Reaction-Based Optical Fingerprinting Strategy in the Recognition of Proteins, Motor Oils, and Estimation of Food Irradiation Doses

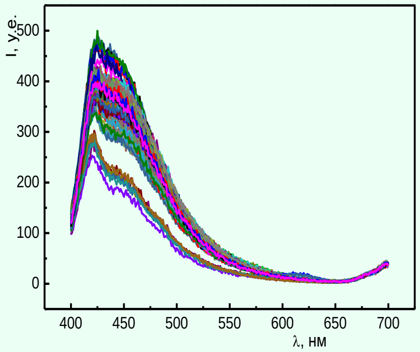
Anna V. Shik, Arseny A. Pypin, Yana V. Zubritskaya,[#] Pavel V. Sobolev, Maksim O. Baytler, Irina A. Stepanova, Irina A. Doroshenko, Ulyana A. Bliznyuk,[#] Tatyana A. Podrugina, Mikhail K. Beklemishev

Lomonosov Moscow State University, Department of Chemistry
[#]Lomonosov Moscow State University, Physics Department
119991 Moscow, Russia
beklem@inbox.ru

Fingerprinting methods
(smart sensing, stochastic sensing, pattern-based sensing, pattern recognition, array sensing)

Basics

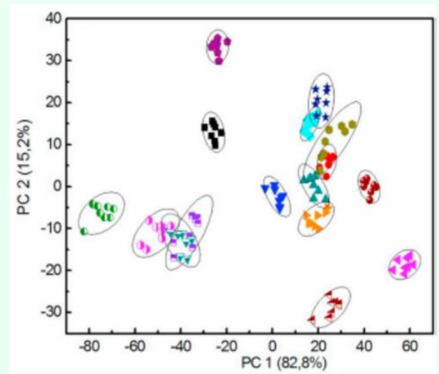
1) Obtain spectra, chromatograms, voltammograms of samples ("multidimensional data")

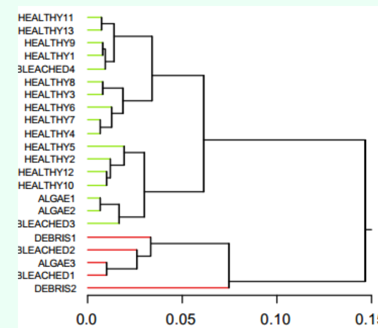


Sensor array

Convolu-
tion of data

2) Process them with chemometric methods (without assigning individual signals)

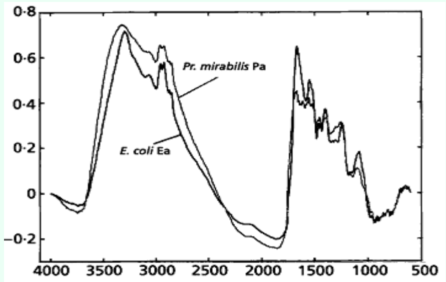




Problems solved:
1. Unsupervised methods – **clusterization** (grouping)
2. **Authentication** (does the sample belong to the class?)
3. Supervised methods – **discrimination** (to what class does it belong?)
4. Using a library – **identification**.
- Sometimes: **quantitation**

Specific features
- The nature and concentration of specific compounds is not determined (exception: combination fingerprint methods, usually chromatographic)
- No sample preparation (in many cases)

Acquisition of multidimensional data

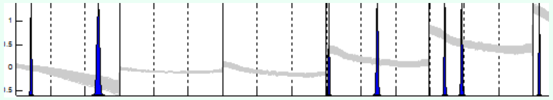


Spectroscopy:
- **UV-visible**
- **fluorescent**
- IR
- Raman
- optical multisensor systems
- NMR

Chromatography:
- HPLC-MS, GC-MS…
- Capillary electrophoresis

Electrochemical methods:
- Voltammetry
- Potentiometry

including multisensor systems
("electronic nose", "electronic tongue")



Tasks solved using fingerprinting methods

- Classification of samples
- Counterfeit detection
- Authenticity assessment
- Detection of additives, impurities
- Identification of manufacturer
- Identification of the source of pollution
- Recognition of stereoisomers
- Medical diagnostics
- Quantitative analysis

Sample types

- Foods
- Beverages
- Food supplements
- Oils and fats, petroleum products
- Soils
- Pharmaceuticals
- Microbiological and medical samples

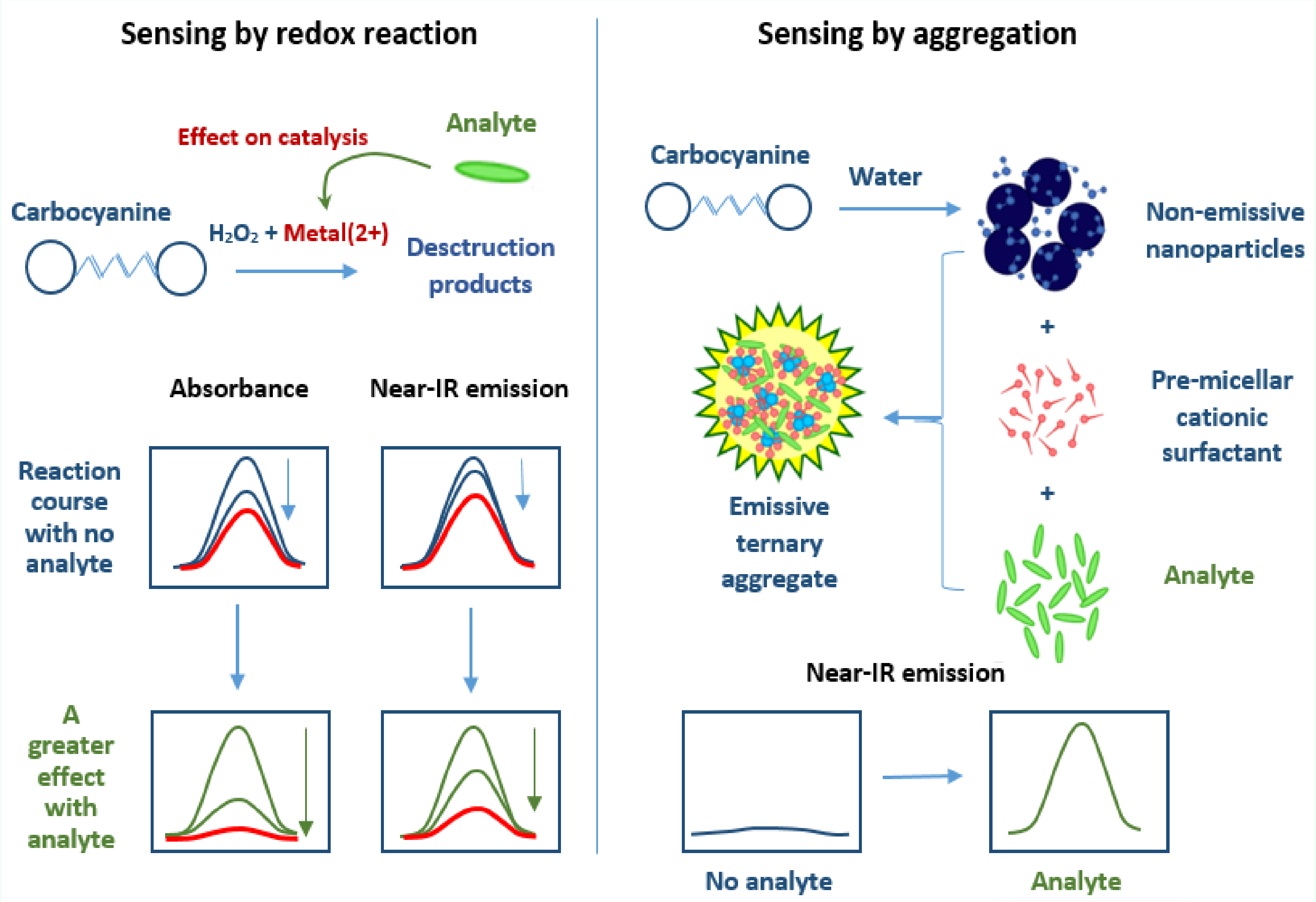
Sample preparation

- without sample preparation
- extraction (soils, plants, pharmaceuticals, blood plasma, food)

Fluorimetric fingerprinting: stages of development

Gene-ration	Based on	References
I	Intrinsic emission spectra of samples	Long known
II	Sample + fluorophore	U. Bunz, 2007 and on; our studies since 2017
III	Sample + indicator reaction	Our studies since 2021; colorimetry: Pargari *J. Anal. Chem.* **77** (2022) 482, Liu L. *Food Anal. Meth.* **14** (2021) 1852, Wang F. *Chem. Commun.* **57** (2021) 4520, Wang L. *Anal. Chim. Acta* **1121** (2020) 26

Signal formation

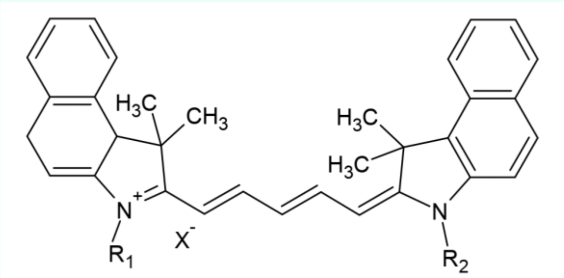


Mechanism 1. The analyte changes the rate of the dye oxidation reaction → fluorescence fading, color change.

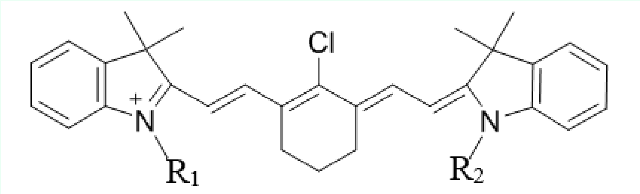
Mechanism 2. The analyte forms aggregates with an oppositely charged surfactant, the dye is **solubilized** in the hydrophobic domains of the aggregate → fluorescence enhancement.

Carbocyanine fluorophores

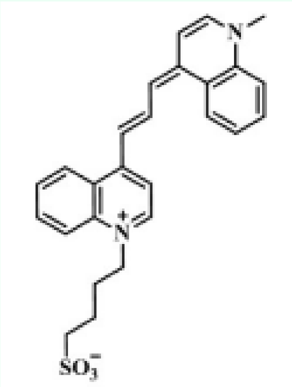
(synthesis: T.A. Podrugina, I.A.Doroshenko, Division of Medical Chemistry)



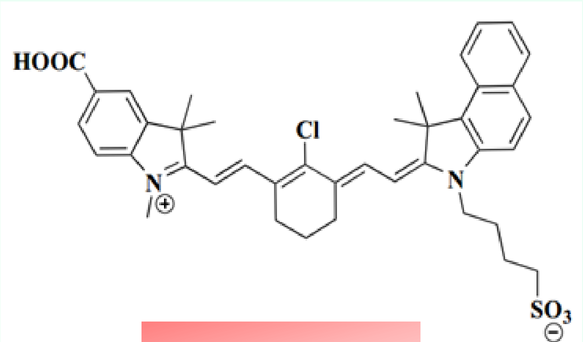
Dye 1: R_1, R_2 = $-(CH_2)_{10}COOH$;
X = Br^-



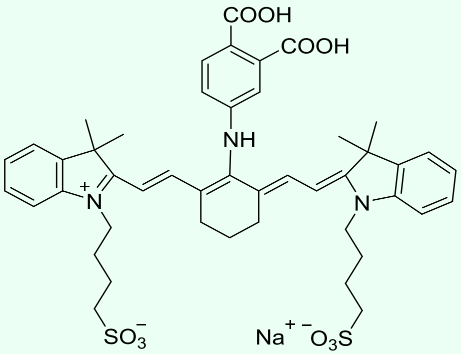
Dye 4: R_1, R_2 = $-(CH_2)_4SO_3^-$
Dye 5: R_1, R_2 = $-(CH_2)_3COOCH_2CH_3$



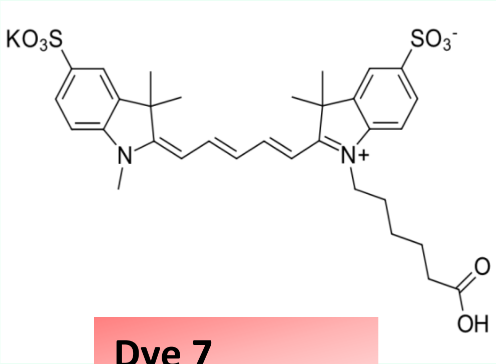
Dye 2



Dye 3



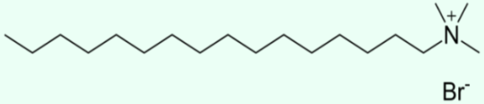
Dye 6



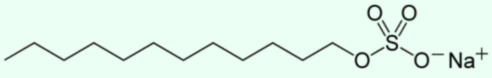
Dye 7

Counter-ions for aggregation reactions

Cetyltrimethyl-ammonium bromide (CTAB)

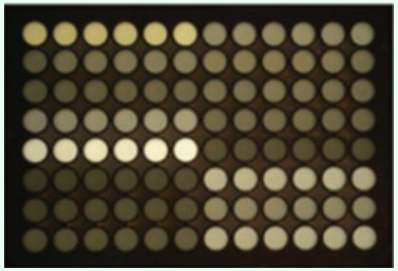


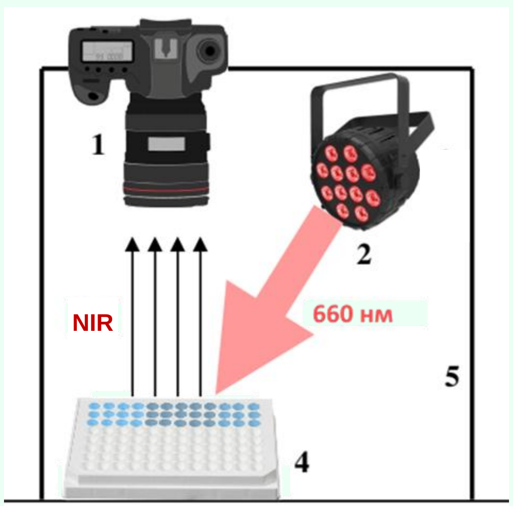
Sodium dodecyl sulfate (SDS)



Signals and instruments

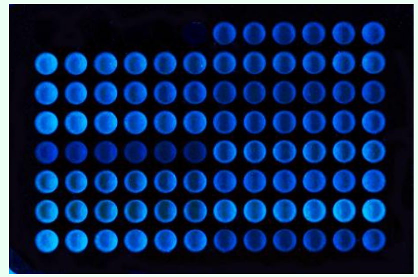
1. Near-IR fluorescence (ex 660 nm, em 700-800 nm)





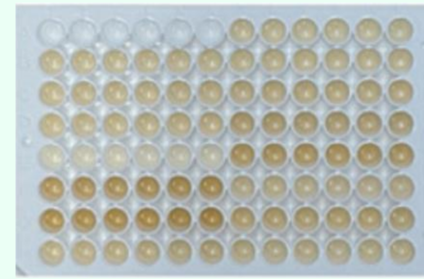
NIR visualizer:
1 – camera with light filter (>**700** nm),
2 – red LEDs (**660** nm),
4 – 96-well plate,
5 – housing.

2. Visible fluorescence (ex 254/365 nm, emission visible)



Camag visualizer:
ex **254** and **366** нм,
measurement – photo camera

3. Absorption/reflectance in visible light



Smartphone camera
or Camag visualizer

Data processing

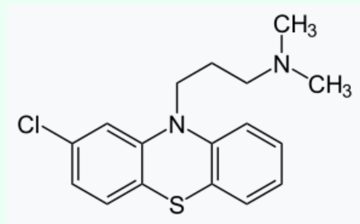
Standard approaches:

- Digitizing of images (ImageJ): fluorescence and absorbance intensities.
- RGB splitting
- Data table: *samples* (rows) – *reaction times* (columns)
- Principal component analysis (PCA), linear discriminant analysis (LDA), k-nearest neighbors algorithm (kNN).
- Score plots (PCA, LDA).
- Validation (83% training set / 17% validation set)
- Accuracy = $\frac{\text{no. of correctly assigned validation observations}}{\text{total number of observations in validation set}} \times 100\%$

1. Recognition of 9 pharmaceuticals

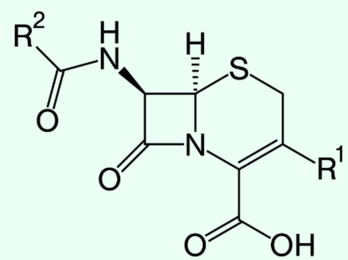
Phenothiazines

- Promazine
- Chlorpromazine
- Promethazine



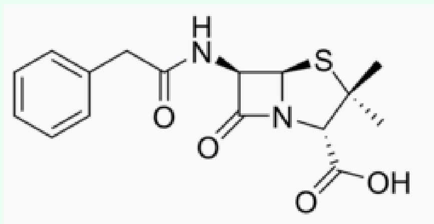
Cephalosporins

- Ceftriaxone
- Cefazolin
- Ceftazidime
- Cefotaxime



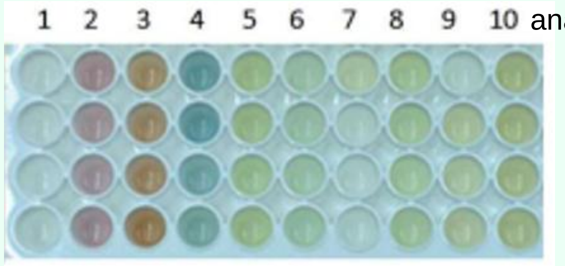
Penicillines

- Benzylpenicillin
- Ampicillin

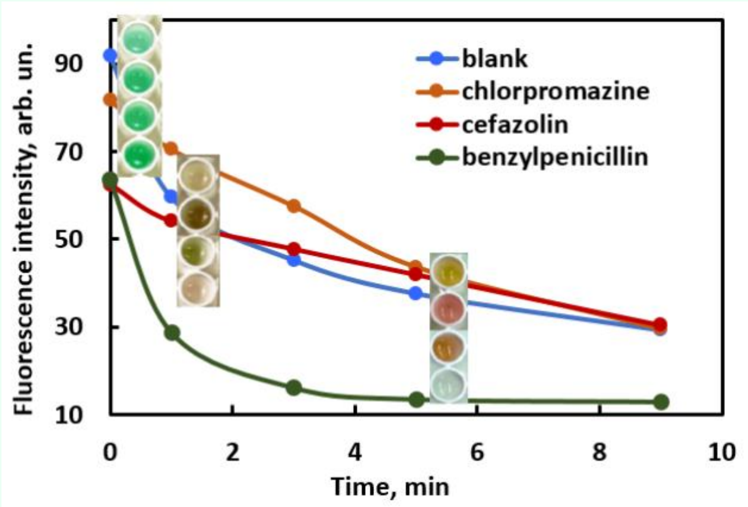


1. **Aggregation-based reaction**
2. **Redox reaction**

Oxidation of a dye in the presence of 9 samples (in 5 min)



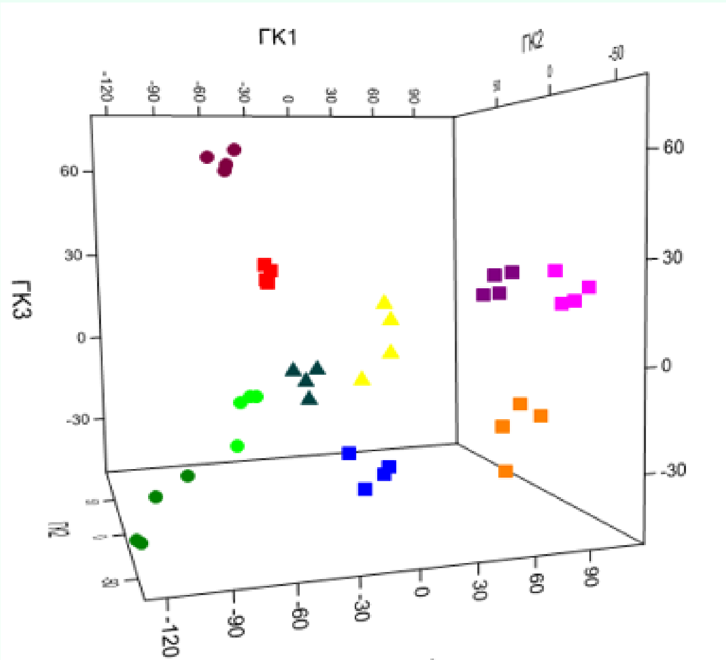
Kinetic curves of NIR fluorescence intensity of dye **3**



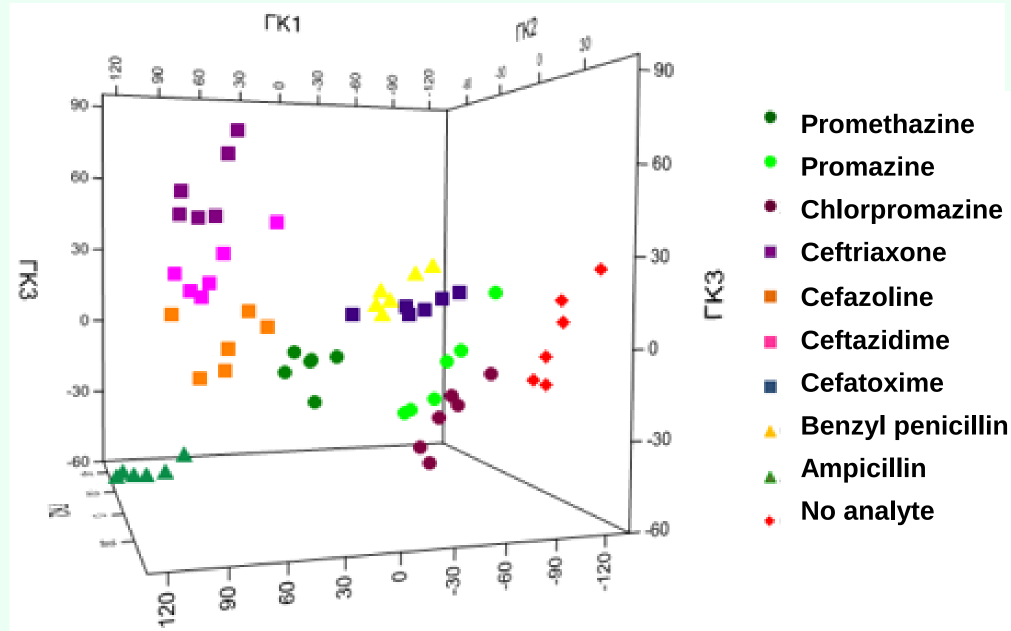
1. Recognition of 9 pharmaceuticals

Principal component analysis score plots

in water



in the presence of turkey homogenate



- All 9 compounds are distinguished in the space of three principal components.
- Compounds not recognized in GC1-GC2 coordinates are recognized in GC1-GC3 coordinates.

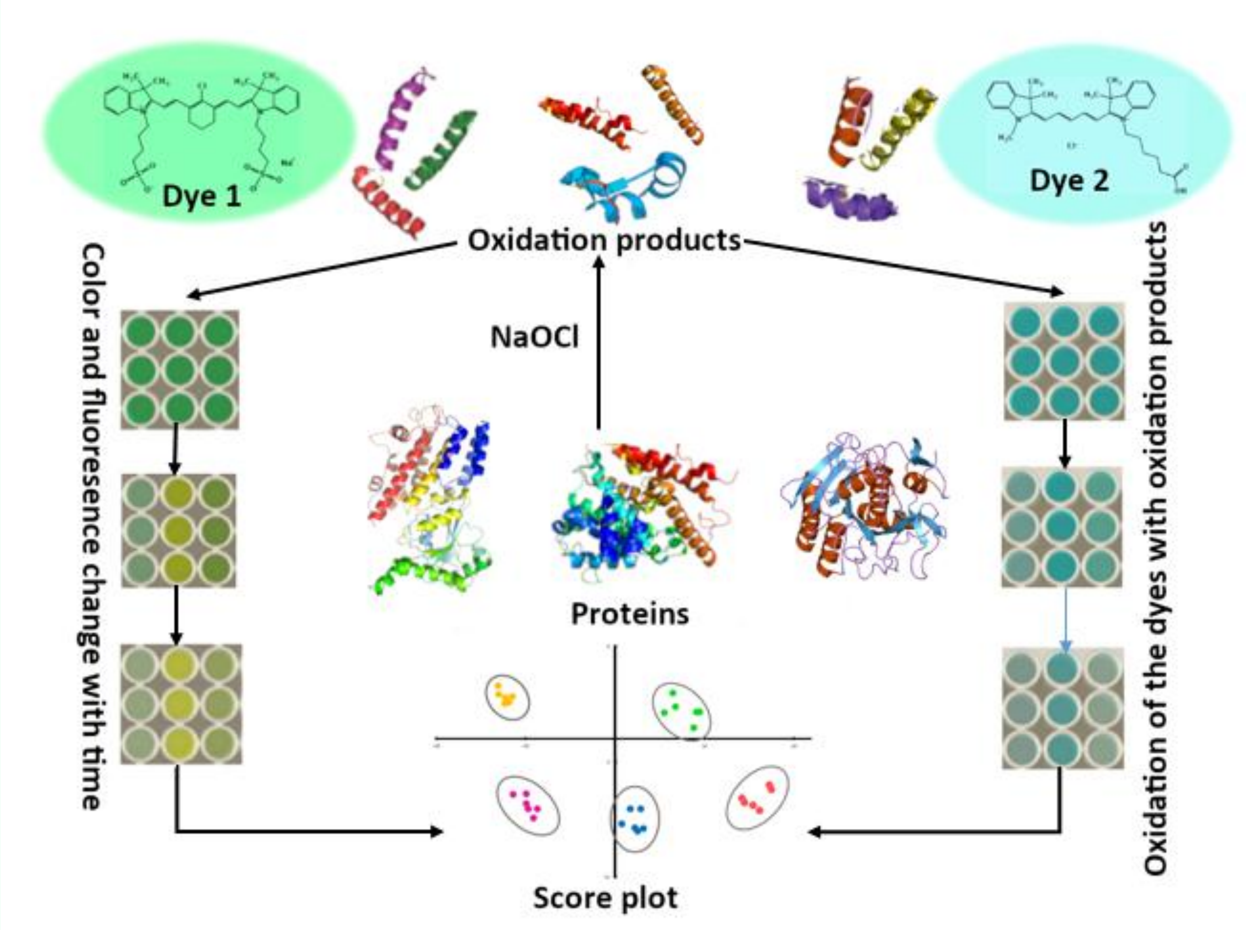
- 7 out of 9 substances are recognized in the space of three principal components.
- Discriminant analysis: 100% recognition accuracy.

Shik, A.V. et al. *Chemosensors* **2022**, *10*, 88

2. Preoxidation of samples

Example: Lin, H.; Jang, M.; Suslick, K.S. **Preoxidation** for Colorimetric Sensor Array Detection of **VOCs**. *J. Am. Chem. Soc.* **2011**, *133*, 16786–16789.

Recognition of proteins and rennet samples by oxidation with hypochlorite

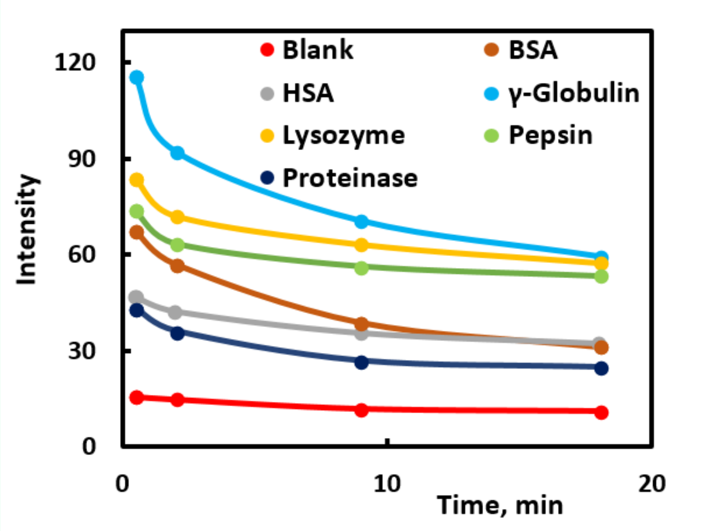


Chloramines:
R-NHCl
R-NCl_2

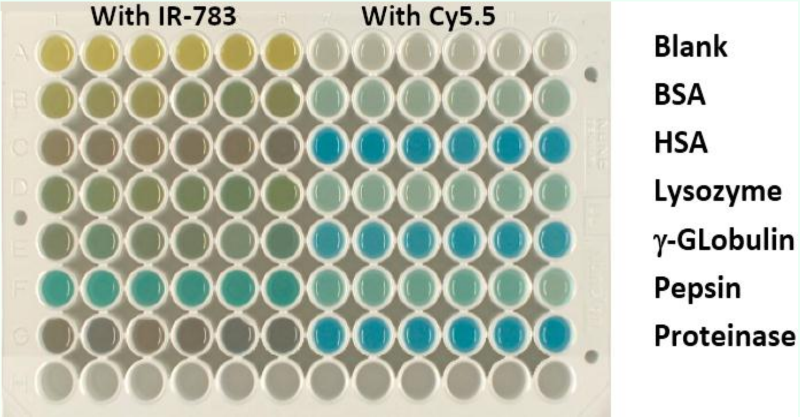
Shik A.V. e.a. *Sensors*, **2023,** 4299.

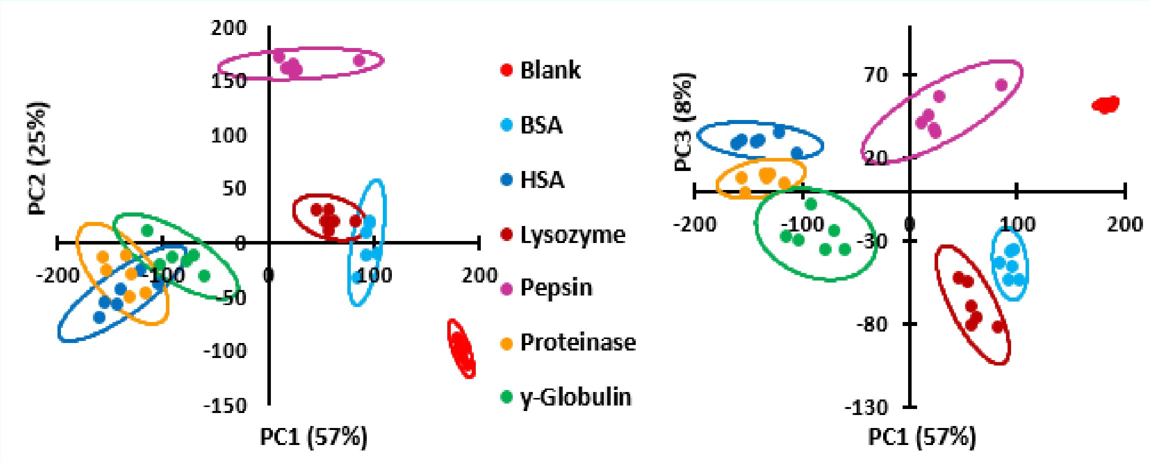
2. Recognition of 6 proteins using hypochlorite oxidation

Kinetic curves of dye oxidation in the presence of proteins (NIR emission)



A 96-well plate with reaction mixtures (photo in the visible range; 6 replicates)





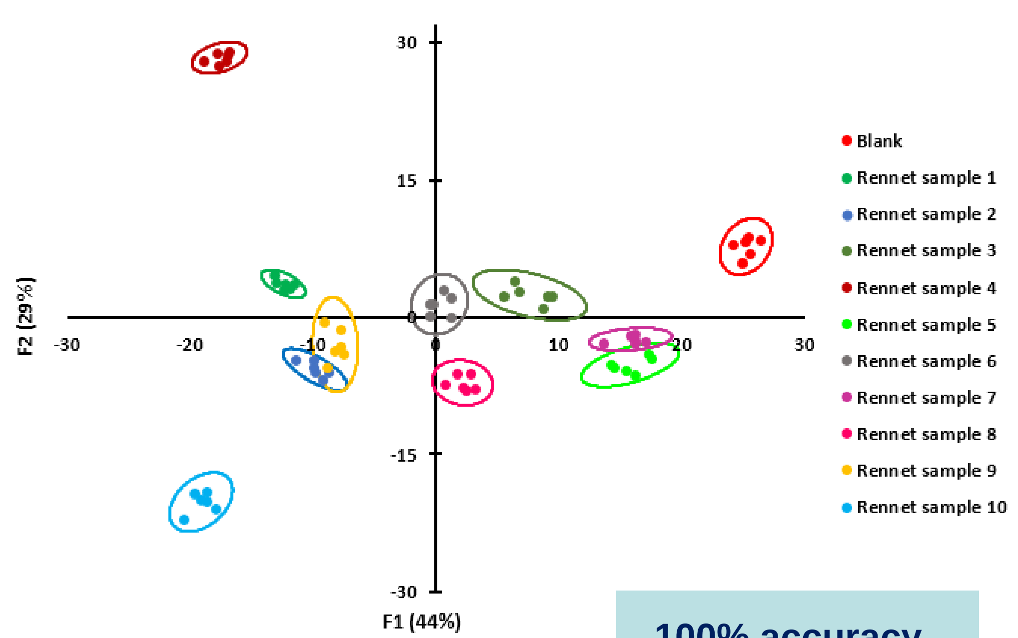
Principal component analysis (PC1 – PC2, PC3) score plots for the oxidation of two dyes by oxidation products of 6 model proteins

2. Discrimination of Rennet Samples using Hypochlorite Oxidation

List of rennet samples (**chymosin + pepsin**)

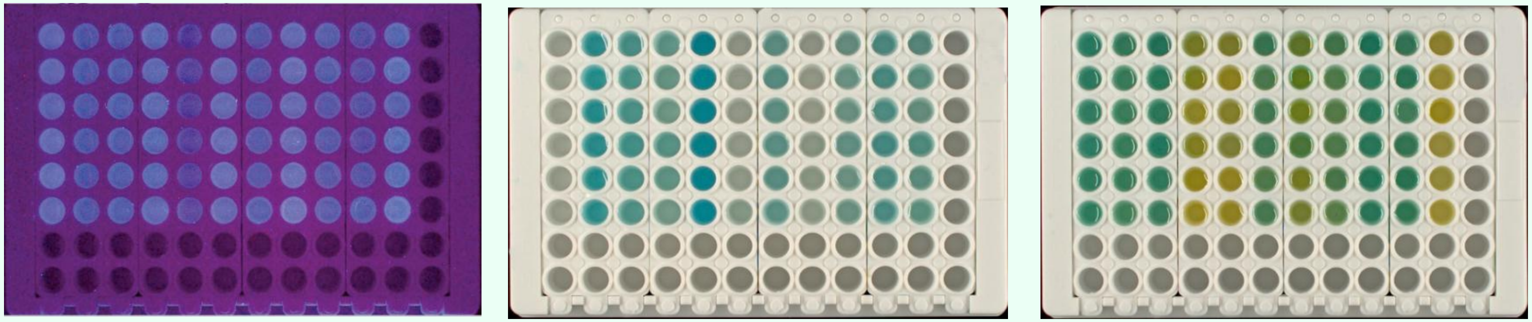
No	Composition of samples (country of origin)
1	Chymosin from *Kluyveromyces lactis* yeast (Spain)
2	Chymosin 90% / bovine pepsin 10% (Russia)
3	Chymosin from *Rhizomucor miehei* fungus (China)
4	Calf chymosin 50% / bovine pepsin 50% (Italy)
5	Bovine pepsin 50% / avian pepsin 50% (Russia)
6	Chymosin 70-80% / bovine pepsin 20-30% (Russia)
7	Pepsin from *Rhizomucor miehei* fungus (Italy)
8	Chymosin 90% / bovine pepsin 10% (France)
9	Chymosin 90% / bovine pepsin 10% (Russia)
10	Pepsin from *Rhizomucor miehei* fungus (Japan)

Linear discriminant analysis (LDA) score plot for 10 rennet samples



100% accuracy

Images of the plates with reaction mixtures in the recognition of rennet samples



3. Food treatment with ionizing radiation

- Suppression of microorganisms, shelf life extension, storage and transport at higher temperatures
- Destruction of pests
- Fruit ripening delay
- Stimulation or delay of germination, etc.

Exceeding the allowed dose is harmful! Dose control:

- Irradiation of the dosimeter *simultaneously* with the product (thermoluminescent dosimeters)
- **Determination of the dose *after* irradiation: methods used:**

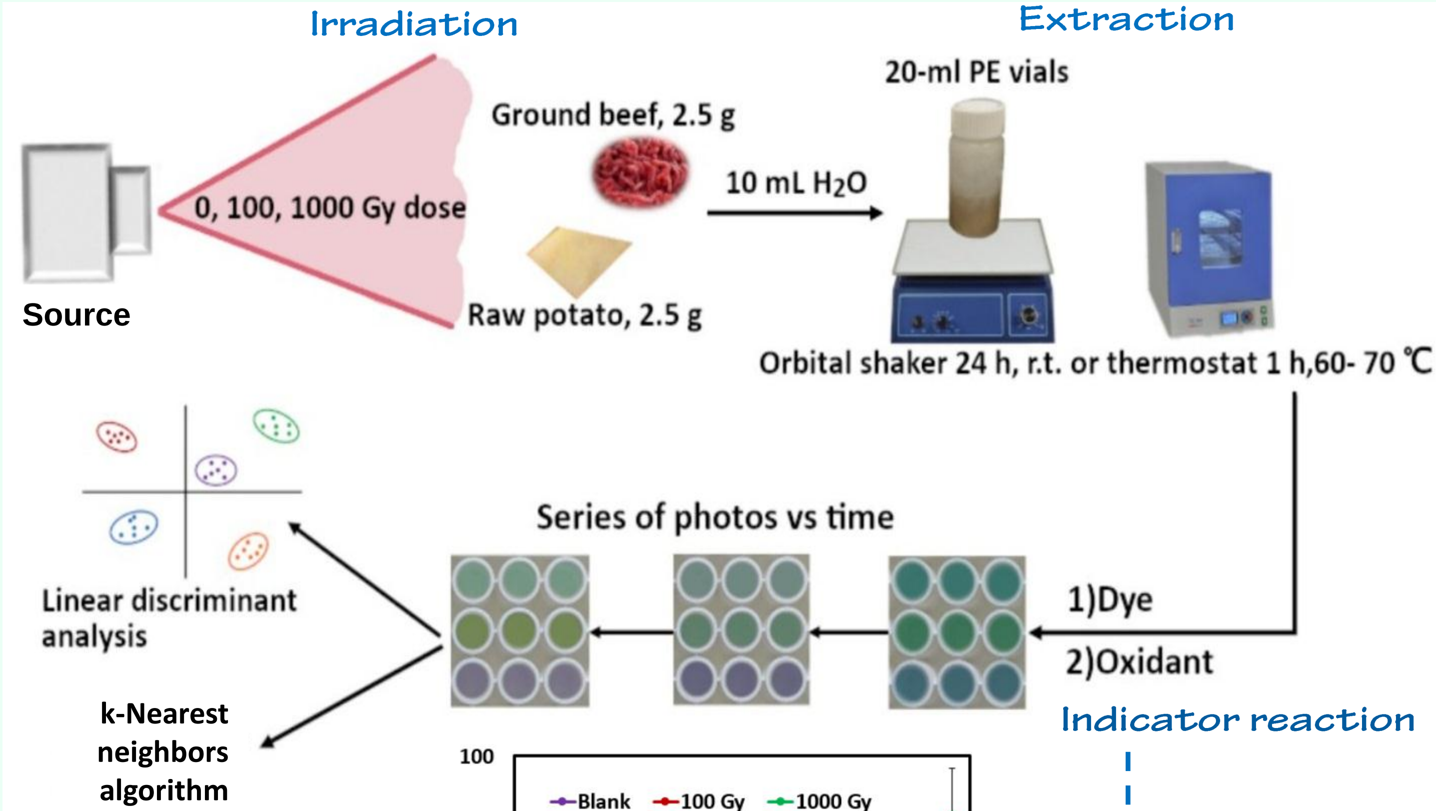
✓ Photoluminescence, thermoluminescence and ESR (if there are solid particles)
✓ IR spectroscopy (near IR)
✓ Determination of antioxidants (DPPH, FRAP, etc.)
✓ Determination of markers (alkylcyclobutanones, dihydrothymidine, etc.)
✓ Determination of DNA decay products (gel electrophoresis)
✓ Real time PCR
✓ Electronic nose

General disadvantages: sophisticated instrumentation, complicated protocol, long sample pre-treatment, low sample throughput. Cannot detect low doses.

"Post treatment" determination of the dose absorbed by irradiated food

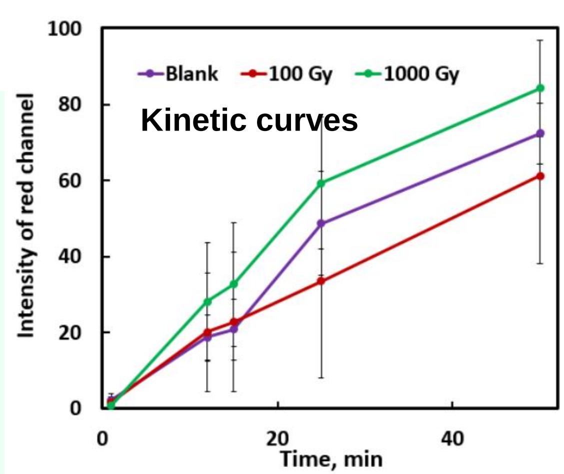
Fingerprinting strategy	Result
By intrinsic absorption and emission spectra	None
By addition of fluorophores to samples	None
By conducting indicator reactions (photometric and fluorimetric control)	**Positive**

Testing of Irradiation Dose in Beef and Potatoes by Reaction-Based Optical Sensing Technique



Data processing

Accuracy of dose estimation



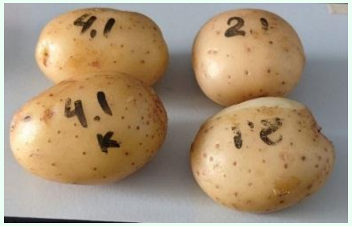
Dyes: carbocyanines, Rhodamines, Crystal Violet
Oxidants: H_2O_2, $KBrO_3$, NaOCl
Catalysts: Mo(VI), Cu(II), V(V)

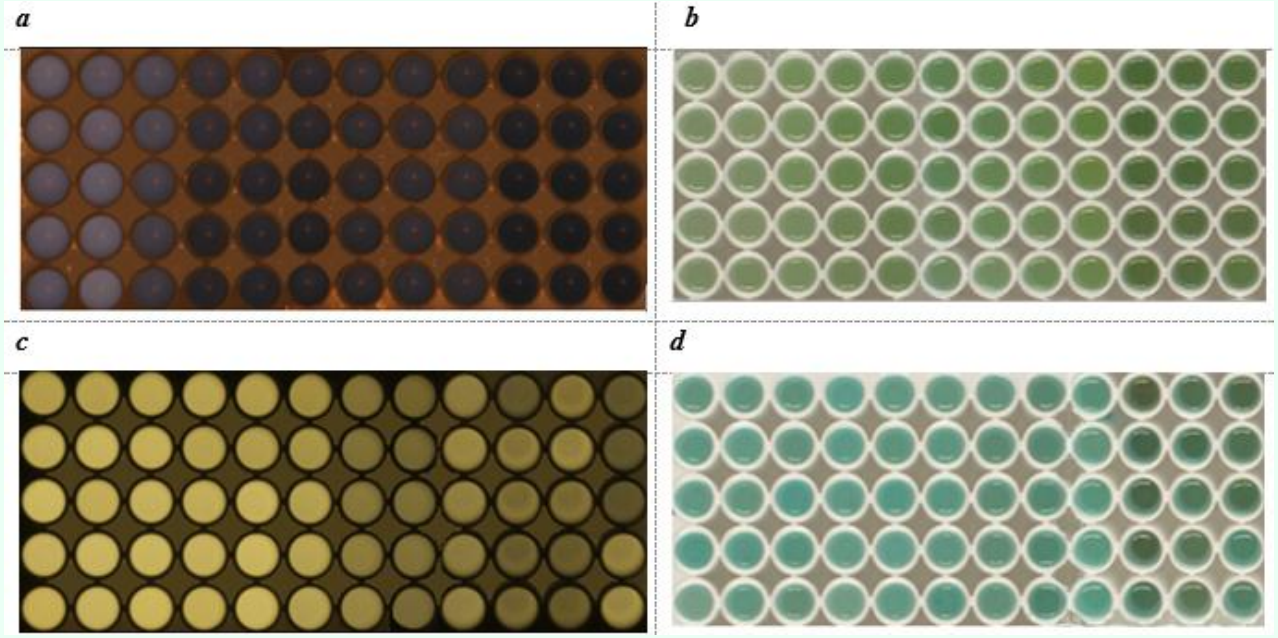
Potatoes: irradiation and measurement protocol

- Electron beam **irradiation**
- Removing and discarding the peel
- Using the top layer of pulp (1 mm)
- Adding antioxidant (ascorbic acid or sulfite)
- **Extraction** with water (4 g / 10 mL, 24 h, r.t.)
- Mixing with indicator **reaction** components
- **Photographing** every several minutes (NIR and visible / absorbance and fluorescence)
- Image processing by LDA, kNN techniques

Irradiation of potato tubers

- Damage prevention, **shelf life** extension
- Delayed **sprouting** of tubers during storage
- 1000 Gy is the **maximum** permitted dose





Examples of images

a, b – dye 1,
c, d – dye 2,

a, c – NIR fluorescence,
b, d – absorbance/ reflectance (smartphone photo)

Discrimination of potato samples of the same variety irradiated with electrons

(doses of 10, 100, 1000 and 10000 Gy)

Reactions used:

- Two oxidation reactions (with dyes **2** and **3**)
- Aggregation reaction with dye **1.**

22 data columns («full dataset»: absorbance, fluorescence at different times for 3 reactions)

Discrimination accuracy for various datasets

Dataset	Number of data columns*	Accuracy, %
Full dataset	**22**	**100**
Without reaction of dye **1**	20	78
Without oxidation of dye **3**	16	78
Without oxidation of dye **2**	16	85
Only reaction of dye **2**	12	85
Only the data with highest standard deviation	7	64
Data from photographs selected based on the largest visual difference	5	57

* Each column contains data for one indicator reaction at certain time.

100% accuracy – only for the full dataset

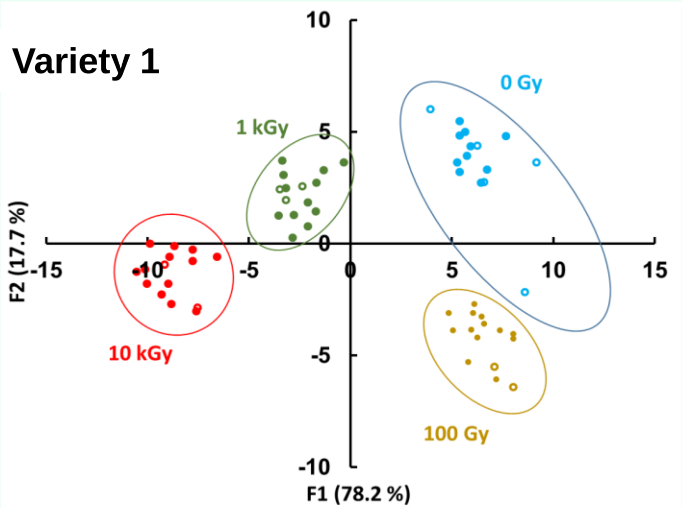
Using fewer data columns worsens the discrimination results

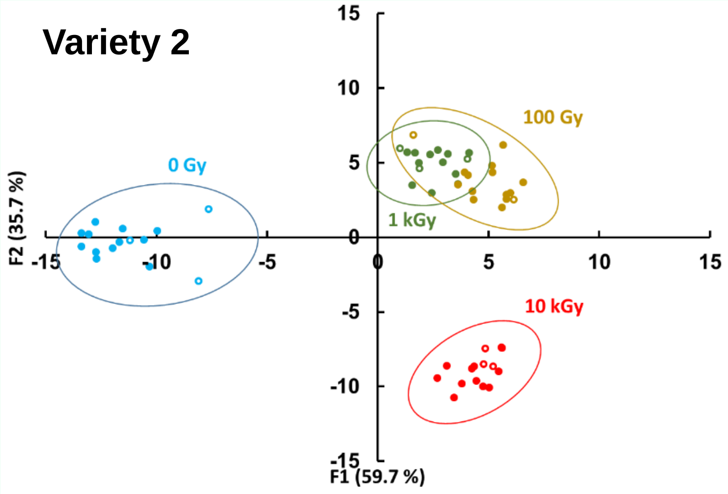
Shik A.V. et al. *Food Chem.*, 414 (2023) 135668

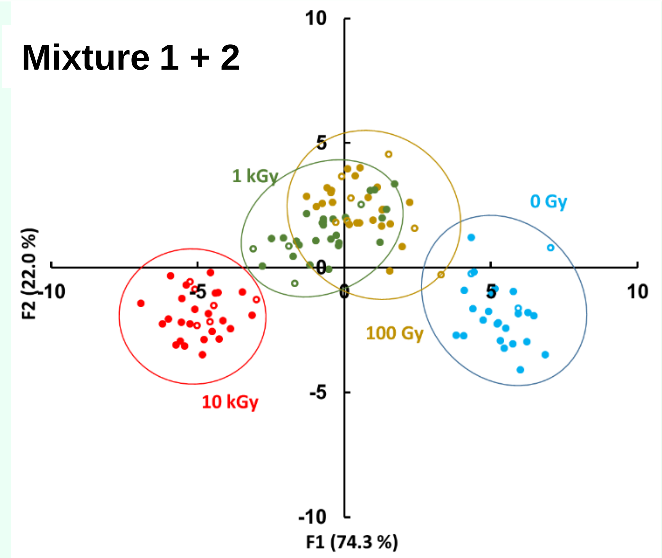
Two potato varieties

(0, 100, 1000, and 10,000 Gy doses)

LDA score plots







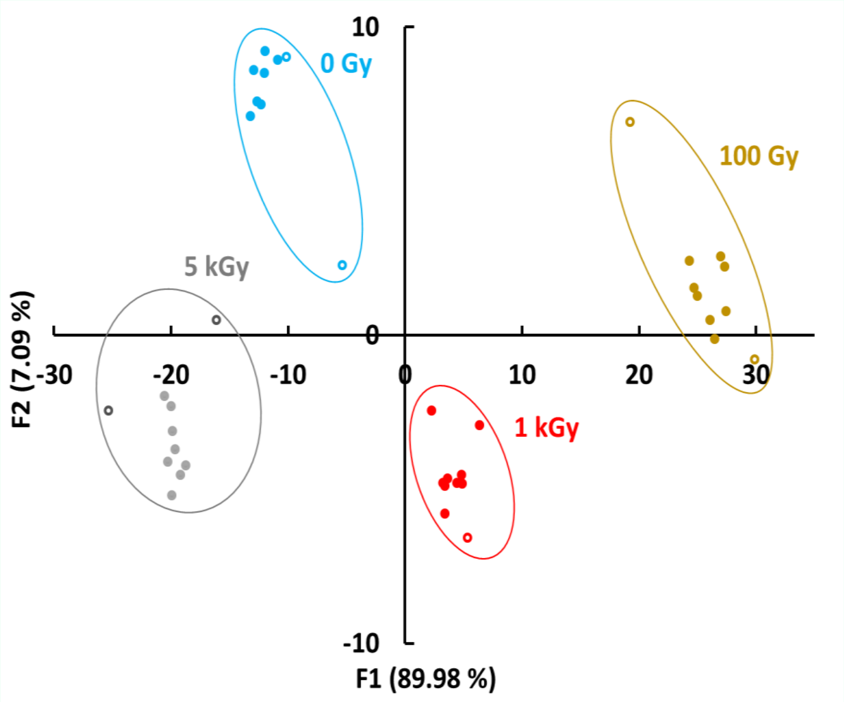
.

Training sets: full symbols, validation sets: empty symbols.

Discrimination of a mixture of tubers of **two varieties** is less efficient than one variety **(85% accuracy)**

X-ray-irradiated potatoes (100 Gy, 1 and 5 kGy) (one variety)

LDA score plot

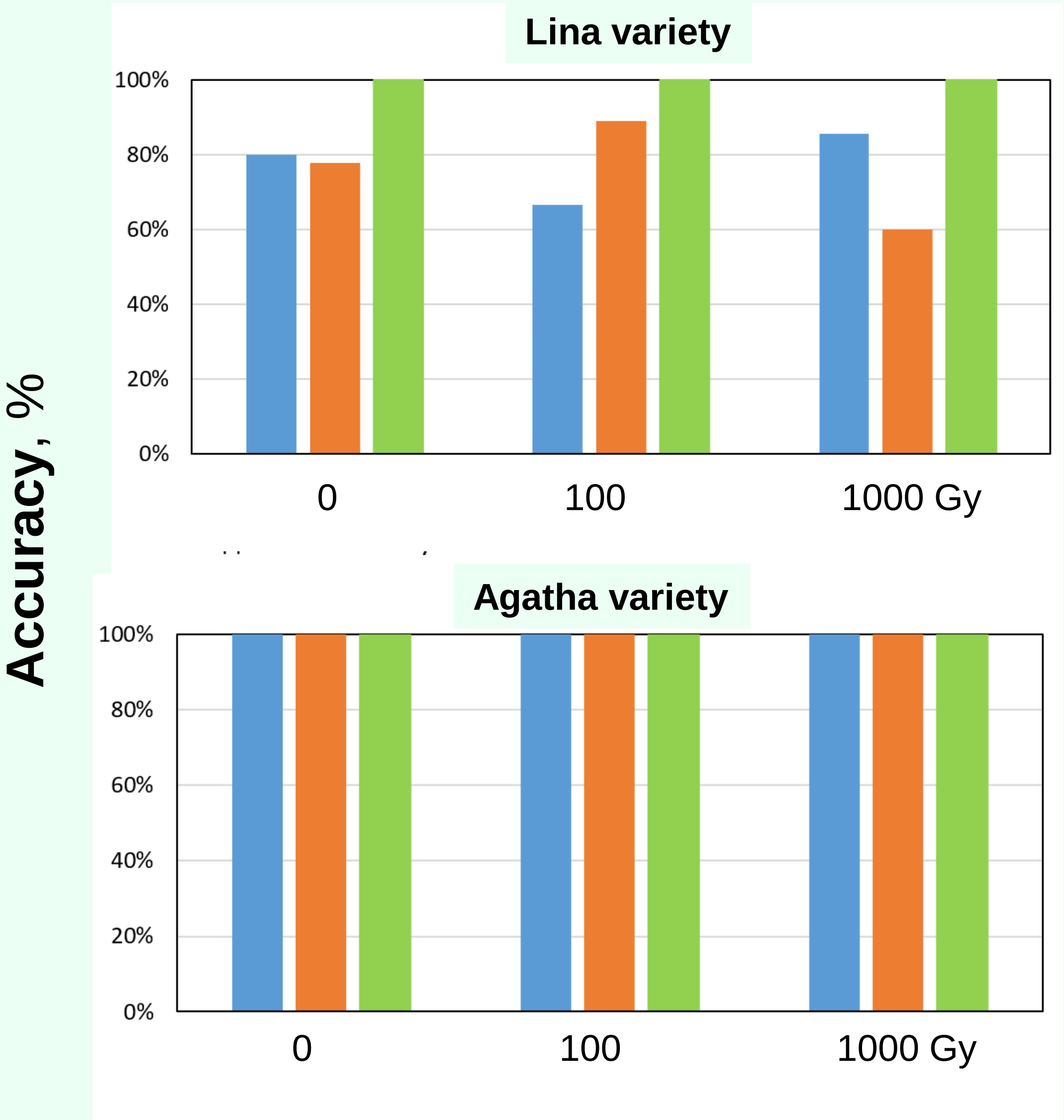


Accuracy of dose estimation: 95%

- Both types of radiation work similarly (**electrons** or **X-rays**)
- **Irradiated** samples are confidently distinguished from **non-irradiated** ones
- The order of dose is estimated with 85-100% accuracy .

Shik A.V. et al. *Food Chem.*, 414 (2023) 135668

Accuracy of dose estimation in irradiated potatoes (X-ray) stored during 0, 2, and 6 days after irradiation



Stable compounds are responsible for the signal

For electron-irradiated samples:
after 6 months in the refridgerator: 100%

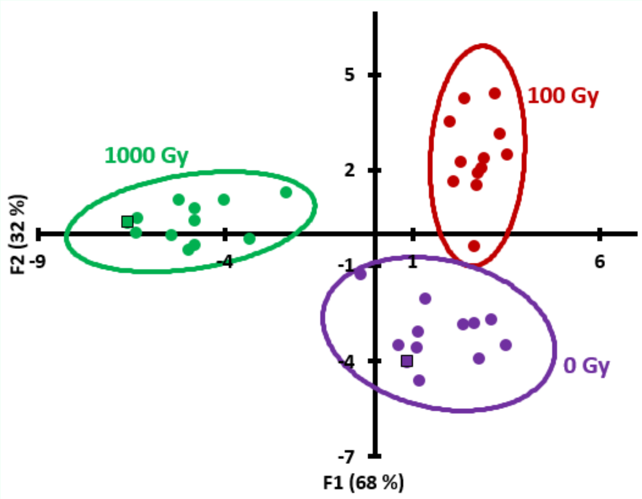
Zubritskaya et al. (submitted)

3. Testing of Irradiation Dose in Raw Ground Beef

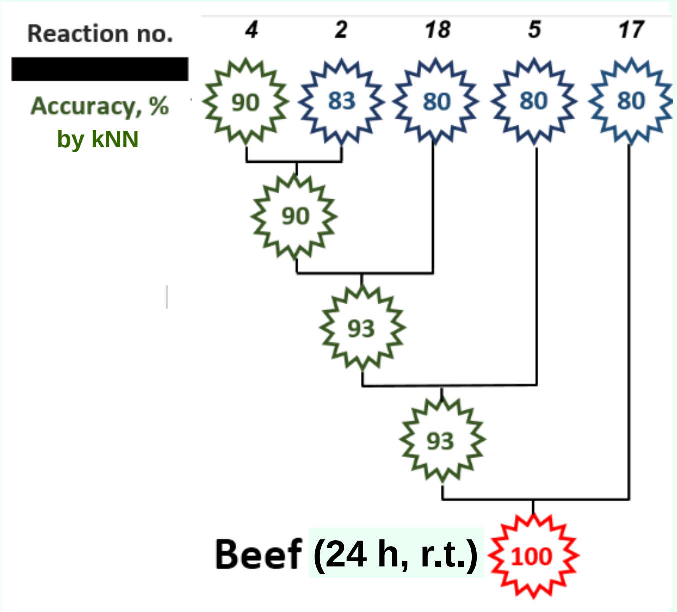
Best indicator reactions for beef

Individual reactions	No.	Accuracy, %, by LDA
Dye 4 - metamizol - H_2O_2	4	**93**
Cy5.5 - NaOCl	18	**93**
Dye 2 - H_2O_2	2	**90**
Dye 3 - lysozyme - NaOCl	17	87
IR-783 - H_2O_2	3	80
TAMRA - bromate	5	63

Beef extracted during 24 h, 23°C (LDA)
(reaction with dye **4**)



Merging data from 5 reactions (kNN):



One indicator reaction:
90-93% accuracy (by LDA);
merged **5 reactions**:
100% accuracy (by kNN)

Shik A.V. et al. *J. Food Compos. Anal.* (submitted)

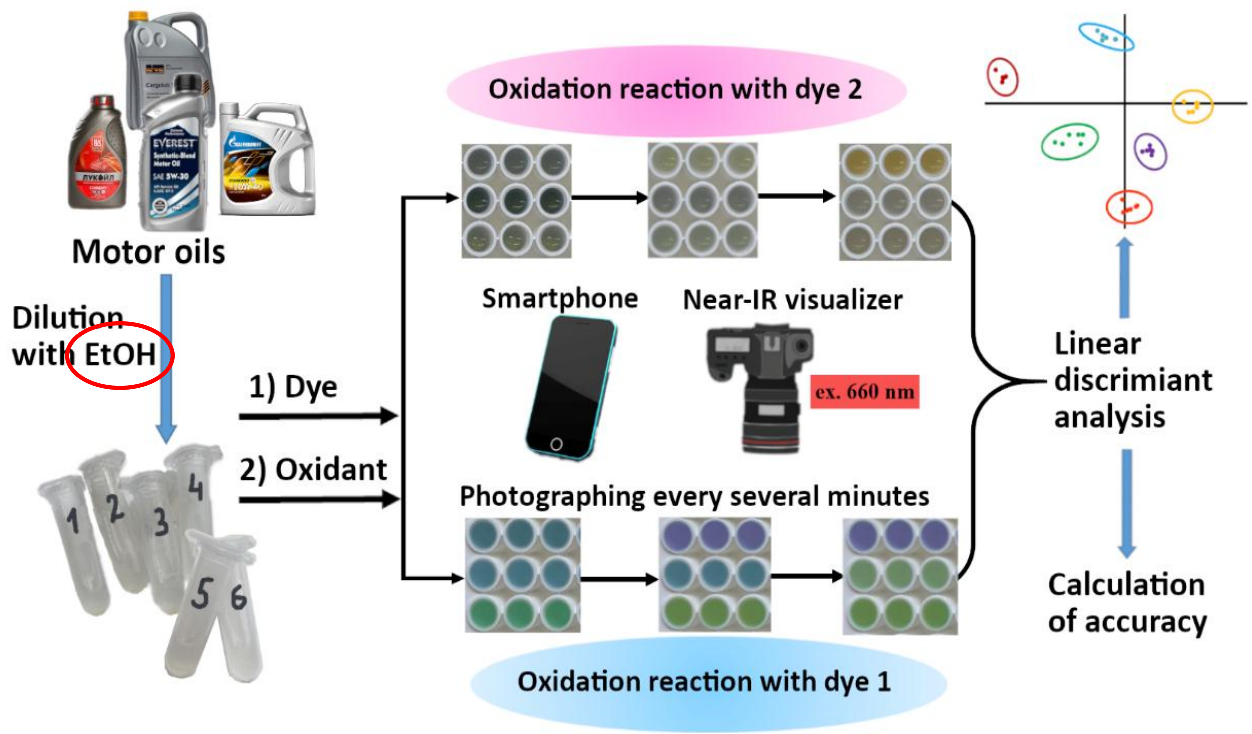
Accuracy of individual observation *vs*. accuracy for the whole sample

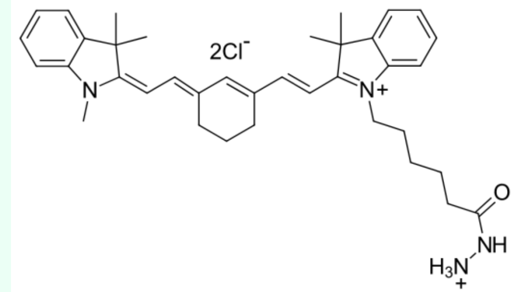
All the reported accuracy values pertain to **single observations**

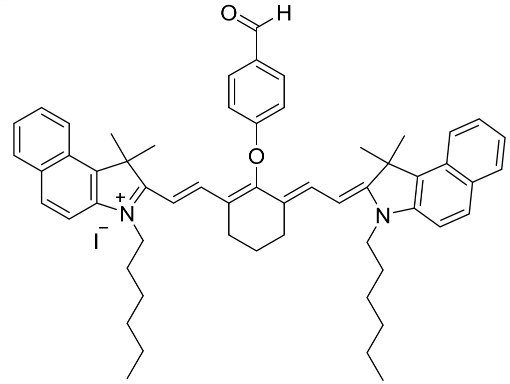
Number of parallel observations for the sample	Accuracy of dose estimation	
	for a **single observation**	for the **whole sample**
6	**90%**	**98.5%**
	93%	**99.4%**

4. Recognition of Fat-Soluble Samples: Discrimination of Motor Oils

- Our indicator reactions are usually carried out in **aqueous** solutions
- Here we developed the reactions that occur in **ethanolic** medium (minimum water)**,** which makes it possible to work with **fat-soluble** samples.







Dyes

Oxidants

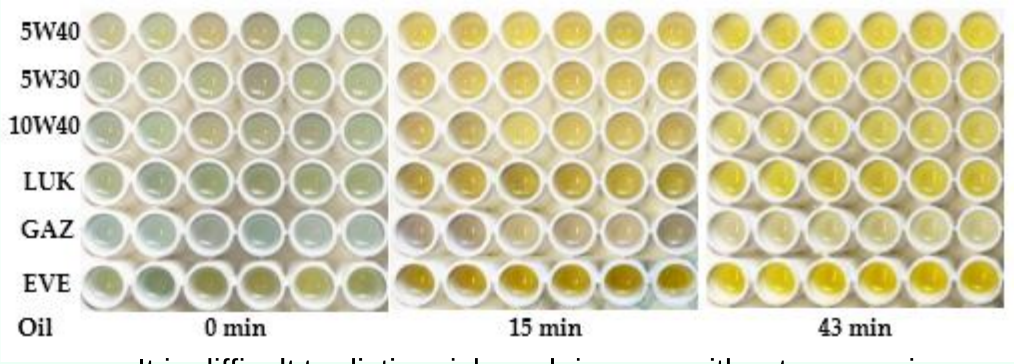
- t-BuOOH
- t-BuOOH (Cu^{2+} as catalyst)
- HNO_3
- Aqua regia
- O_2

4. Discrimination of Motor Oils

Six oil samples:

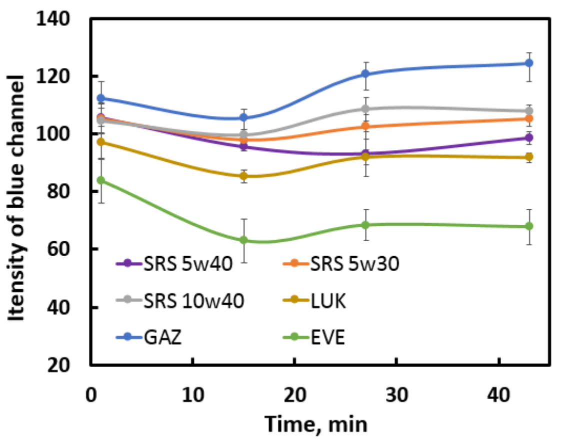
Designation	Name	SAE grade	Manufacturer
SRS 5W30	Cargolub TFX	5W30	SRS Schmierstoff Vertrieb GmbH
SRS 5W40	Cargolub TFX	5W40	SRS Schmierstoff Vertrieb GmbH
SRS 10W40	Cargolub TFX	10W40	SRS Schmierstoff Vertrieb GmbH
LUK	Genesis Armortech	5W30	Lukoil (LLK International)
EVE	Everest	5W40	US Global Petroleum
GAZ	Gazpromneft Premium	10W40	Gazpromneft-SM

Indicator reaction: oxidation of dye **1** with HNO_3: images

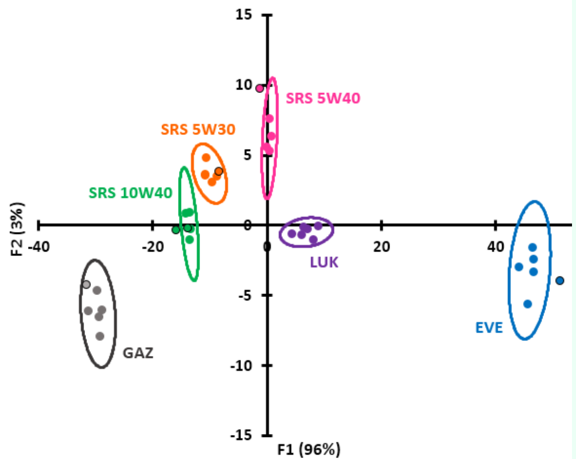


It is difficult to distinguish such images without processing

Kinetic curves



LDA score plot for the reaction of dye **1** + HNO_3



- **100% accuracy** (LDA) by one reaction
- **Other** fat-soluble samples can be possibly recognized

Reaction-based fingerprinting: major totals

Advantages

- Various types of samples
- No prior knowledge on markers required
- Short analysis time
- High sample throughput (5−10 hr^{-1})
- Simple protocol
- No full-spectrum instruments
- No high-skilled personnel
- Standard software for image digitizing and data processing
- Sometimes, only commercially available reagents

Limitations

- For a new type of sample: re-select the indicator reactions
- Always analyze standards in parallel with unknown samples (*common to all fingerprint methods*)
- No understanding of markers ("black box" technique)

Prospects

- New indicator reactions, more sensitive to the composition of samples (improving the accuracy of discrimination)
- Combination with biological techniques

Thank you for your attention!

beklem@inbox.ru