

# Medicinal Plants: DNA-Based Identification of Raw and Processed Materials



Johannes Novak  
Institute for Applied Botany and Pharmacognosy  
University of Veterinary Medicine  
Veterinärplatz 1  
A-1210 Wien, Austria  
e-mail: [Johannes.Novak@vu-wien.ac.at](mailto:Johannes.Novak@vu-wien.ac.at)



**GRAZ**  
2007

Society for Medicinal Plant Research  
Gesellschaft für Arzneipflanzenforschung – GA

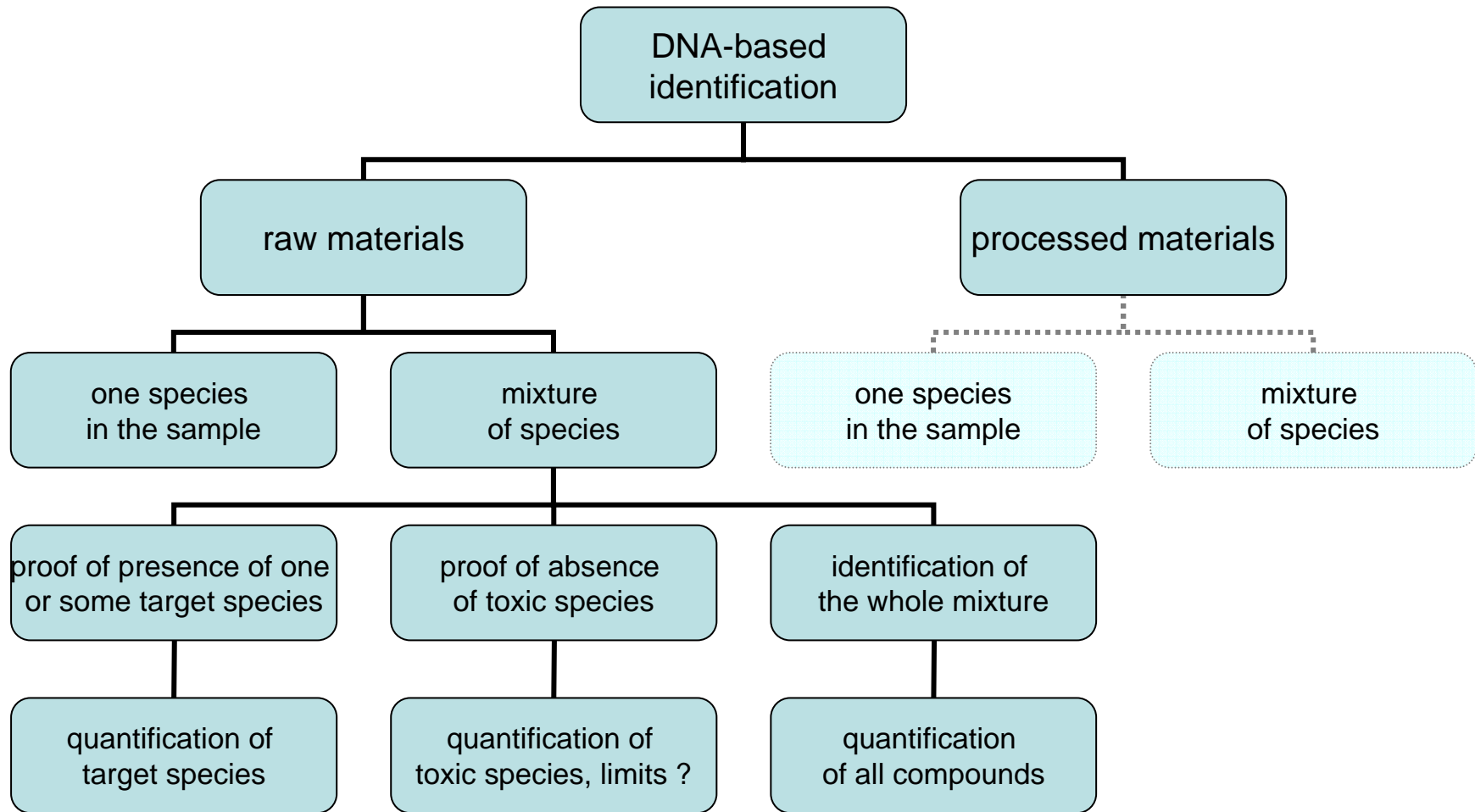
55th International Congress  
and Annual Meeting of the Society for  
Medicinal Plant Research

Graz, Austria  
September 2 – 6, 2007

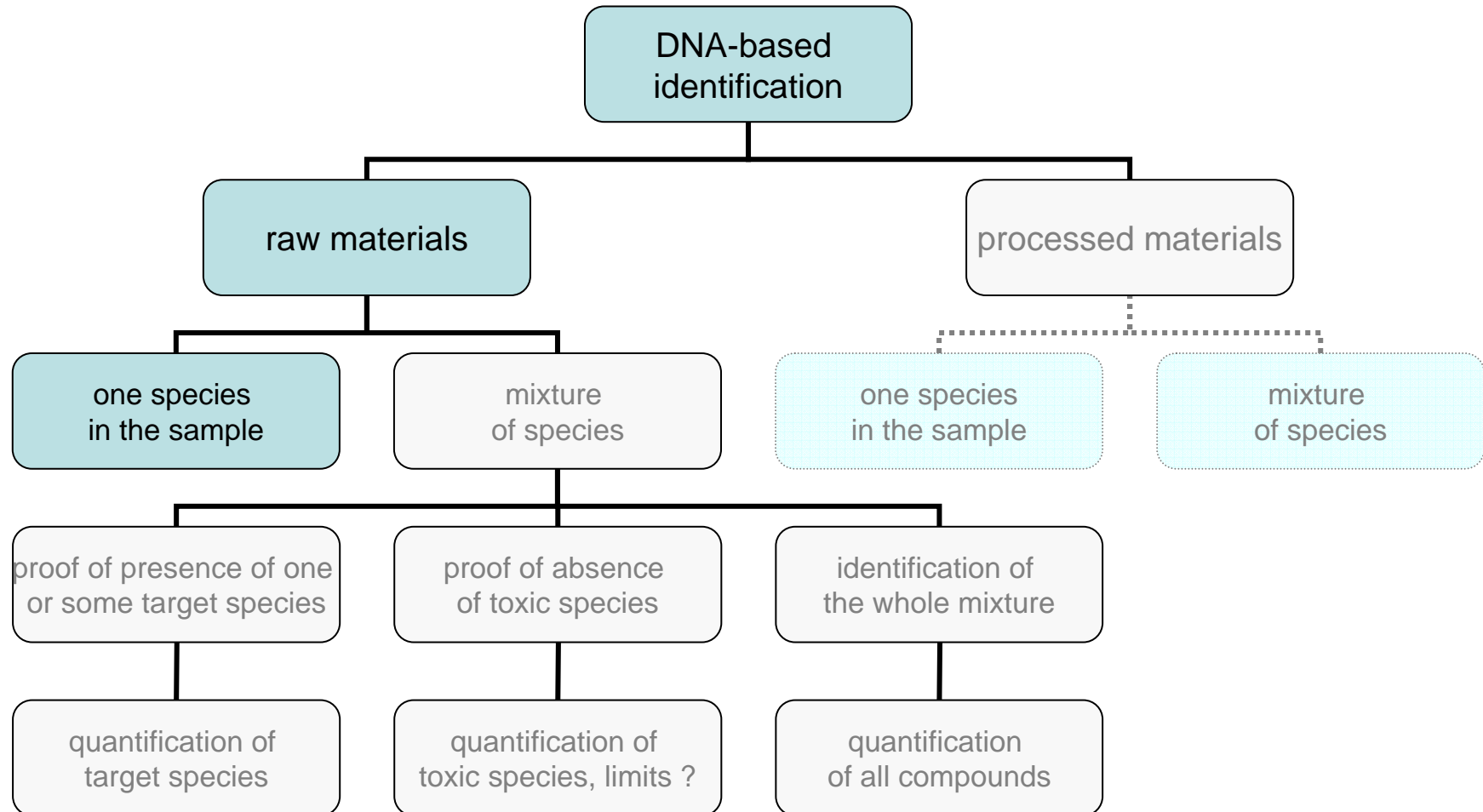


## Lectures and posters at GA2007 about DNA-based methods

- **PL 009 – Sucher NJ: Chips and Qi: microchip-based authentication of traditional Chinese medicinal plants.**
- P 248 – Phonkot et al.: Antioxidant activities and DNA fingerprint of 4 varieties Lotus stamens.
- P 256 – Kersten et al.: The potential of PCR-related methods to identify medicinal plants in herbal medicinal products.
- P 610 – Xue and Li: Molecular identification of the traditional Tibetan medicinal plant *Gentianopsis paludosa* (Gentianaceae) using diagnostic PCR and PCR-RFLP based on nrDNA ITS regions.
- P 619 – Xue et al.: Molecular authentication of the traditional Chinese medicinal plant *Euphorbia humifusa* and *E. maculata*.
- P 644 – Hadian et al.: Genetic diversity of Iranian accessions of *Satureja hortensis* L. using RAPD markers.
- P 646 – Heubl et al.: A molecular approach for the discrimination of medicinal *Astragali radix* used in TCM.

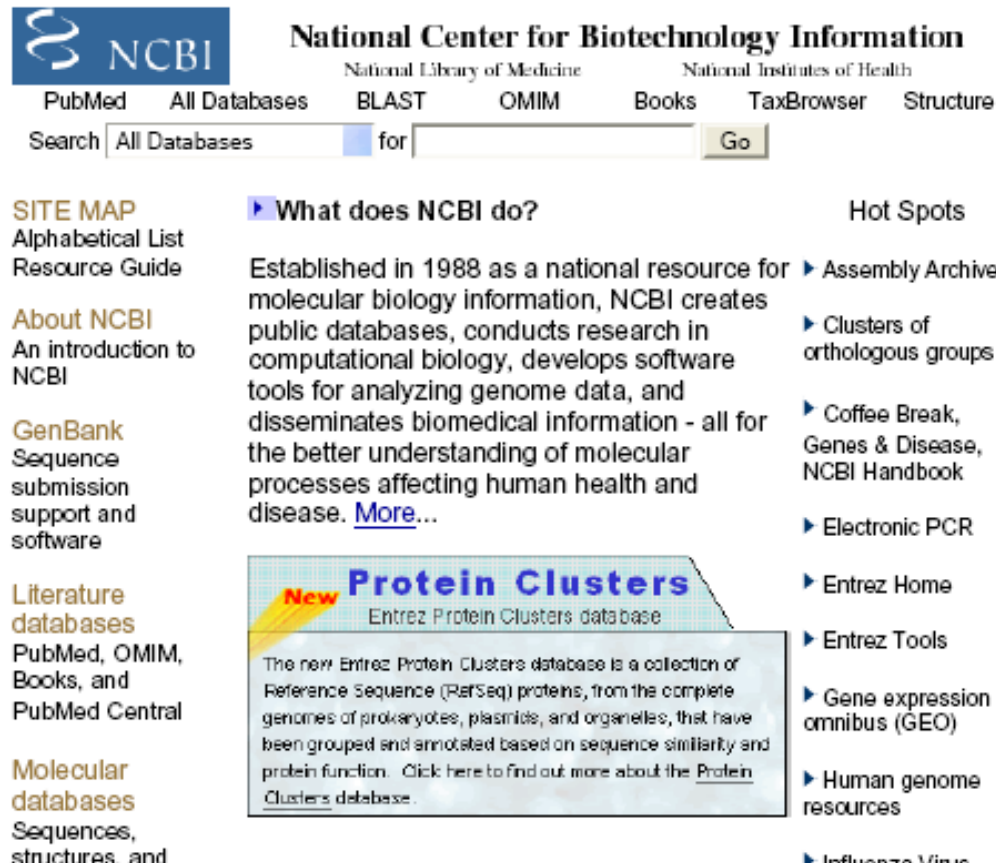


# I: Identifying Species (DNA-Barcoding?)



# Public Database for Sequences

<http://www.ncbi.nlm.nih.gov/>



The image shows the NCBI website homepage. At the top left is the NCBI logo. To its right is the text "National Center for Biotechnology Information" with "National Library of Medicine" and "National Institutes of Health" below it. A navigation bar contains links for PubMed, All Databases, BLAST, OMIM, Books, TaxBrowser, and Structure. Below this is a search bar with "All Databases" selected and a "Go" button. The main content area is divided into three columns. The left column lists "SITE MAP" (Alphabetical List, Resource Guide), "About NCBI" (An introduction to NCBI), "GenBank" (Sequence submission support and software), "Literature databases" (PubMed, OMIM, Books, and PubMed Central), and "Molecular databases" (Sequences, structures, and...). The middle column has a heading "What does NCBI do?" followed by a paragraph: "Established in 1988 as a national resource for molecular biology information, NCBI creates public databases, conducts research in computational biology, develops software tools for analyzing genome data, and disseminates biomedical information - all for the better understanding of molecular processes affecting human health and disease. [More...](#)". Below this is a "New Protein Clusters" banner for the Entrez Protein Clusters database, with a text box stating: "The new Entrez Protein Clusters database is a collection of Reference Sequence (RefSeq) proteins, from the complete genomes of prokaryotes, plasmids, and organelles, that have been grouped and annotated based on sequence similarity and protein function. [Click here to find out more about the Protein Clusters database.](#)". The right column is titled "Hot Spots" and lists: Assembly Archive, Clusters of orthologous groups, Coffee Break, Genes & Disease, NCBI Handbook, Electronic PCR, Entrez Home, Entrez Tools, Gene expression omnibus (GEO), Human genome resources, and Influenza Virus.

# What is „DNA-Barcoding“?

- A single gene region (approx 600 base pairs) is sequenced and used as „barcode“.
- The same sequence will be used for all species to standardise the protocol
- An inventory of all species will be created (without accompanying morphological analysis, etc.)
- The gene sequences will be analysed with distance algorithms to assign an individual to a species.

## Identification of Birds through DNA Barcodes

Paul D. N. Hebert<sup>1\*</sup>, Mark Y. Stoeckle<sup>2</sup>, Tyler S. Zemlak<sup>1</sup>, Charles M. Francis<sup>3</sup>

1 Department of Zoology, University of Guelph, Guelph, Ontario, Canada, 2 Program for the Human Environment, Rockefeller University, New York, New York, United States of America, 3 National Wildlife Research Centre, Canadian Wildlife Service, Ottawa, Ontario, Canada

PLoS Biol 2(10): e312 (2004).

# BOLD: The Barcode of Life Data System (<http://www.barcodinglife.org>)

RATNASINGHAM S. and HEBERT P.D.N., Molecular Ecology Notes, 7, 355-364, 2007

**BOLDSYSTEMS** | Management & Analysis

Bats of Southeast Asia - first paper [BM]

**Barcode Identifiers**

Barcode ID: BM258-04      Sample ID: ROM 101996  
 Gene: COX1      GenBank Accession: Record is unpublished  
 Last Updated: 2005-08-26      Translation Matrix: Vertebrate Mitochondrial

**Sequencing Runs**

Run Date	Run Site	Direction	Trace File	PCR primers	Seq Primer	Status
2004-11-18 15:18:37	University of Guelph	Reverse	BM258-04R2_H01.ab1	VR2VF2	VR2	high qual
2004-11-18 11:55:58	University of Guelph					

**Nucleotide Sequence**

Length: 619  
 Comp. A: 190  
 Comp. G: 95  
 Comp. C: 169  
 Comp. T: 165  
 Updated: 2005-08-26

GGACAACCGAGGCC  
 TTCTTCATAGTAATA  
 ATAGCATTCCCGGA  
 ACAGTAGAAGCTGG  
 TCTGTAGATCTAGCA  
 ACCATCATCAATATA  
 GCAGTCCTACTATTA  
 ACAACTTCTTCGAC

**Amino Acid Sequence**

Length: 206

MAFFRMINMSFWLLP  
 TIINMKPPALSQYQT

**Illustrative Barcode**

**Sequencing Run Details:**

Run Site: University of Guelph  
 Date: 2004-11-18      PCR primers: VR2VF2      Mean: 45.3249      C300  
 File: 04R2\_H01.ab1      Seq Primer: VR2      View: 58683.9  
 Status: high qual      Direction: Reverse      Slides: 244.6774      1382 2436 18 20 22 24 26 28 30 32 34

**Sequencing Run Details:**

Run Site: University of Guelph  
 Date: 2004-11-18      PCR primers: VR2VF2      Mean: 44.1869      C300  
 File: 04F2\_H01.ab1      Seq Primer: VF2      View: 79163.9  
 Status: high qual      Direction: Forward      Slides: 279.579      1382 2436 18 20 22 24 26 28 30 32 34

**BOLDSYSTEMS** | Management & Analysis

Specimen Identification Request

**Identification Summary:**

Taxonomic Level	Taxon Assignment	Probability of Placement (%)
Phylum	Arthropoda	100
Class	Insecta	100
Order	Lepidoptera	100
Family	Arctiidae	100
Genus	Neonerita	100
Species	Neonerita obscurata	100

**Distance Summary:**

Similarity scores of the top 100 matches

**TOP 10 Matches:**

Phylum	Class	Order	Family	Genus	Species	Specimen Similarity (%)
Arthropoda	Insecta	Lepidoptera	Arctiidae	Neonerita	obscurata	100
Arthropoda	Insecta	Lepidoptera	Arctiidae	Neonerita	obscurata	99.93
Arthropoda	Insecta	Lepidoptera	Arctiidae	Neonerita	obscurata	99.83
			Arctiidae	Neonerita	obscurata	99.83
			Arctiidae	Neonerita	obscurata	99.83
			Arctiidae	Neonerita	obscurata	99.83
			Arctiidae	Neonerita	obscurata	99.83
			Arctiidae	Neonerita	obscurata	99.83
			Arctiidae	Neonerita	obscurata	99.83
			Arctiidae	Neonerita	obscurata	99.83

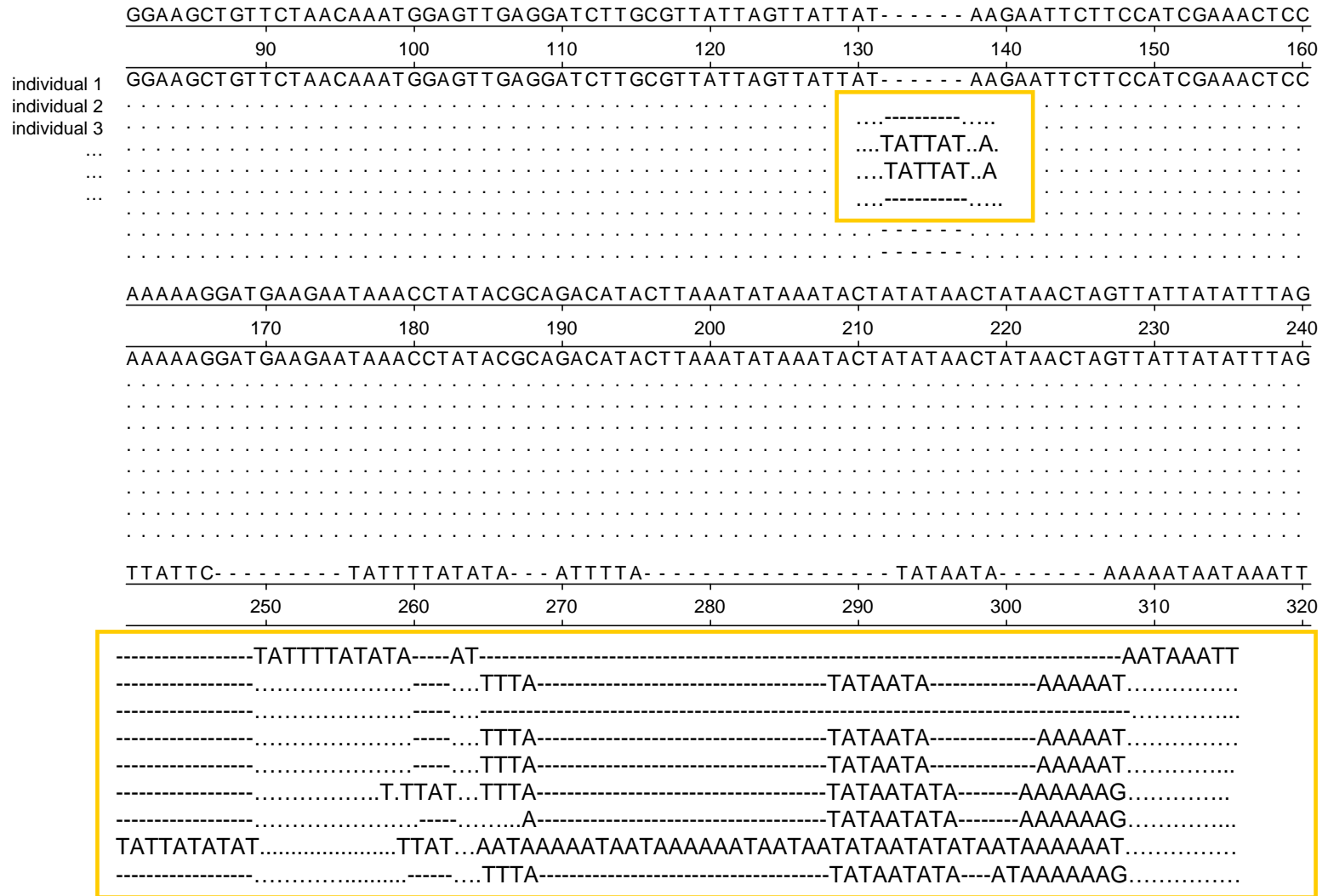
**Species Page:**

Phylogenetic tree showing relationships between species. Includes a pie chart and a map of the world.

**Species Images:**

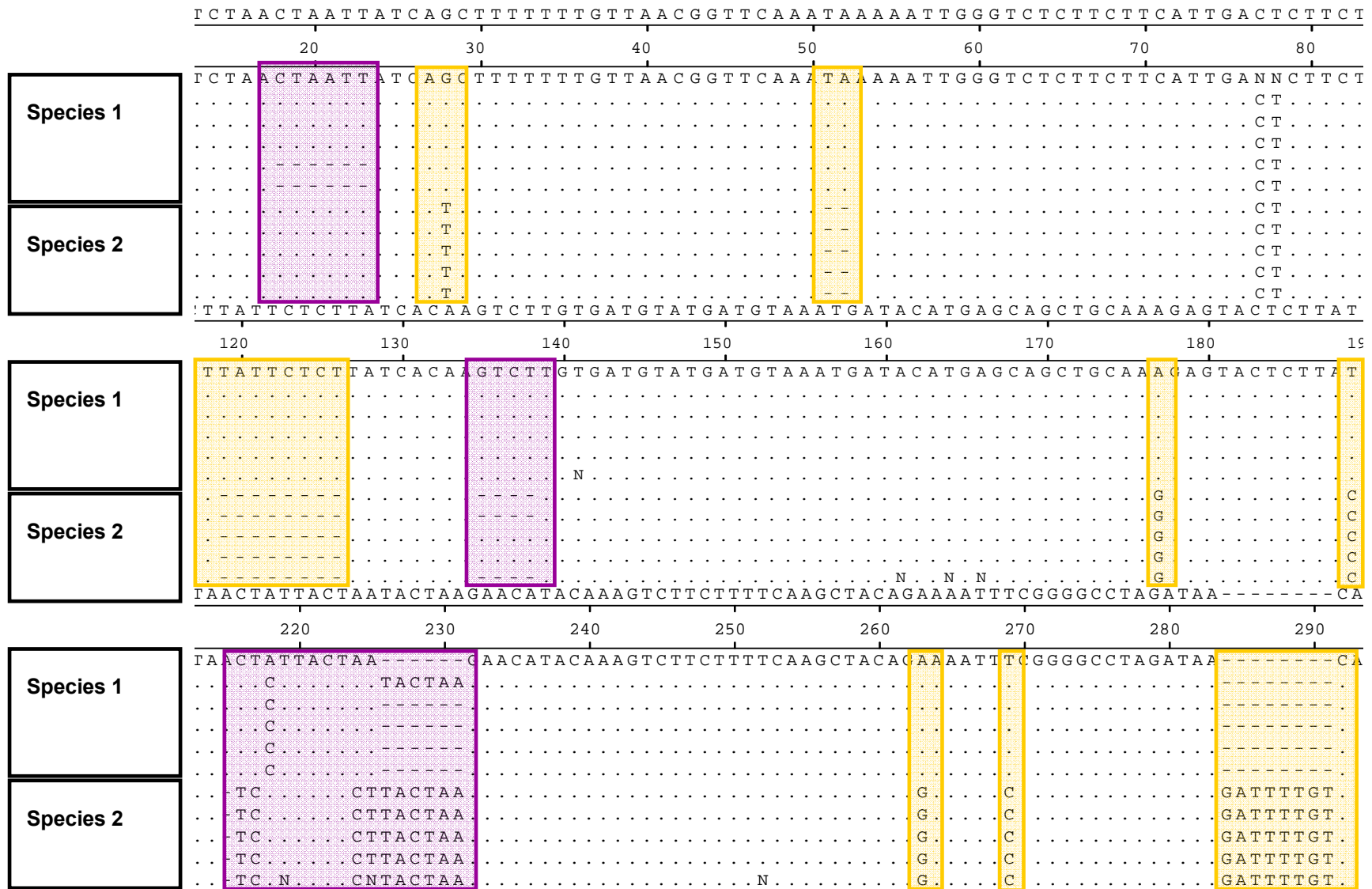
# Caveat: Intraspecific Variability

example: several individuals of one species



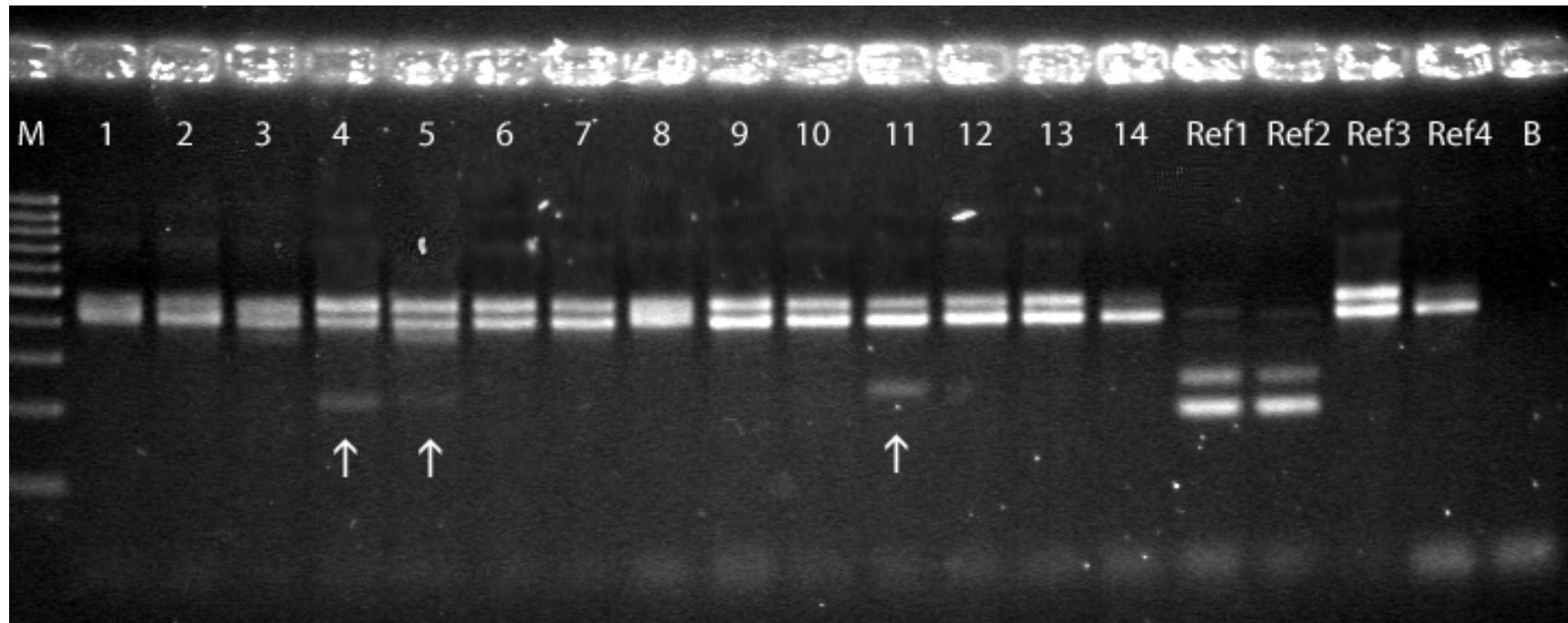


# Intraspecific Variability



yellow boxes: intraspecific (usable) variability, violet boxes: interspecific variability

# Subgenus-specific PCR Based on Sequence Information

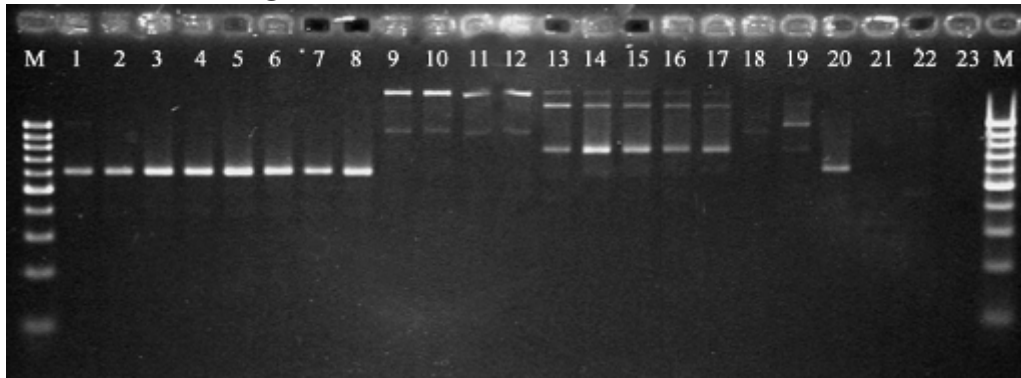


lanes 1-8: samples (drug)

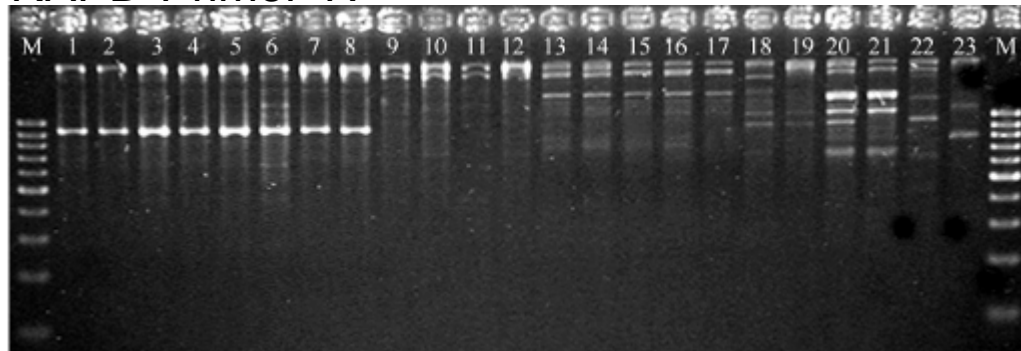
No.	Species	Subgenus
Ref1	<i>Species 1</i>	Subgenus 1
Ref2	<i>Species 2</i>	Subgenus 1
Ref3	<i>Species 3</i>	Subgenus 2
Ref4	<i>Species 4</i>	Subgenus 2

# Species-specific ,Anonymous' Molecular Markers (not based on sequence information)

RAPD-Primer X:



RAPD-Primer Y:

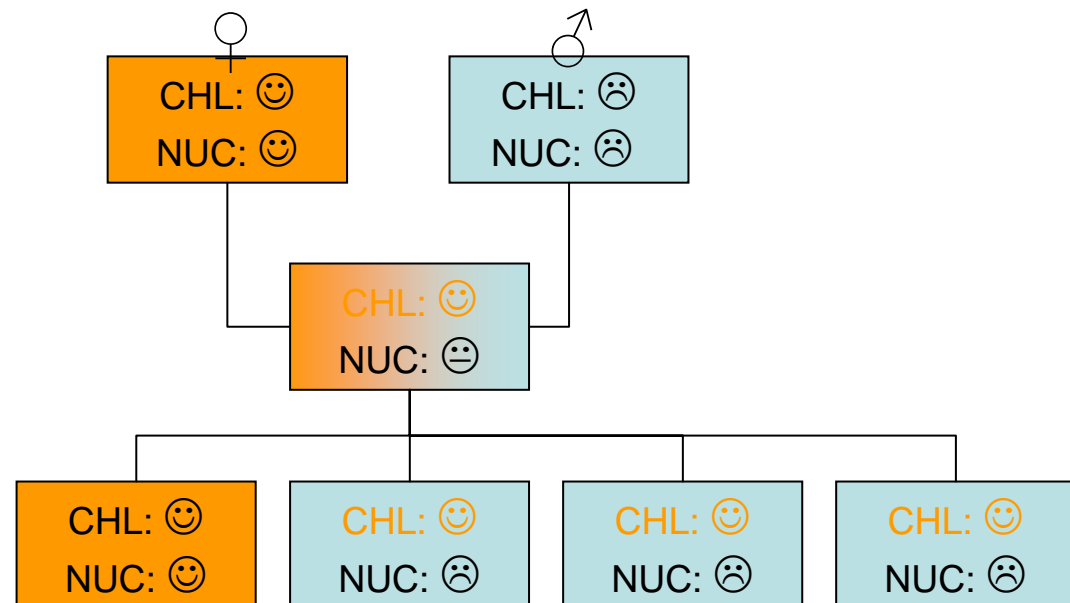


lanes 1-8: target species

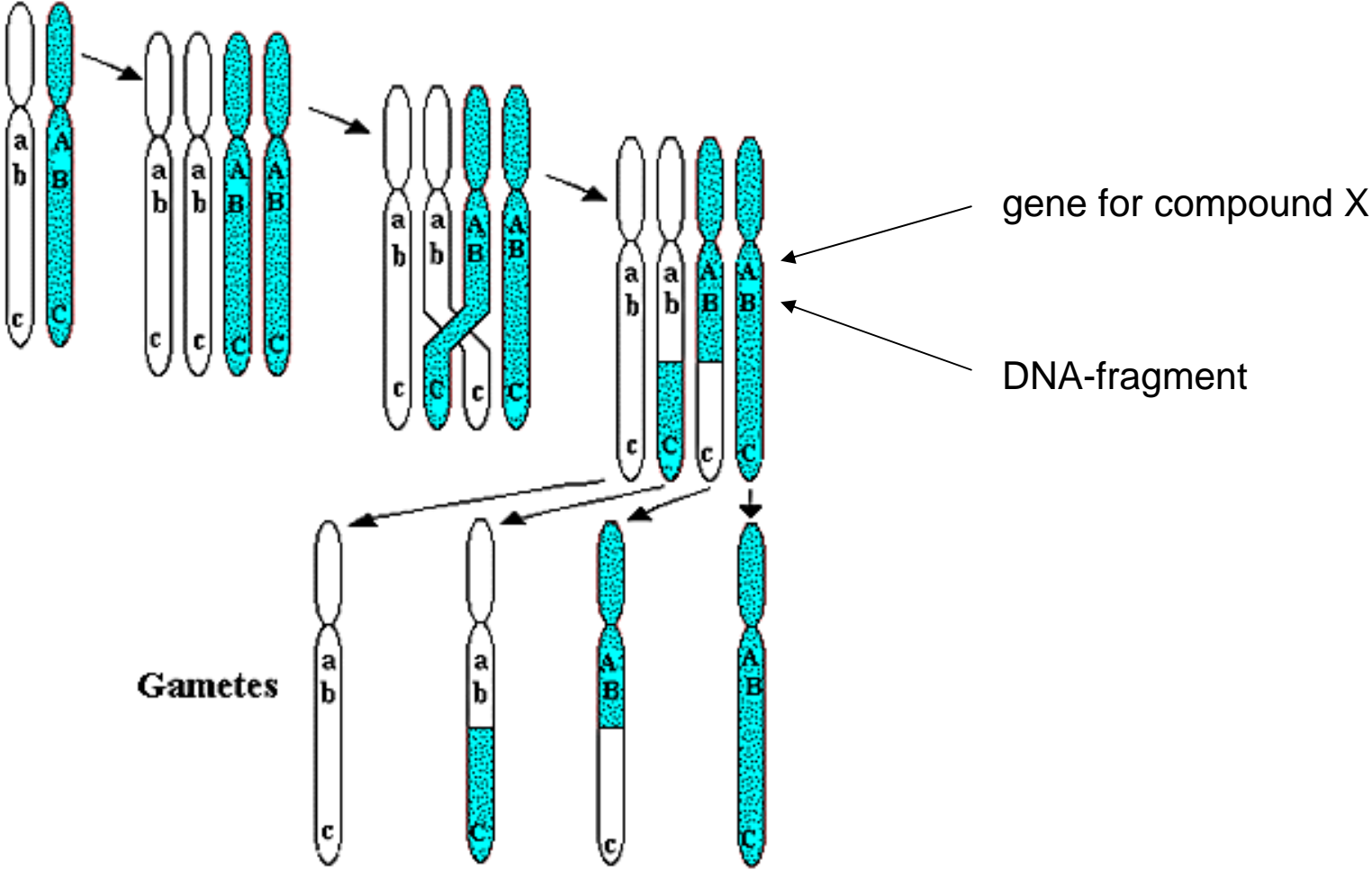
# Problem Hybridisations

(especially Introgression - Chloroplast vs. Nuclear Genome)

- Taxonomy is often working with chloroplast sequences
- These sequences are only of limited use in detecting hybridisations because chloroplasts are maternally inherited
- Nuclear sequence information necessary detecting recent hybridisations



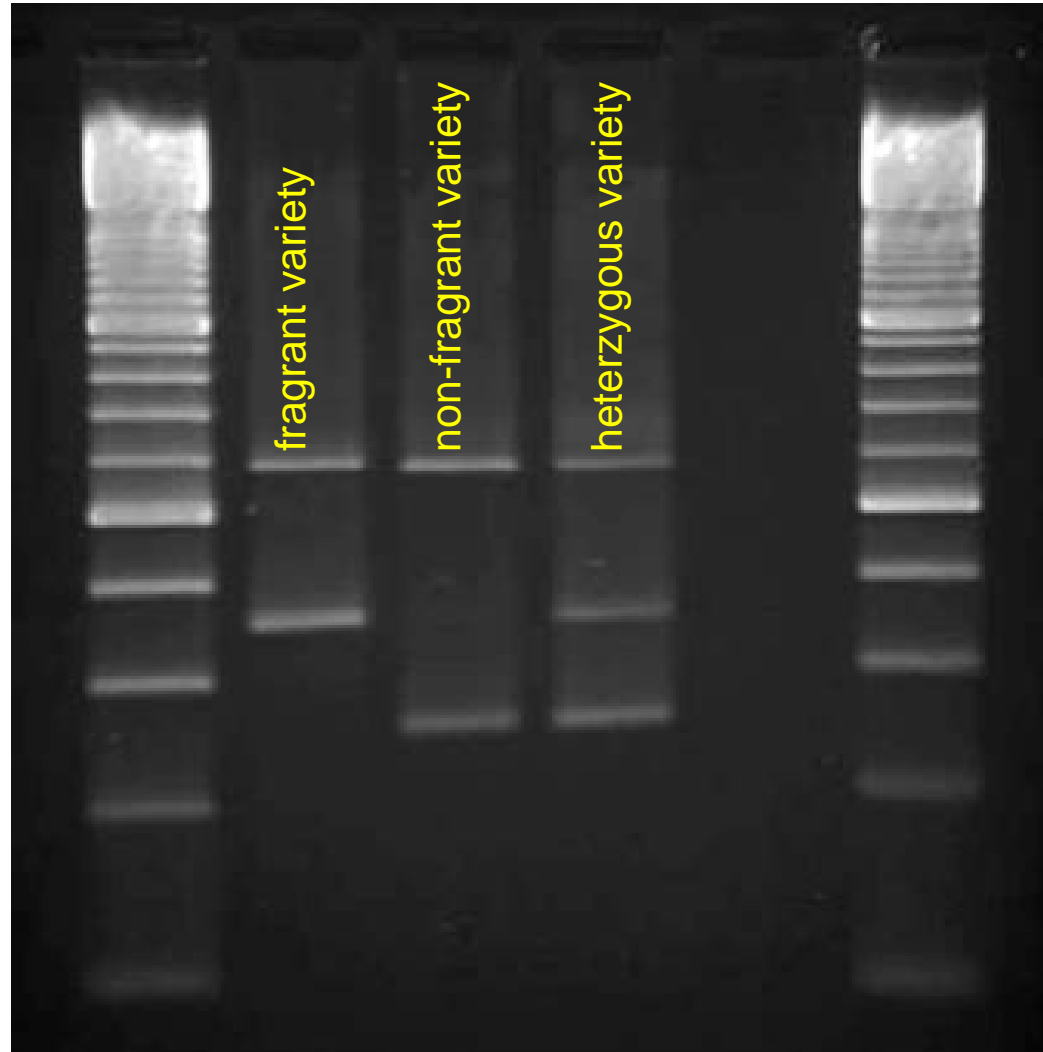
# DNA-based Identification of Chemotypes (‘Molecular Markers’)



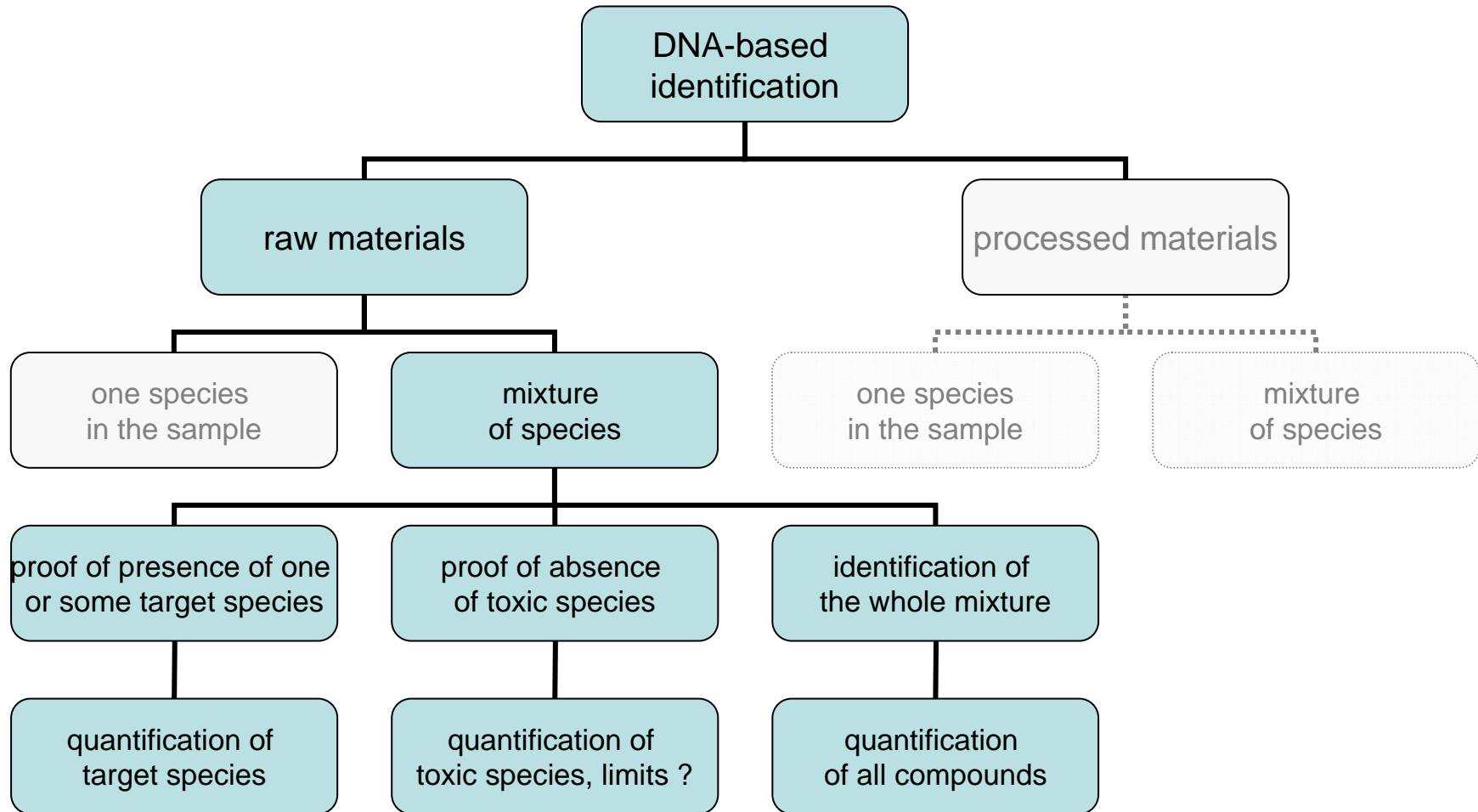
Crossing-over and recombination during meiosis

**Example for Molecular Markers**

**Bradbury LMT et al.: A perfect marker for fragrance genotyping in rice. Molecular Breeding (2005) 16: 279–283**

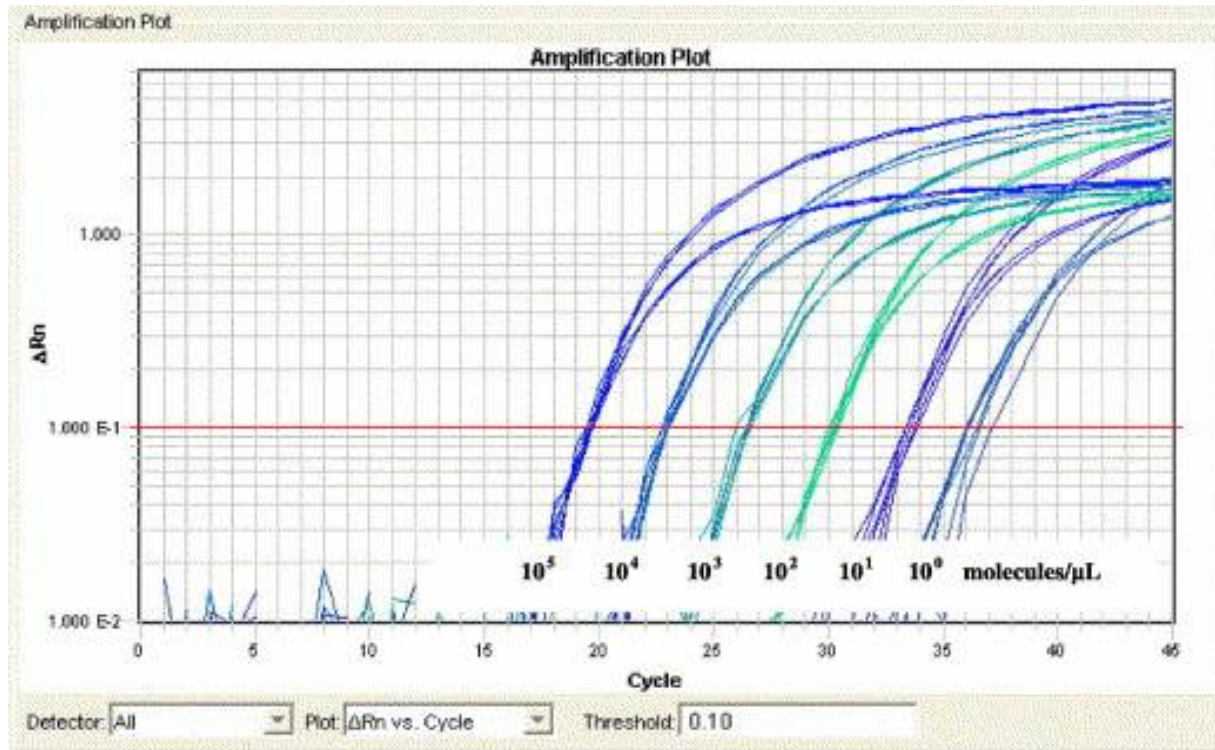


## II: Mixtures



## qPCR (quantitative PCR)

Engel et al.: Quantification of DNA from genetically modified organisms in composite and processed foods. Trends in Food Science & Technology 17, 490-497, 2006.



← GMO-specific signals  
(FAM-labeled probes)

← Rapeseed-specific signals  
(VIC-labeled probes)

### Calibration curves:

LibertyLink assay

$$\text{Log Qty (recDNA)} = -3.414 * C_t + 38.914$$

$$\text{Log Qty (reference)} = -3.432 * C_t + 39.335$$

SeedLink assay

$$\text{Log Qty (recDNA)} = -3.492 * C_t + 39.387$$

$$\text{Log Qty (reference)} = -3.353 * C_t + 38.874$$



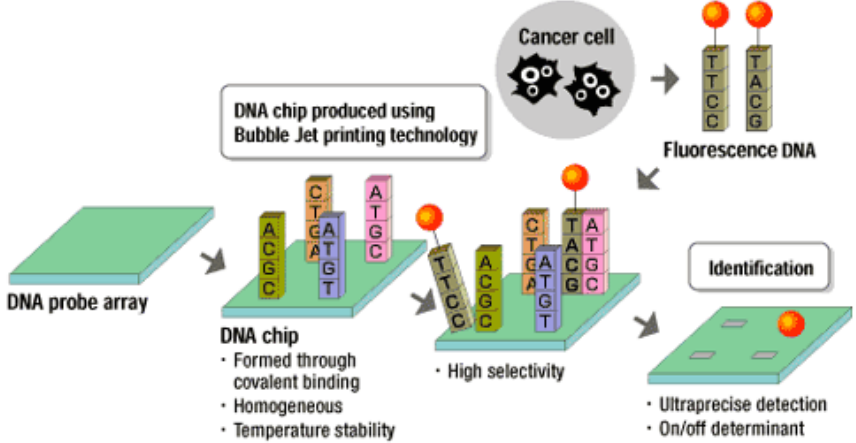
# DNA-Microarrays

DNA-Microarrays

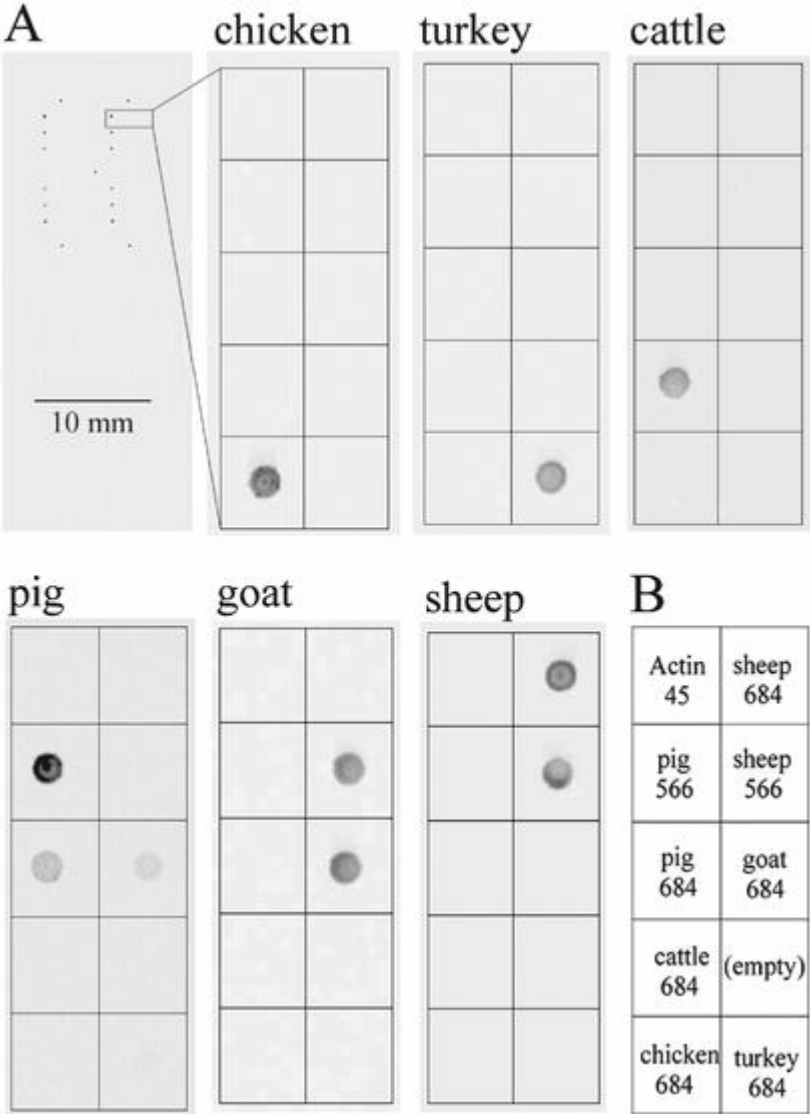
Pharmacodynamics

Pharmacogenomics

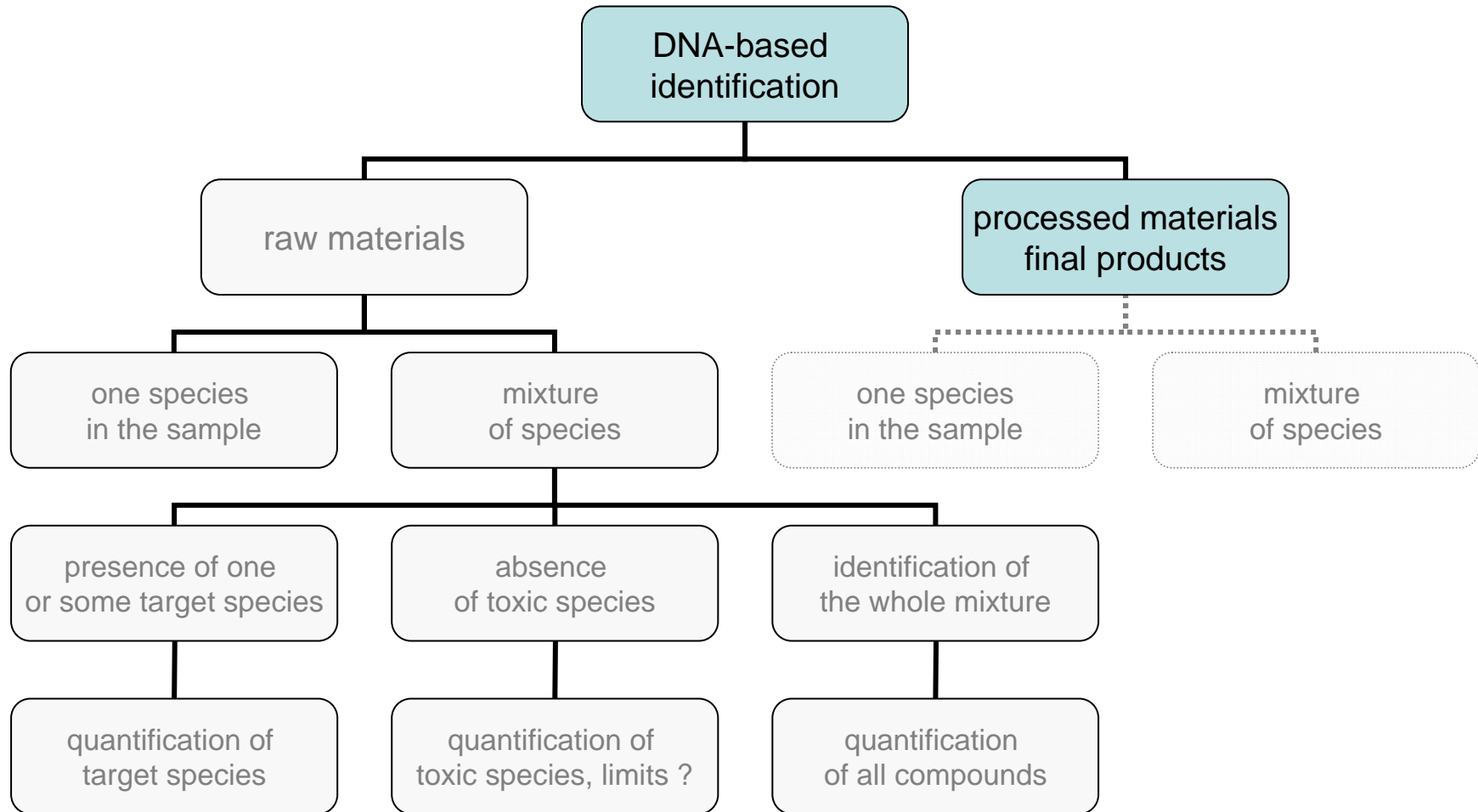
Pharmacognosy



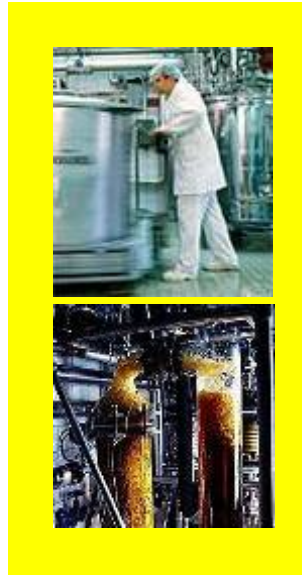
**Example for Species Detection by DNA Microarray:  
 Peter et al., Eur Food Res Technol 219:286-293 (2004)**



### III. Processed Materials / Final Products



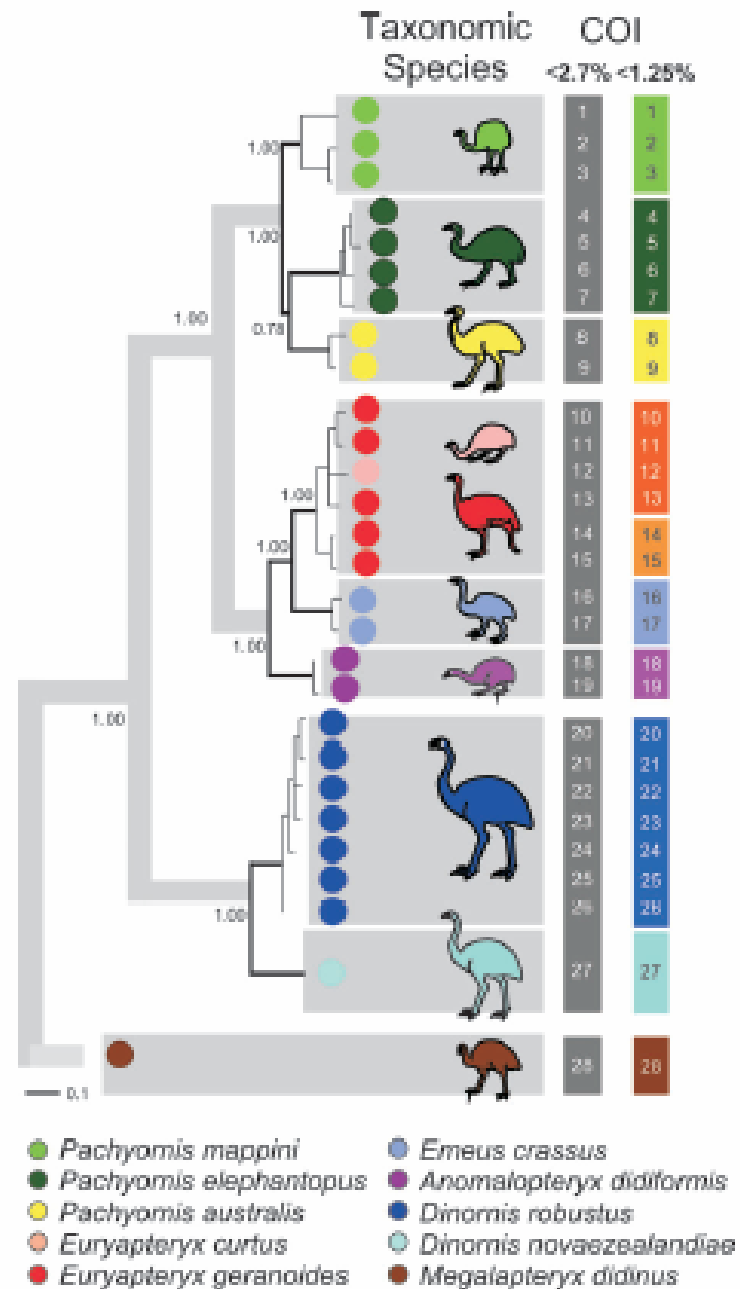
### III. Processed Materials / Final Products



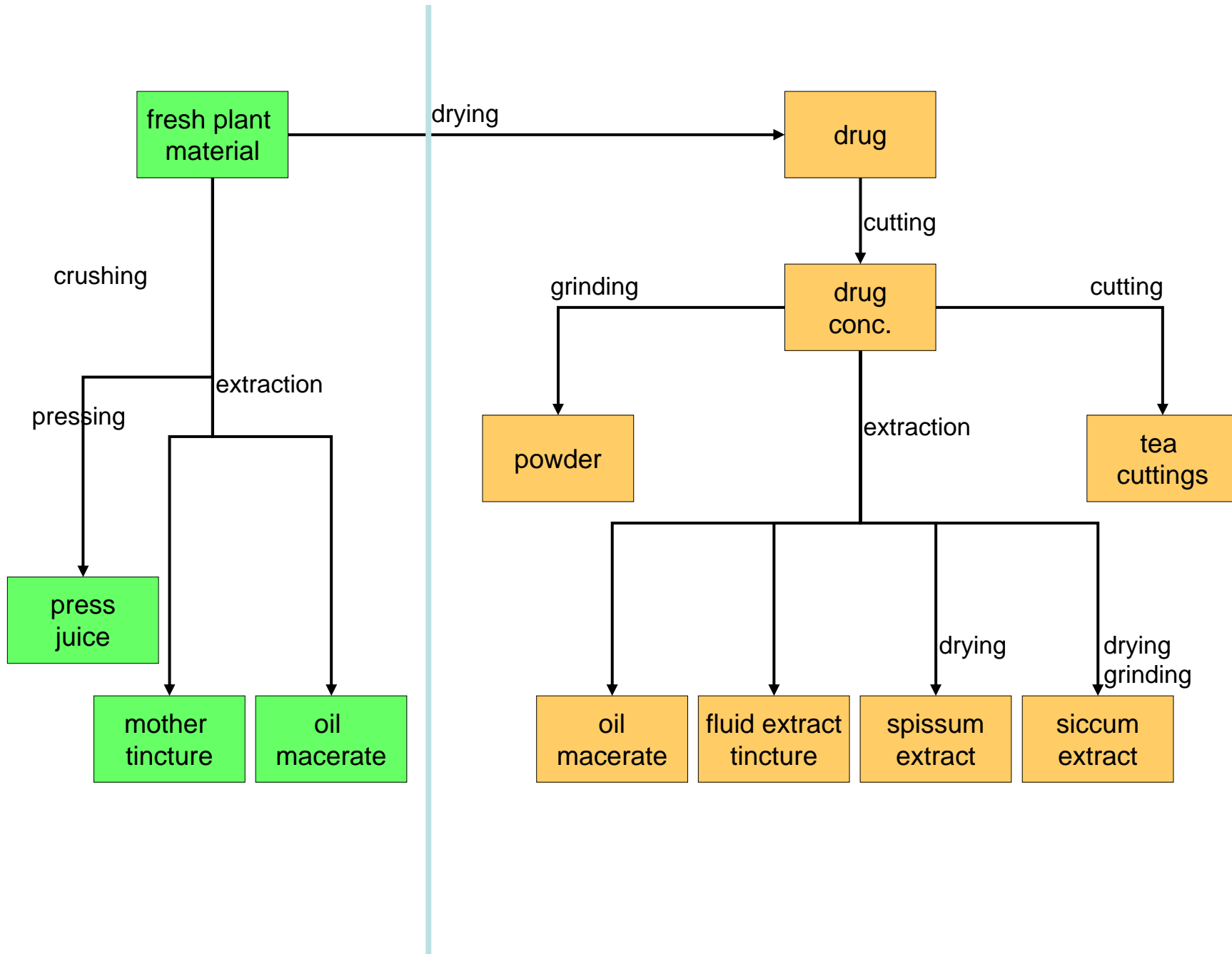
# Stability of DNA Extinct, Subfossilised Species

## Is a Large-Scale DNA-Based Inventory of Ancient Life Possible?

D. M. LAMBERT, A. BAKER, L. HUYNEN, O. HADDRATH, P. D. N. HEBERT, AND C. D. MILLAR

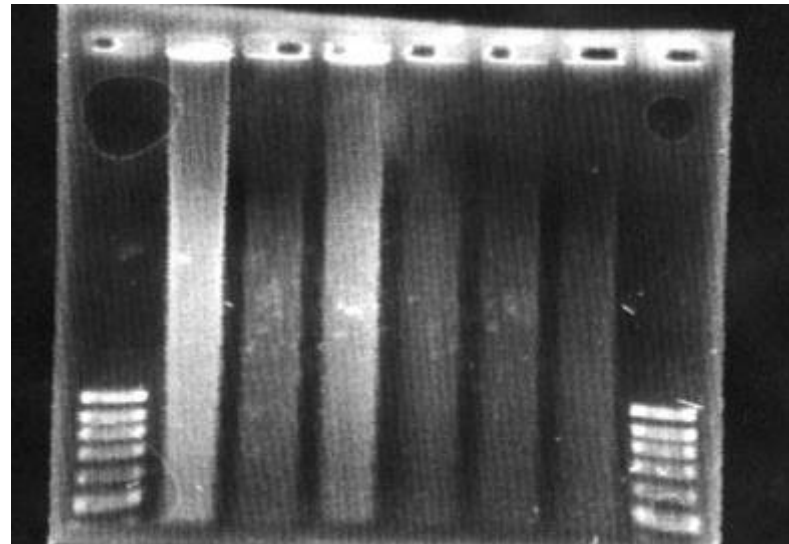


Journal of Heredity 96:279–284 (2005)



## Processed Materials: Powder

- DNA already strongly degraded, but DNA-based identification still possible



# Extracts

## *Echinacea* sp. - spissum ethanol extract

Novak et al.: DNA-Based Authentication of Plant Extracts. Food Research International 40: 388-392, 2007.

	130	140	150	160	170	180
plant extract	TTAAAGGGCTTGTGCTGTTATGCCCGTCA-CTGGTGTGCATAGT-GTGC GTTGCTTCTTT					
<i>Echinacea purpurea</i>	.....-.....-.....					
<i>Echinacea paradoxa</i>	.....C.....-					
<i>Echinacea pallida</i>	.....-.....-					
<i>Echinacea atrorubens</i>	.....A.....C.....-					
<i>Echinacea simulata</i>	.....A.....C.....-					
<i>Echinacea tennesseensis</i>	.....-.....C.....-					
<i>Rudbeckia fulgida</i> var. <i>fulgida</i>	A	G	T	CC	T	.....NN
<i>Rudbeckia subtomentosa</i>	..	G	T	CC	.....A	..C
<i>Sanvitalia fruticosa</i>	.....	CC	.....	A	.....	-
<i>Rudbeckia missouriensis</i>	A	G	T	CC	T	.....T
<i>Rudbeckia heliopsidis</i>	..	G	T	CA	-	.....C
<i>Oblivia mikanioides</i>	.....	G	-	.....	T	G
<i>Idiopappus quitensis</i>	.....	A	.....	A	G	C

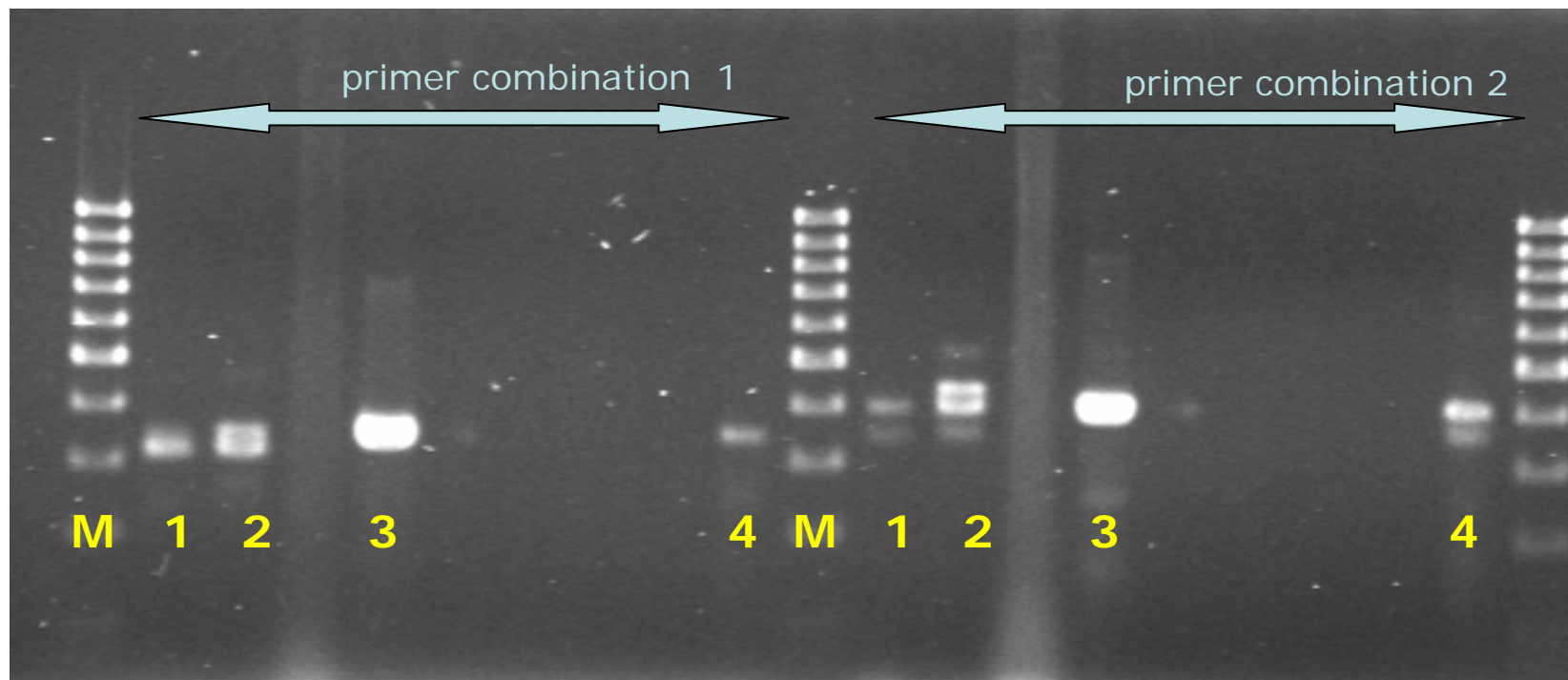


# Extracts

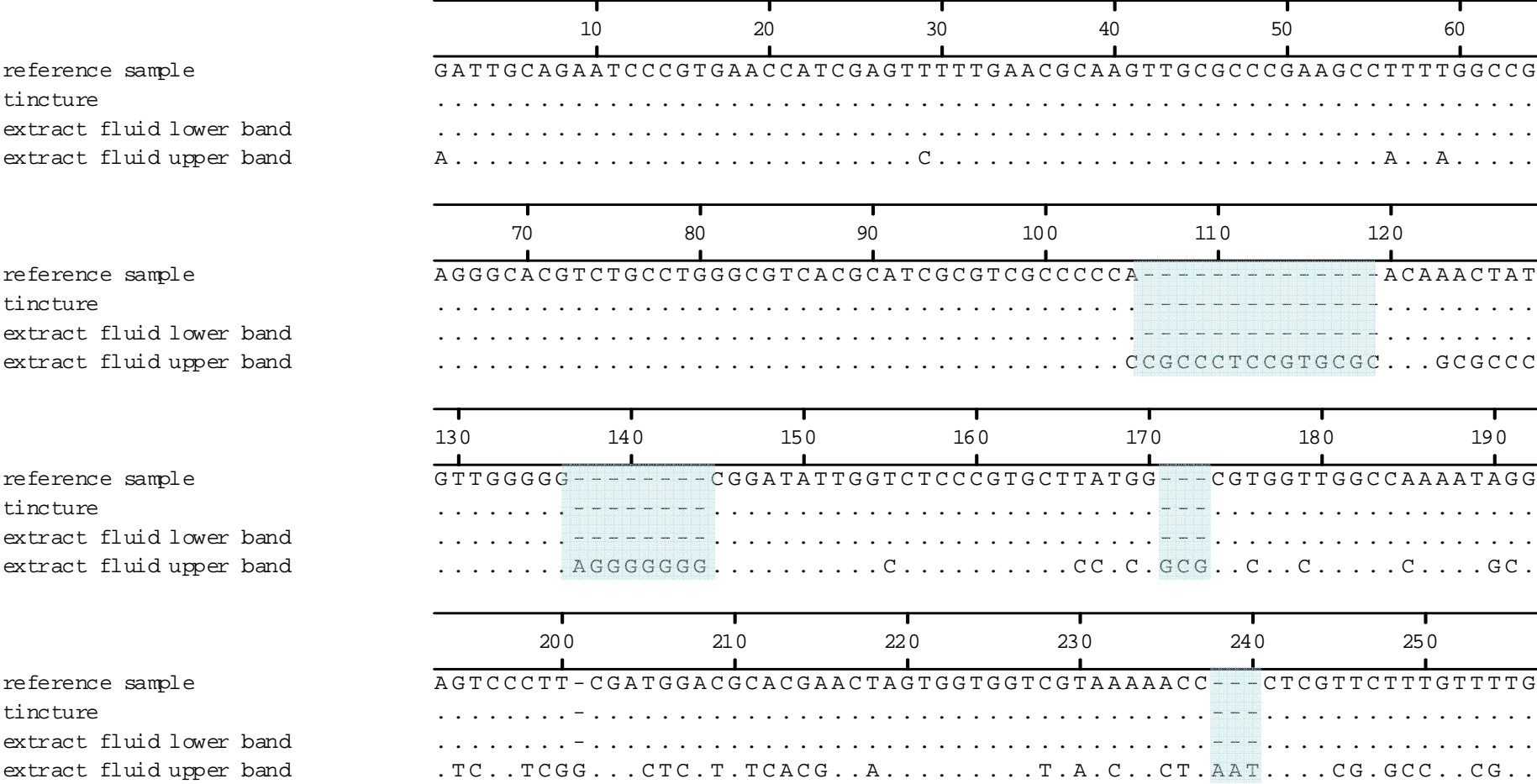
## *Chamomilla* tincture / extract fluid / tea bag

Novak et al.: DNA-Based Authentication of Plant Extracts. Food Research International 40: 388-392, 2007.

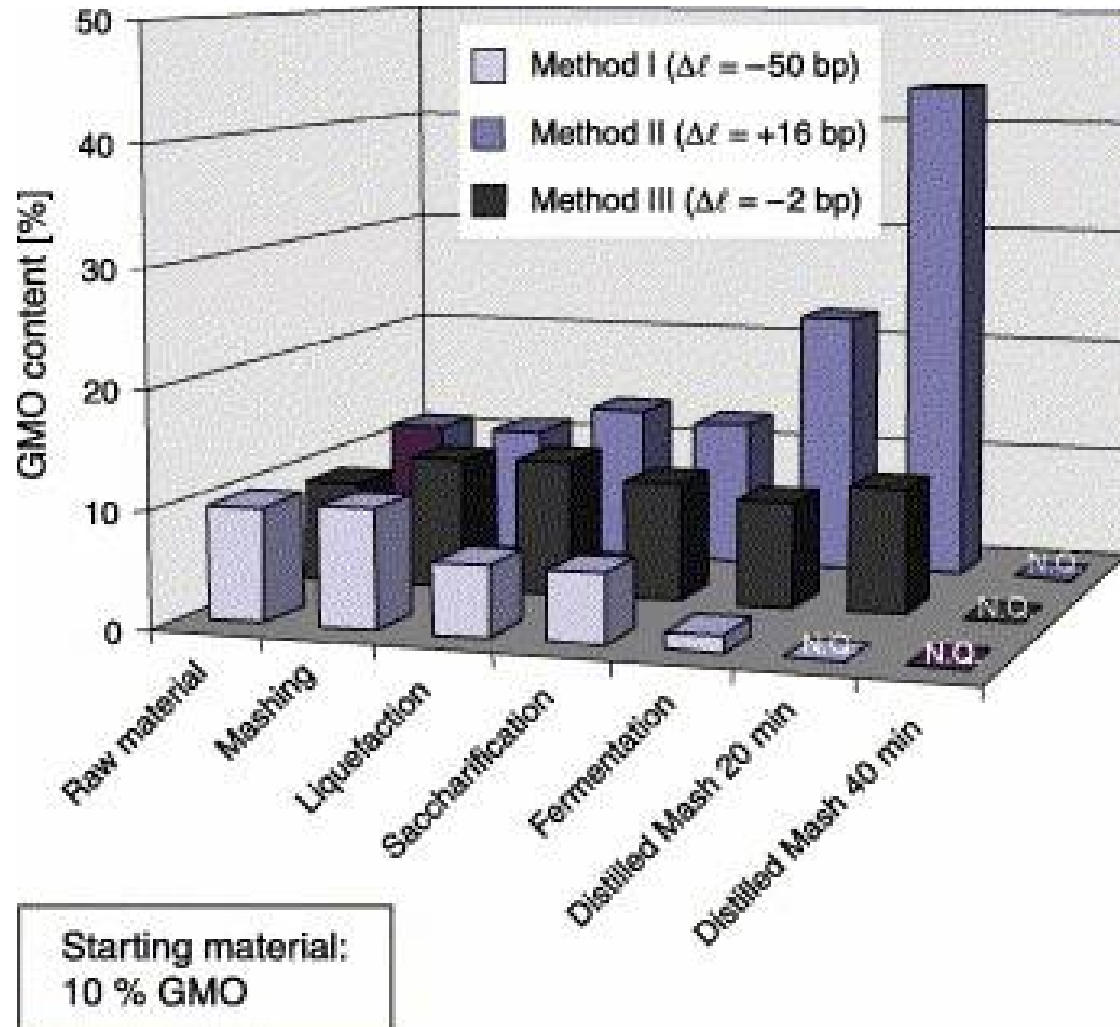
- 1: Chamomilla tincture
- 2: Chamomilla extract fluid
- 3: Chamomilla reference plant
- 4: Chamomilla teabag



# Chamomilla: primer combination 2



Relative quantification of DNA from corn Bt-176 in samples from intermediate stages of the process of ethanol production; NQ, not quantifiable. (Moreano, 2005 and Moreano et al., 2005b).



We kindly appreciate Bionoricas interest in our work

