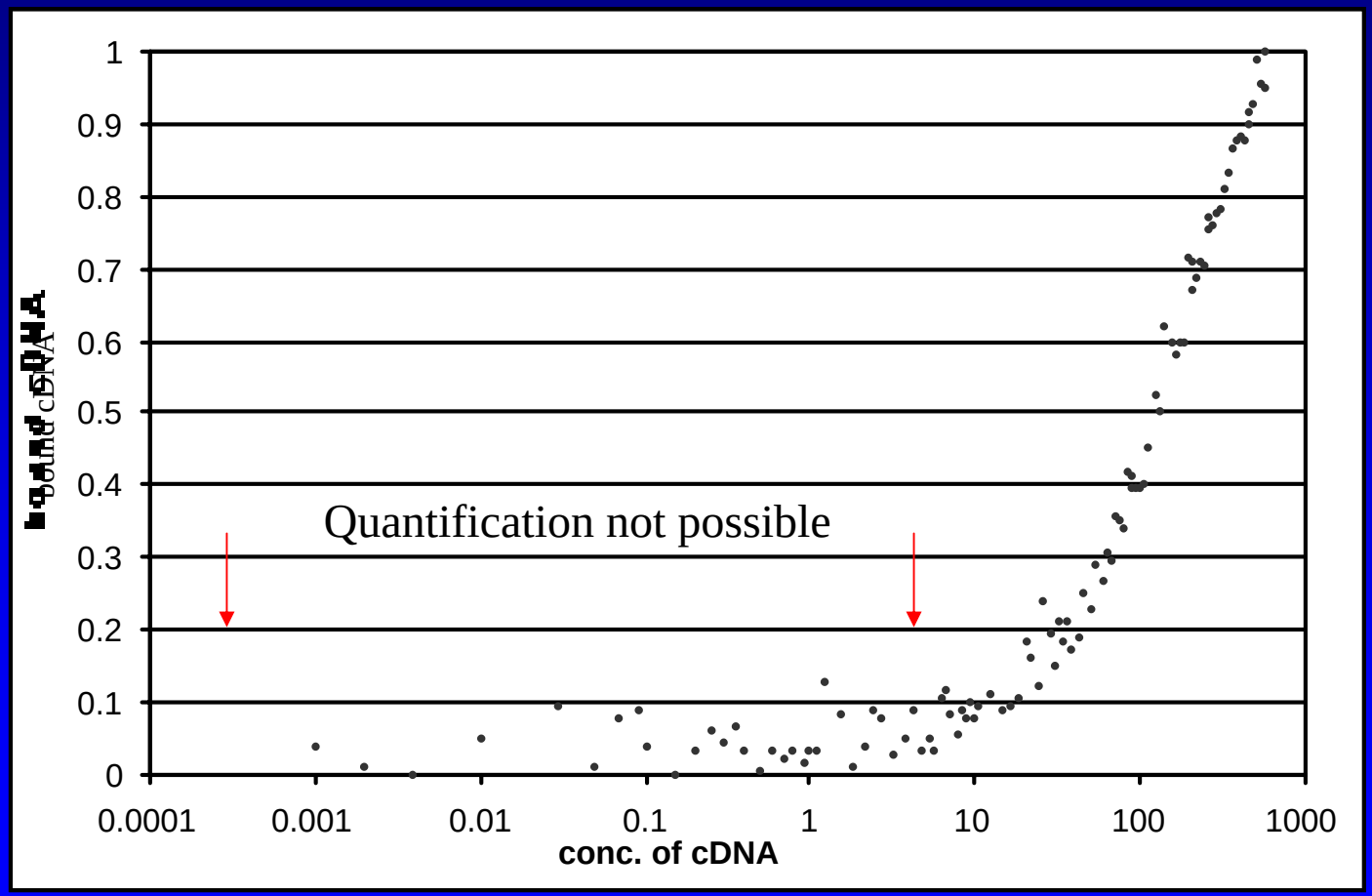
The left side, the fraction of remaining unbound, ssProbe, goes from 1 to 0 with increasing time. One could think of the amount of bound cDNA increasing from 0 to 1 at the same time

$$\text{bound cDNA} = 1 - \frac{[\text{ssProbe}]}{[\text{ssProbe}_o]} = 1 - e^{-kT[\text{cDNA}_o]}$$

Let's graph, keeping T constant and varying [cDNA]



Let’s add some noise (S/N = 20)



Hold this thought: an RNA profile on a microarray is a hybridization experiment with a mixture of multiple different, sequence-specific cDNAs at widely different concentration (> 6 logs).

In other words, the experiment we just graphed

Atwood model of primer-limited PCR

$$C_{n+1} = C_n \times \left(1 + \frac{\text{eff.}}{1 + \frac{aC_n}{P_o - C_n}} \right)$$

C_n = product copy# @ cycle n

eff. = initial replication efficiency

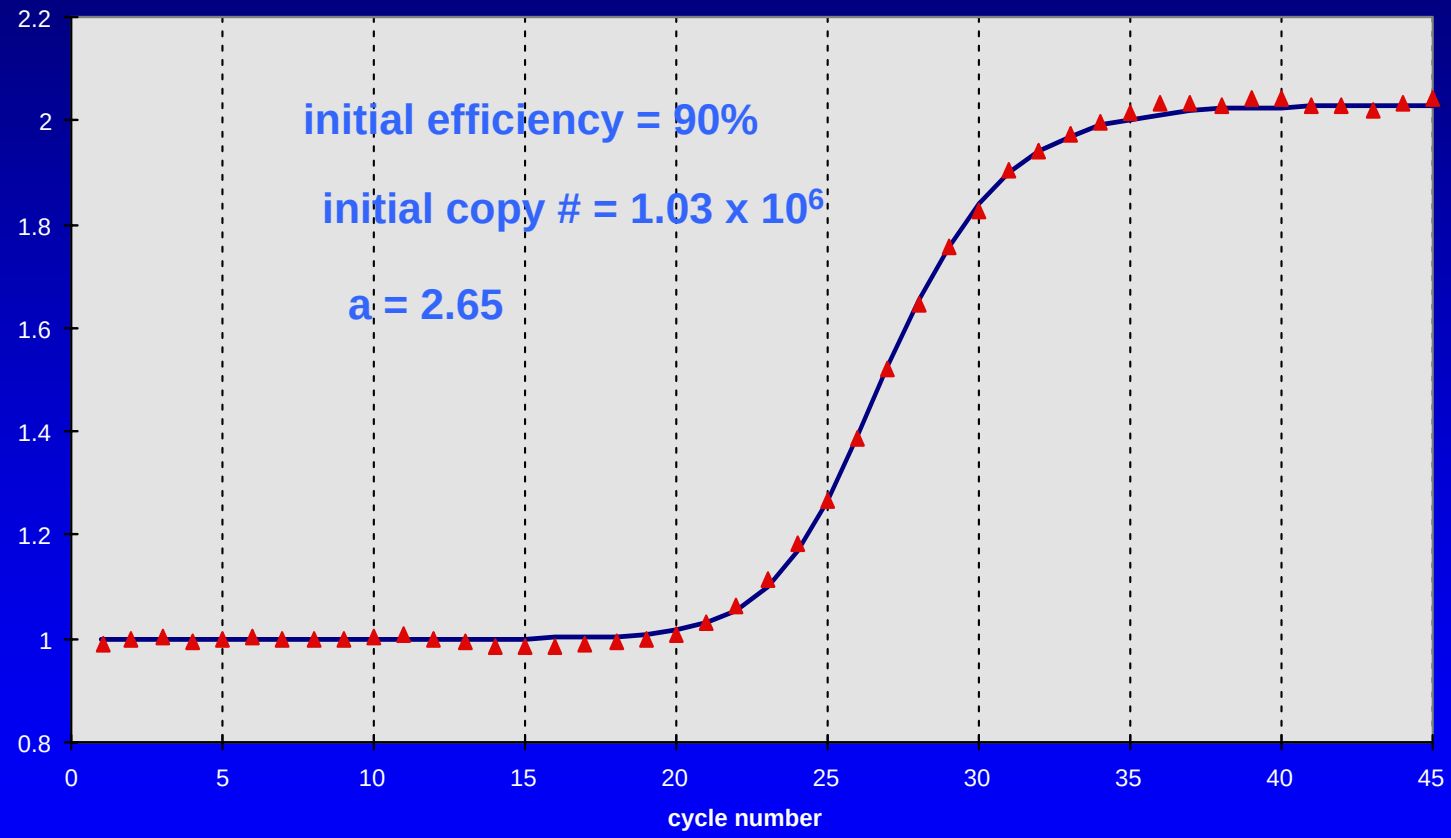
P_o = initial primer concentration

a = strand reassociation vs. primer annealing

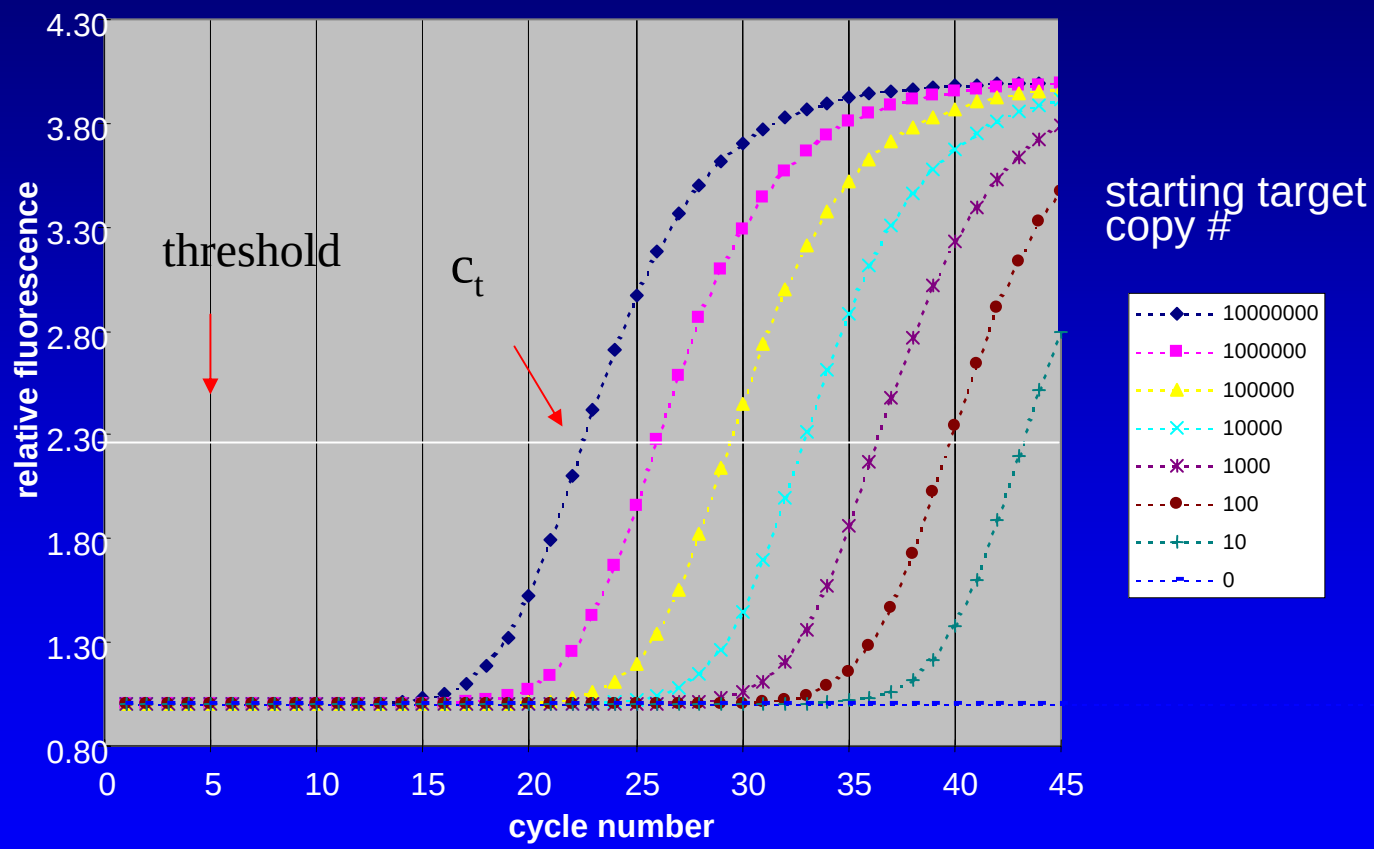
Conversion to detected relative flourescence

$$\text{rel. fluor}_{\text{cycle i}} = 1 + (\text{copy\#}_{\text{cycle i}} / \text{copy\#}_{\text{cycle infinity}})(X-1)$$

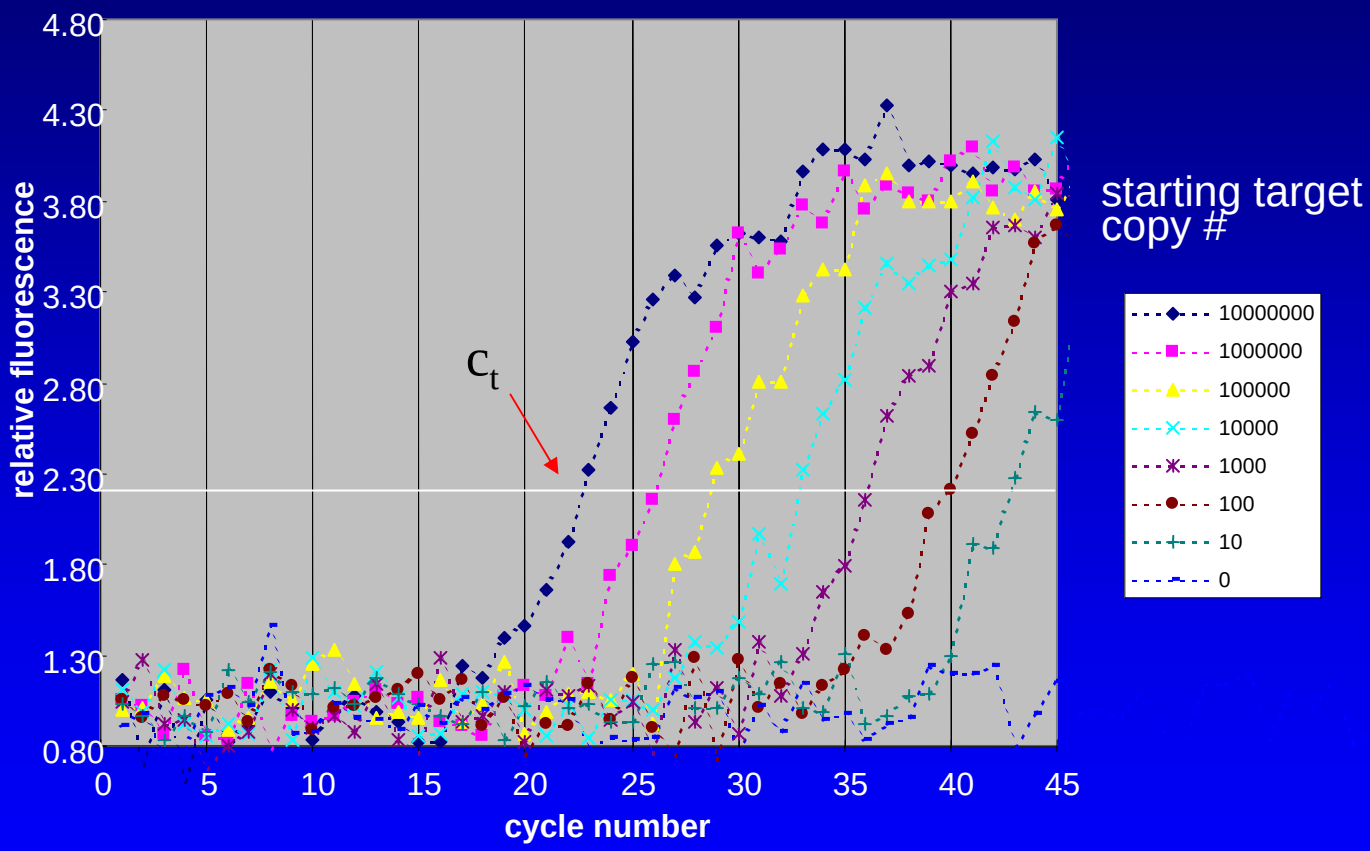
Curve-fitting the Atwood model to data



Vary intial copy number in the PCRs



Let’s add some noise (S/N = 20)

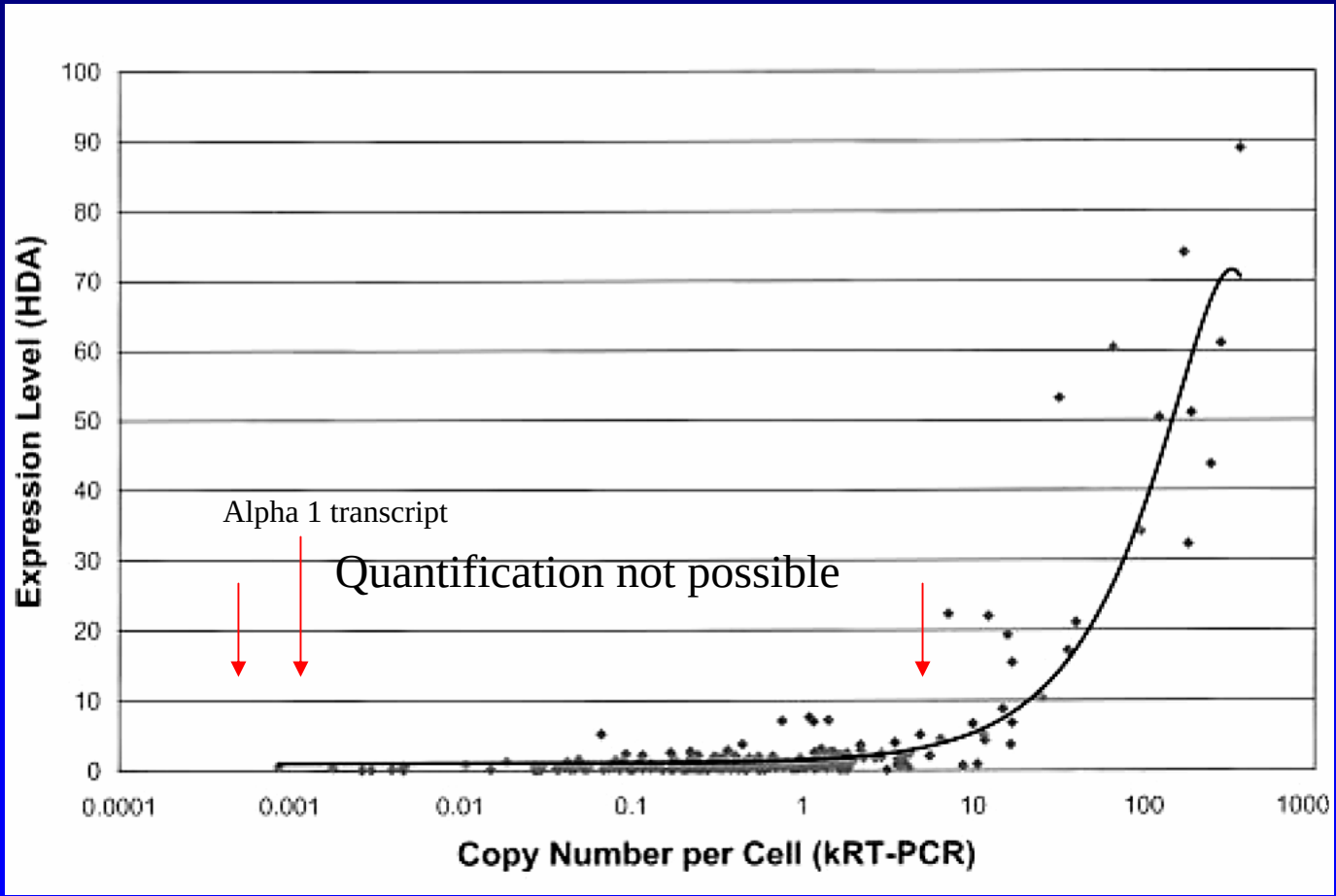


Dynamic range is unaffected

Mike Holland's experiment

- RNA profile yeast gene expression of 250 genes using kinetic RT-PCR
- Early log phase cultures of strain BY4742 in YPD medium
- Compare to published profile of same genes from same strain under same conditions measured using microarray hybridization

Microarray hybridization measurement of mRNA level in yeast vs kinetic RT-PCR measurement



M. Holland, *J. Biol. Chem.* **277**:14363-66 (2002)

One can conclude from this:

- Yeast mRNA transcripts vary over six orders of maginitude
- Lowest level transcripts are at 1 copy per 1000 cells
- These transcripts are probably functional
- Microarrays are three orders of magnitude in sensitivity away from detecting these lowest level transcripts
- Problem should be worse in mammalian cells

How can mRNAs at < 1 copy per cell mean anything?

- mRNAs are labile; proteins are the more stable effectors of function
- Measuring the average level in many cells at different stages of cell cycle - catching only a few cells “in the act” of transcription
- Tissues are heterogeneous
- Regulatory proteins are low-level (lac repressor at few copies/cell)

To enhance dynamic range, could you do the same trick with hybridization, *i.e.*, make measurements over time?

- Yes, but to get greater sensitivity need to hybridize longer
- 10-fold greater sensitivity requires 10-fold longer hybridization
- Hybridization times are already 24 hr
- Longer hybridization = higher background fluorescence

Could you increase the concentration of cDNA?

- Yes, but many microarray experiments have limited mRNA to begin with and use as small a volume hybridization as is reasonable to work with
- 10-fold gains in sensitivity require 10-fold increases in concentration
- “Volume excluders”, *e.g.*, PEG, are already used

Conclusions

- Many interesting mRNAs are below the level of detection by microarrays
- The dynamic range of quantification by microarray hybridization is limited
- Kinetic RT-PCR has a much broader dynamic range and much greater sensitivity
- The number of mRNAs testable with microarrays is much larger in one experiment (bigger “bandwidth”)

Having your cake and eating it too

- A dense microarray of PCRs
- Microfluidic partitioning of sample and master mix

Thing II - why is thermocyling good for quantification? (or why is the “q” in qPCR as good as it is?)

- Exponential processes like PCR are not supposed to be easily reproducible
- Small efficiency differences early on (due, say, to temperature differences) can be magnified into large differences by late cycles
- Common misperception
- The “stop/start” of temperature cycling synchronizes all reactions
- Advantage of PCR over isothermal amplifications

Big Problem in gene discovery by "genome-wide SNP association" - Huge number of genotype determinations required

- 200,000 to 1,000,000 SNPs required for genome-wide coverage
- 500 case + 500 control x 200,000 SNPs
  = **2×10^8** genotype determinations
- Total output of RMS + collaborators, last 10 years maybe **5×10^6**

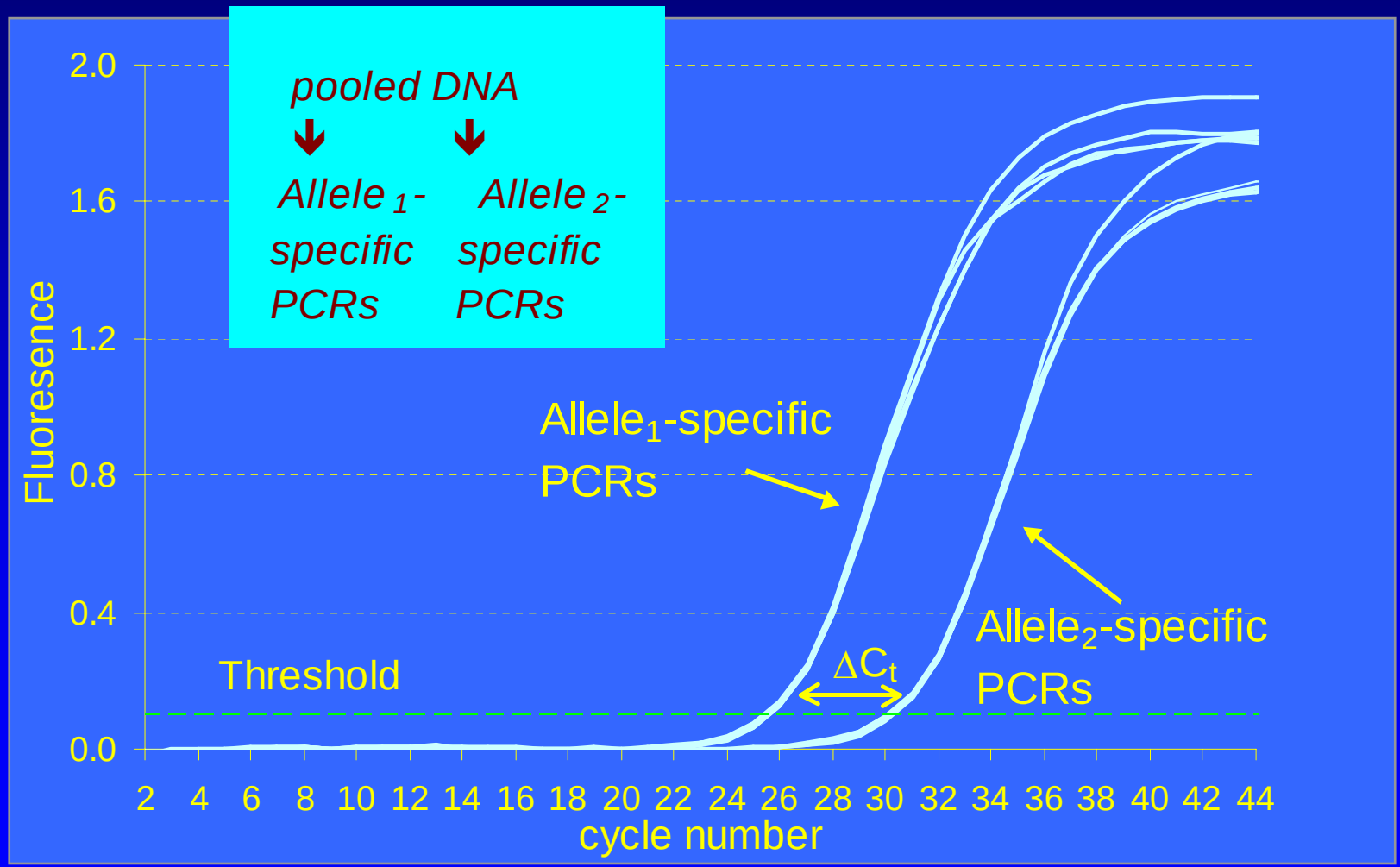
Promise of genome-wide SNP screens: genes for genetically complex disorders - heart disease, cancer, osteoporosis, etc.

How to do a billion genotypes

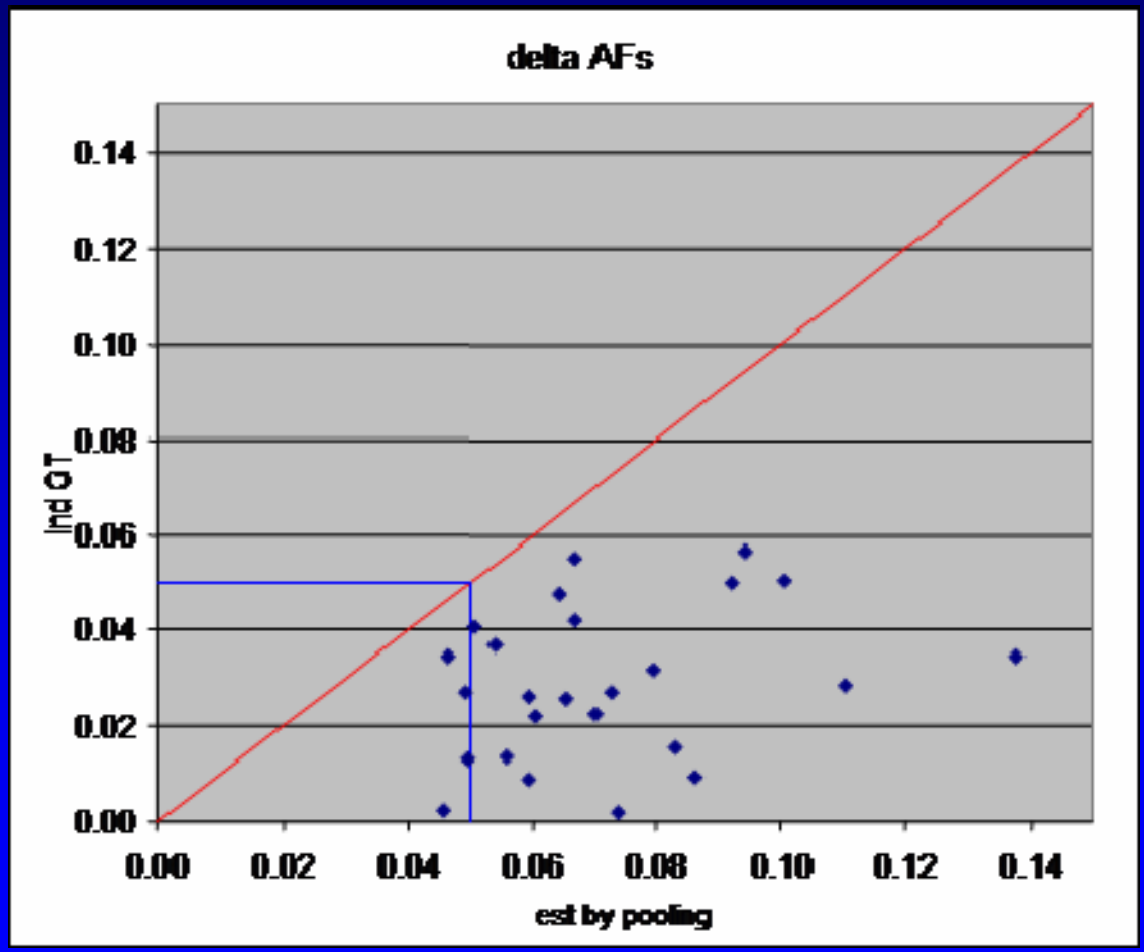
- Multiplex the DNA samples (i.e., pool equal amounts of them)
  - Use allele frequency in pool as measure of allele frequency in population
  - Look for allele frequency difference between case and control larger than expected “by chance”
  - Erlich, H, Arnheim, N and Strange, C (1985). Proc. Natl. Acad. Sci. 82: 6970–6974. – use of Southern Blot RFLP analysis
  - Germer, S, Holland, M and Higuchi, R (2000) Genome Res. 10(2):258-66. – use of real-time PCR
- Multiplex the assays (as many genotypes as possible per assay) - AffymetrixTM, ParalleleTM
- Multiplex both - PerlegenTM

Principle of Our Pooling assay

$$\text{Frequency of allele}_1 = \frac{1}{(2^{\Delta C_t} + 1)}$$



Exacerbation of “winner’s curse”



How about limiting to amino acid substitutions?

- Some 20,000 missense SNPs in databases
- Botstein and Risch
  - Most known Mendelian (single-gene) disorders are due to misense mutations – why should complex (multi-gene) disorders be different?
  - Ascertainment bias (you get what you look for)
- Celera Diagnostics (J. Sninsky)
  - Allele-specific, kinetic PCR



“Mama always said life was like a box a chocolates, never know what you're gonna get.”

- Forrest Gump