# Molecular Biology in Food Hygiene


## Ivo Papoušek


16 July 2019

# Why molecular biology?

Today, there is a number of issues of importance due to public health, economic and legal concerns, which are impossible or unneccessarily difficult to solve without molecular biology. Among the most important are:

- pathogens in food (bacteria, viruses, parasites…)
- composition of food (authenticity, allergens…)

In general, molecular biology is used to **detect presence, identify and/or quantify a specific agens** in food based on analysis of **biomacromolecules**.

- an overview of selected molecular techniques currently used
- examples of problems to be solved

# Why molecular biology?

Historically, tests directed to identification of **proteins**

But: - low specificity

- often unsuitable for processed food (***heating***, chilling, salting, seasoning) – structural modification of proteins

Current molecular techniques are based on analysis of **DNA**

- DNA provides more information
- is independent of the tissue it resides in -> it does not matter whether you sample milk, blood, meat, liver… (it does matter for proteins)

DNA-based techniques are generally **more robust, very sensitive** and can be **extremely specific**

# DNA – deoxyribonucleic acid

- the (almost) universal genetic material

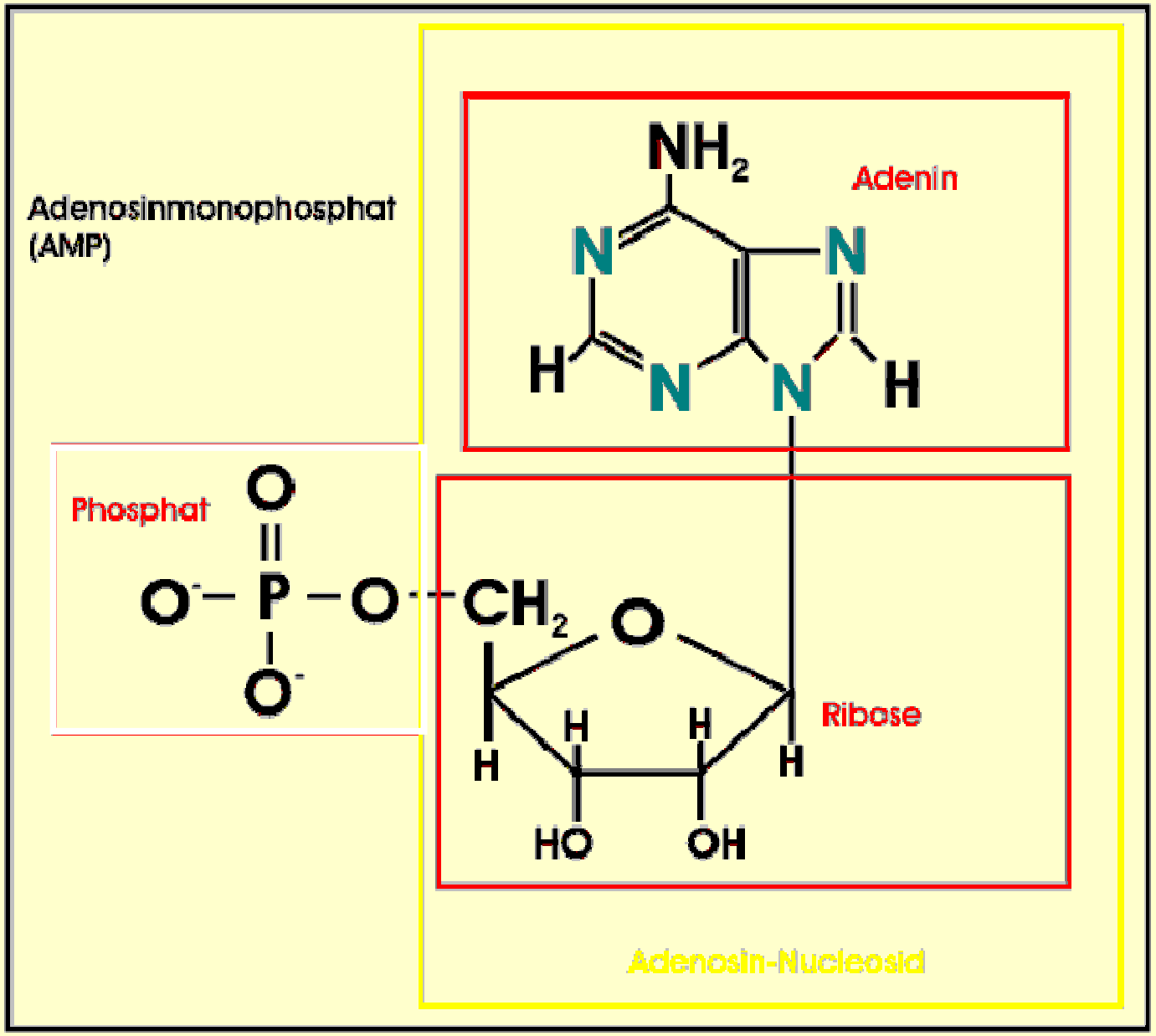
- biopolymer: polynucleotide chain consisting of monomers (*nucleotides*) linked by phosphodiester bond

- nucleotide: phosphate + deoxyribose + nitrogenous base

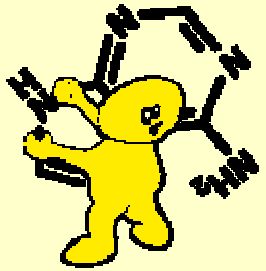
- sequence of nucleotides in DNA is called primary structure and encodes the genetic information

- double-stranded (consisting of two chains which carry complementary nucleotides)

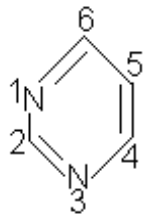
# Nucleotide


Adenosinmonophosphat
(AMP)
Adenin
Phosphat
Ribose
Adenosin-Nucleosid

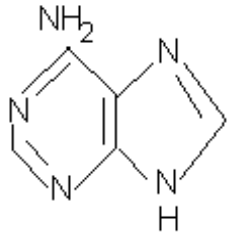
# Bases



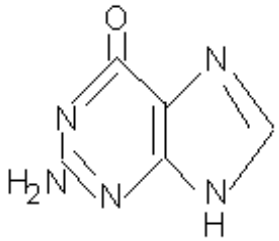
purine



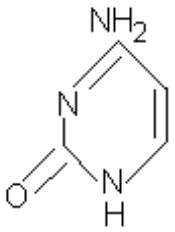
pyrimidine



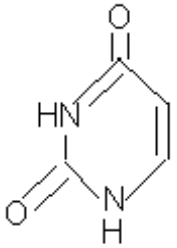
adenine



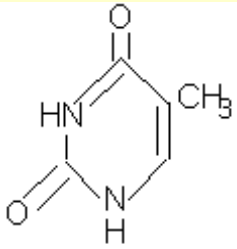
guanine



cytosine



uracil



thymine

# DNA

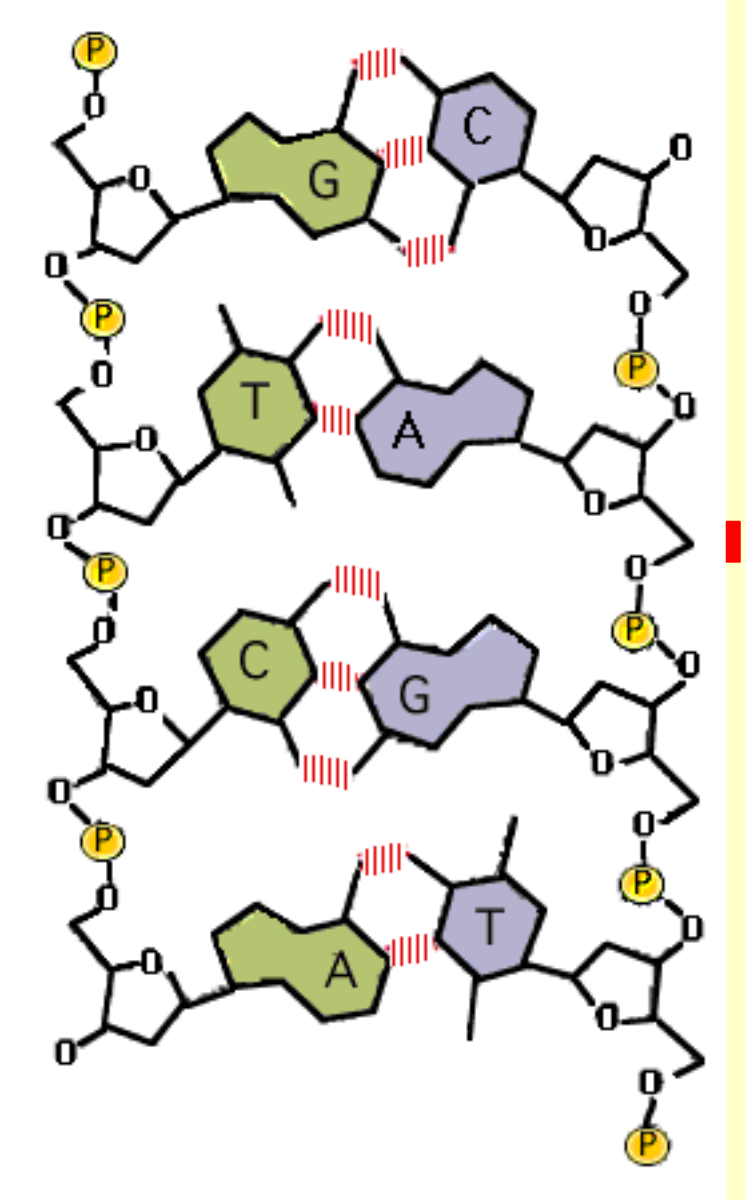
**primary structure**

**=**

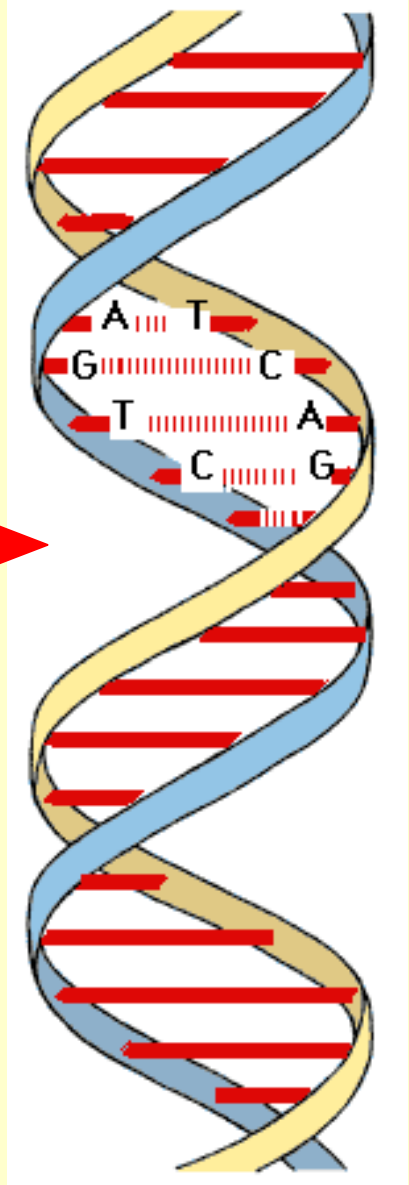
**nucleotide sequence**

**=**

**genetic info.**




**secondary structure**



# DNA in food hygiene

Basically we are interested in detection of genetic information, which is **very specific**:

if we look at certain genes, we find that their sequence variants are **species-specific** (relatively low intra-species and high inter-species variability)

–> they enable us to **distinguish** material coming from **distinctive species** (or other taxonomic units)

# Pitfalls and challenges

1. **small quantities** of target DNA - either absolute (small sample) or relative (mixed samples)
2. many compounds in food can be **inhibitory for molecular methods**

solutions:

a) **samples enrichment** (centrifugation of milk, extraction of only bacterial/viral DNA from food samples…)

b) **highly efficient DNA extraction** – recovery of quality DNA from vast array of complex food matrices (meat, milk, canned food…)

c) **selection of a suitable marker/target gene**

- nuclear vs mitochondrial DNA: mtDNA present in multiple copies -> increased likelihood of detecting specific sequences

- however number of copies of mtDNA variable among species, individuals or even tissues of the same individual –> problematic for quantification methods

- **both are used, in different situations**

# Pitfalls and challenges

3. **degradation of samples** due to thermal treatments (cooking), high pressure, pH modification… -> fragmentation of DNA
   -> sometimes we can use only **very short fragments** (100-200 bp), **suitable for selected techniques only**; less specific

   interesting fact: even meat boiled (100 °C) or roasted (200 °C) for several hours may provide DNA usable for analysis

4. **sensitivity** – generally desired, but if too sensitive, false positive results (e.g. due to cross-contamination – processing of different products in the same factory)

# Polymerase chain reaction (PCR)

- developed in 1983 (Kary B. Mullis) → 1993 Nobel Prize in Chemistry
- method generating thousands to millions of copies of a particular (specific) DNA sequence ***in vitro***; based on replication
- absolutely crucial, basic technique, often followed by subsequent methods.

# What we need for PCR (reaction compounds)

- **DNA template**: DNA to be copied/detected. New copies are synthesized based on principles of complementarity from ssDNA, which serves as a matrix.

- **primers**: artificially synthesized oligonucleotides (15-30 nucleotides) which are complementary to the beginning and to the end of the target region of DNA – they flank the target region that will be amplified. **They are crucial for the specificity of reaction!!!**

- **DNA polymerase**: a thermostable enzyme which performs actual synthesis of new copies of DNA

- **nucleotides**: building blocks linked by DNA polymerase into a chain

PCR is a cyclical process (chain reaction) :
each new copy can be copied again

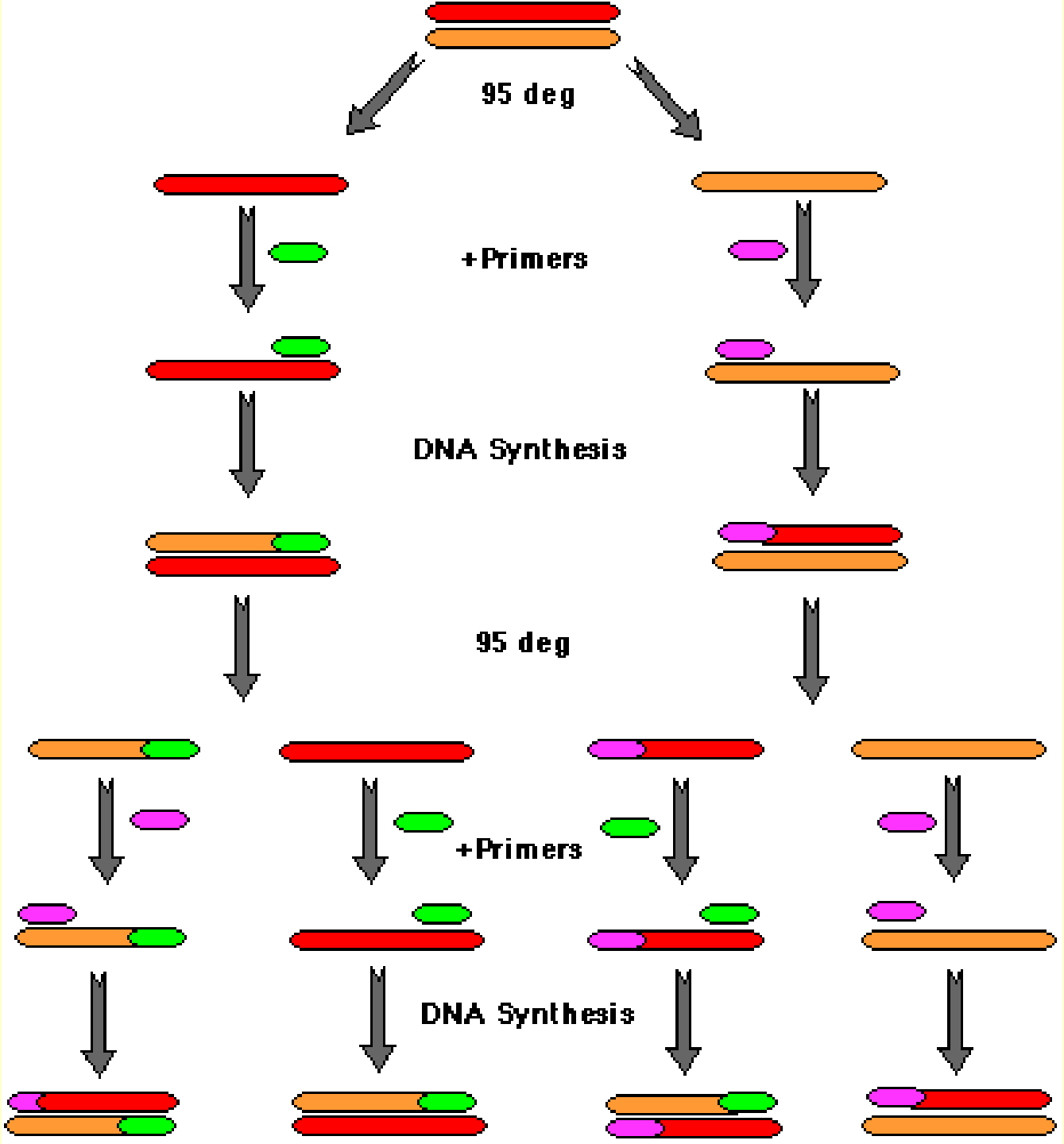
Denature DNA sample
to separate DNA strands
(94°C, 5 min)

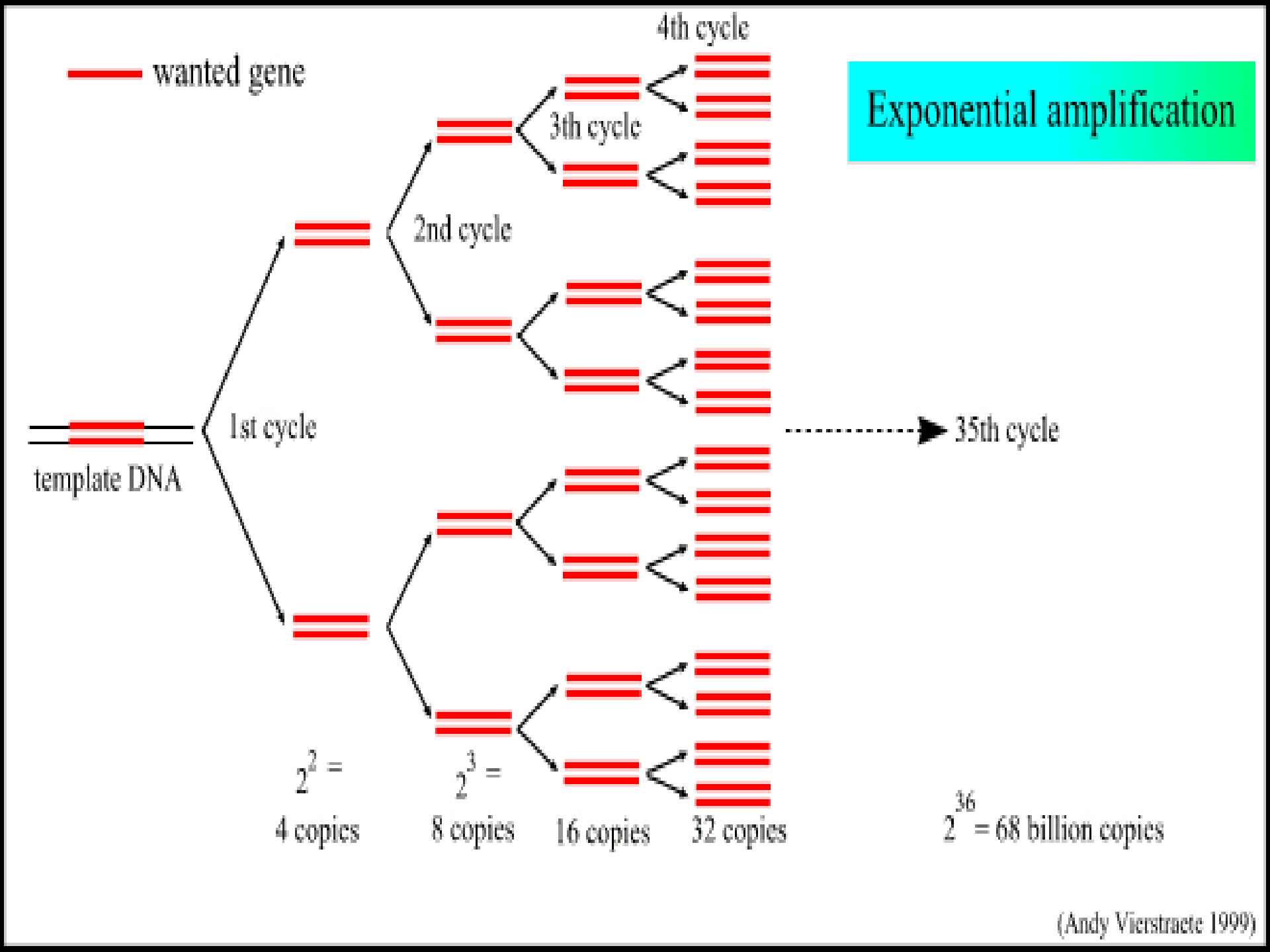
Primers bind to
DNA strands
(30–65°C, 30 s)

25 – 35 ×

Polymerase synthesizes
new DNA strands
(72°C, 45 s–2 min)

Denature to separate
DNA strands
(94°C, 30 s)


95 deg
+Primers
DNA Synthesis
95 deg
+Primers
DNA Synthesis




# Exponential amplification

(Andy Vierstraete 1999)

# Detection of amplified PCR product

- Usually by **gel electrophoresis**:
- Sample is put into a gel (agarose, polyacrylamide) and electrical field is applied
- DNA is negatively charged –> will travel through the gel to positive electrode
- Smaller fragments will travel faster (i.e. we can separate fragments of different sizes)

# Detection of amplified PCR product

- DNA fragments (if present) are stained (usually by a dye dissolved in the gel; the dye binds to passing molecules of DNA) and subsequently visualized (e.g. by UV-light transillumination; the stain fluoresces)

- Generally we are interested in **detection of fragments of the correct size** - we compare their position on the gel relative to a „marker“ (set of fragments of known size)


1
2
3
4
5
6
7
8
M
9
10
11
12
13
14
longer fragments
shorter fragments

# Selected types of PCR

- **Simplex PCR**: you use one pair of primers specific for the target sequence/species. Either you obtain desired product or not -> the target agens is present/absent in tested sample.

- **Multiplex PCR**: several sets of primer pairs within a single reaction

  -> simultaneous detection/identification of several target sequences in a single sample

  -> discrimination of several species at the same time

  - different species identified by presence of products differing in length (position on electrophoretic gel)

  - respective products should differ **at least by 40-50 bp** for adequate resolution

# PCR + RFLP

## RFLP – restriction fragment length polymorphism

- Principle: DNA amplified by PCR is digested (cleaved) by an appropriately selected restriction endonuclease; products are analysed by electrophoresis

- Restriction endonucleases: enzymes cleaving DNA in specific sequences (restriction sites)

- Because different organisms differ in sequence of DNA, they will differ in presence/absence of restriction sites -> will differ in spectrum of resulting fragments

# PCR + RFLP

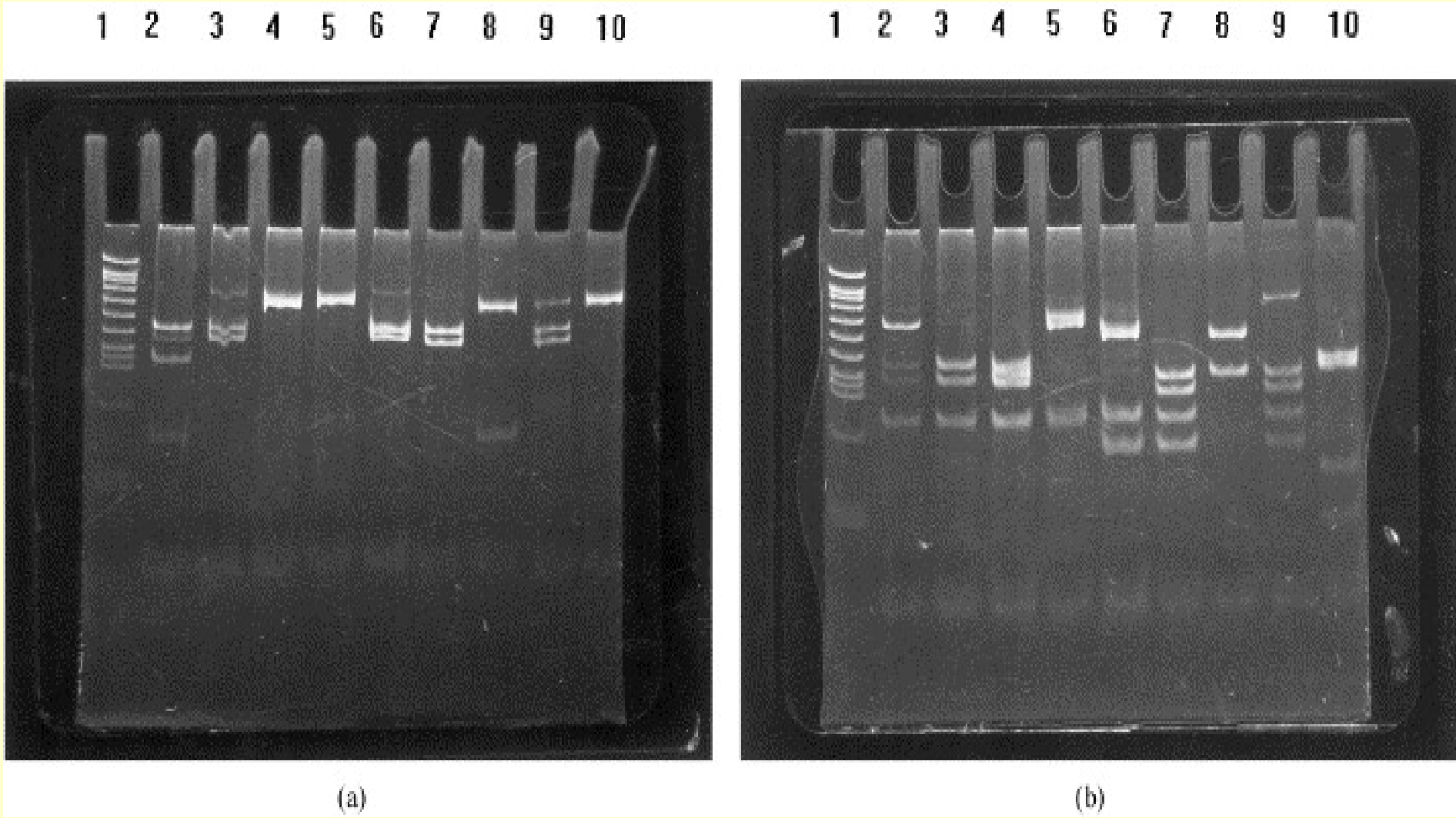
## RFLP – restriction fragment length polymorphism

Advantages:

- I can use a **single set of primers** universal for all tested species
- With properly selected enzyme, **relatively short PCR product is sufficient** (i.e. method applicable for cooked/fried/pateed meat etc.)

Disadvantages:

- **Difficult** discrimination between **closely related species** – often necessary to combine more enzymes – time and money-consuming
- Risk of incomplete digestion – problematic inconclusive results




(a) *EcoRI1* and (b) *BamH1* restriction profiles obtained from PCR-RFLP analysis of Centrometric Satellite DNA in cooked meat from nine animal species. Lane numbers 1=molecular marker; 2=cattle; 3=zebu; 4=banteng; 5=buffalo; 6=bison; 7=wisent; 8=water buffalo; 9=african buffalo; 10=Indian cattle.

# PCR + sequencing

- We obtain direct information about primary structure (sequence of nucleotides)
- Sequence is compared with database – direct identification of respective species

Advantages:

- Many available markers sequenced for most species of interest
- Reliable, reproducible

Disadvantages:

- Relatively expensive
- Can‘t be (easily) used with universal primers (e.g. for all bacteria) for mixed samples – provides unreadable mixed sequence
- Problematic in processed food – to distinguish between closely related species you often need to sequence several hundreds bp

# Next Generation Sequencing (NGS)

• demand for low-cost sequencing caused development of „high-throughput“ methods; today the preferred term is „High Throughput Sequencing“ (HTS)

• principle – **parallelization of sequencing process**

• production of **thousands to millions sequences** at once!

• a **huge volume of data** to be evaluated/sorted

• for the first time, it is harder to **interpret data** than to **obtain them**

• general use: large quantity of sequence information (**whole genome data**; or **many copies of a particular short DNA sequence**; or **many different genomes at once** - *metagenomics*)

# Real-time (quantitative) PCR

- Enables both **qualitative analysis** (presence/absence) and **quantification** of agens in tested sample

- During PCR amplification, fluorescent dye-labelled probes release the dye which fluoresces (several different systems, principles)

- Intensity of fluorescent signals is proportional to the quantity of the PCR amplicon produced in each reaction cycle

  -> higher content of target sequence -> more PCR product -> stronger signal

# Real-time (quantitative) PCR

Advantages:

- **Direct detection of PCR product** – we can skip electrophoresis and gel staining
- **Extremely sensitive** – detection of microquantities (pg of DNA!)

Disadvantages:

- **Extremely sensitive** – danger of contamination and false positive results
- **Detection of nonspecific amplicons**
- **Relatively expensive**

# Problems solved by molecular methods

## 1. presence of pathogens in food

Microbial contamination can lead to serious health problems!!!

a) Bacteria

- *Listeria* - soft cheese, raw milk, salads
- *Campylobacter* – foodborne gastroenteritis – dairy products, meat
- *Escherichia coli* – most thoroughly studied; colitis, diarrheas, some strains potentially fatal! (enterohemorrhagic E. coli – 2011 outbreak, >50 fatalities, mostly in Germany; source: sprouted foods from an organic farm)
- *Salmonella* – undercooked food products from livestock or poultry, eggs, derived products
- *Clostridium botulinum* – produces extremely potent neurotoxins, can be fatal! Home-canned food substances, sausages, fermented dishes

Molecular methods **faster** and **more reliable** in comparison with standard microbiologic cultivation: 24-48 hours of growth time, often less than 100% specificity/sensitivity vs. several hours (PCR)

# Problems solved by molecular methods

## 1. presence of pathogens in food

b) Viruses – human enteric viruses (hepatitis A…)
challenge: present in low levels in contaminated food, cannot replicate there

c) Molds – production of mycotoxins – gastrointestinal disturbances, alteration of immune system, cancer!

d) Parasites

- trichinosis (roundworm *Trichinella spiralis*) – undercooked game, home-reared pigs
- *Toxoplasma gondii* (unpasteurized ovine/goat milk, improperly cooked meat…)

Commonly employed methods:

- Specific PCR – if the pathogen is present, we will obtain a PCR product
- PCR+sequencing – e.g. bacterial 16S-rDNA gene
- Real-time PCR

# Problems solved by molecular methods

## 2. Composition of food

**a) Food fraud identification** – the fraudulent misdescription of food contents on product labels is a widespread problem, particularly with high-value products. Main concerns are:

- Substitution of one species/product by a similar but cheaper/more readily available
- Protection of endangered species

Examples:

1) fish

a) substitution of **bluefin tunas** (prized for size, texture, colour, fat content and taste of meat) by much more common **yellowfin** and **skipjack tuna**

b) **caviar** – beluga caviar (*Huso huso*) – most expensive, rare; both economic and ecological concerns. Substituted by *Acipenser spp.* (sturgeon)

# Problems solved by molecular methods

Fresh fish – identification generally based on morphological characteristics – unreliable – many species are quite similar and show morphological plasticity

Processed fish - morphological examination fails -> molecular methods are necessary; due to ways of processing (fillets, smoked, marinated, canned, frozen, caviar…) there is no universal solution, number of different methods

Most frequently used methods:

- species-specific PCR
- PCR+sequencing (*DNA-barcoding* – sequencing of mitochondrial *COI* gene – reliably distinguishes most commercial species)
- PCR+RFLP

# Problems solved by molecular methods

2) meat - substitution of meat by cheaper/unwanted additions – previously detected based on organoleptic or histological properties (colour, texture, odour; fiber length, density, pattern) – impossible in mixtures

a) recent „affair“ of horse meat in hamburgers – beef „contaminated“ by horse meat, which is sort of taboo in many cultures

b) Italian „Mortara“ salami – goose meat replaced by other poultry (turkey, duck)

c) venison and derived products often targets for fraudulent labelling due to high commercial value

Routinely employed methods:

- simple PCR (+ sequencing)
- multiplex PCR
- real-time PCR – allows identification and quantification of trace amounts (0.5-5 pg!!!) even in multi-species mixtures (e.g. cattle, horse, pig, goat, sheep)

# Problems solved by molecular methods

## b) Assurance to consumers about their choices according to:

- lifestyle (vegetarianism, veganism)
- religious practices (e.g. absence of pork for Jews and Muslims)
- health concerns (absence of allergens – eggs, fish, crustaceans, peanuts, soybean, hazelnuts….) – often cross-contamination in foods processed in the same factory
- GMOs – specific PCR of transgene, mainly in soybean and maize-derived food products

Same methodic principles as before.

# Problems solved by molecular methods

## 3. Animal feedstuffs

- European Community Regulation banned animal derived meals (particularly meat and bone meal, MBM) in the manufacturing of ruminant feedstuffs as a preventive measure to avoid the spread of **Bovine Spongiform Encephalopathy (BSE)**

- Classic method – microscopic detection of animal bone fragments – time consuming, only class-specific (mammal, avian, fish bones), not species

- Molecular methods – much more sensitive, specific

- Species-specific PCR (bovine tissue); multiplex PCR (ruminants, pork, fish, poultry)

# Conclusions

## Which technique to use?

There is no universal  technique suitable for all conditions and purposes.

We have to decide based on:

- Available material – Raw? Processed? How?
- Required specificity/sensitivity of detection – closely related species? Presence/absence of a single certain species? How many species need be discriminated?
- Expenses/Budget – PCR is quite cheap, sequencing and real-time PCR expensive
- Numbers of samples