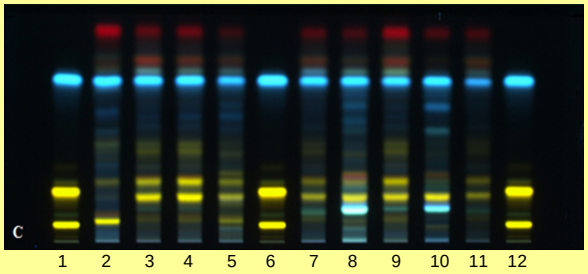
# Pattern Recognition in HPTLC-Fingerprints of Medicinal Plants [1]

Rebekka Ambühl[1], Anita Ankli*[2], Nicolas Richerdt [2], Eike Reich[2]
[1] Institute of Pharmaceutical Biology, Department of Pharmaceutical Sciences, University of Basel, Switzerland
[2] CAMAG Laboratory, Sonnenmattstrasse 11, 4132 Muttenz, Switzerland; Corresponding author: anita.ankli@camag.com

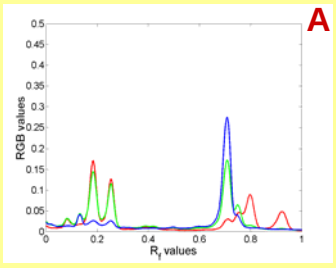
**Summary**
Identification of medicinal plants based on visual comparison of images of HPTLC fingerprints must take into account the natural variability of plants in addition to the variance of the analytical process. For reliable identification of herbal species an appropriate electronic database of many fingerprints is therefore mandatory. Such database could be the basis for an objective and investigator-independent identification tool.
This poster presents a multivariate statistic approach (PCA-MANOVA) that is capable to cluster and identify sets of HPTLC data. The first ten principal components of 64 samples have been used for evaluation. This lead to satisfactorily distinct clusters and it was possible to allocate an unknown sample to the correct species. The allocation was successfully performed with the multivariate distance value.

**Introduction**
Herbal products are complex compositions of various compounds. Their analysis is a challenge given the natural variability of ingredients due to a large number of variables like climate conditions, soil composition, harvesting, etc. For quality control and identification purposes characteristic profiles called fingerprints are commonly used. One established method for generating such fingerprints is HPTLC, which also offers the option of presenting the results as colorful images.
The goal of this work was to find an approach to compare and cluster HPTLC fingerprints of herbal species with the use of standard statistical methods.

**Material**
ADC2 with humidity control, immersion device, plate heater, ATS4, new HPTLC software in beta version (all CAMAG), MATLAB (The MathWorks Inc.); HPTLC plates Si 60 $F_{254}$ (Merck); p.a. solvents (Acros, Merck, Roth); standards (USP); 64 different plant samples from the Lamiaceae family, including *Thymus vulgaris* (TH), *Salvia officinalis* (SA), *Oregano vulgare* (OR), *Rosmarinus officinalis* (RO), *Melissa officinalis* (MA), and Melissa dry extract (MAE) from Iteipmai, Chrüterhüsli, Migros, Coop, PhytoLab.




Track 3 →




RGB →

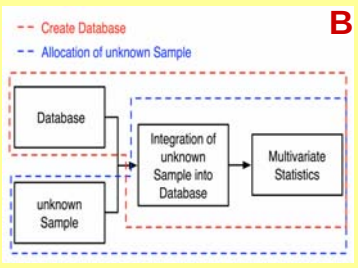



**Fig. 1: Data acquisition**
Documentation under UV 366nm after derivatization, Tracks 1, 6, 12: rutin, hyperoside and rosmarinic acid with increasing Rf values;
2: *Thymus capitata*; 3, 4, 7, 9, 11: *T. vulgaris*; 5, 8 ,10: *T. serpyllum*.

**Fig. 2: Data analysis**
**A:** Individual track of one thyme sample defined as RGB profile
**B:** Data analysis process

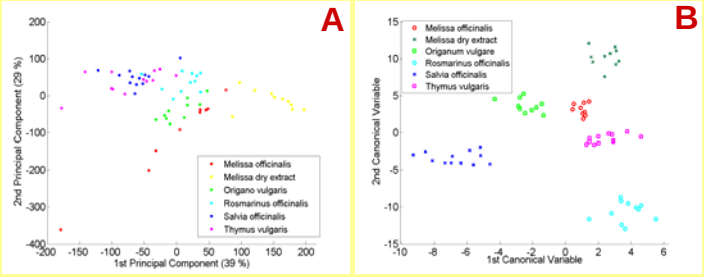



**Fig. 3: PCA and MANOVA of fingerprints under UV 366nm after derivatization**
**A:** PCA score plot of first two principal components
**B:** Canonical analysis score plot of the first two variables

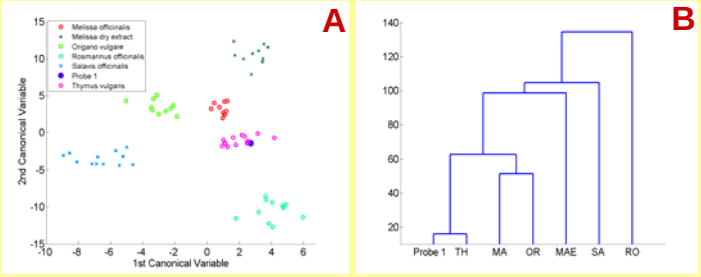



**Fig. 4: Identification of an unknown sample (Probe 1) by MANOVA**
**A:** Canonical analysis score plot of the first two variables
**B:** Dendrogram of group means

**Method**
500mg of powdered plant material and 5mL of methanol were mixed and sonicated for 10min, then centrifuged. 1mg of the standards rutin, hyperoside and rosmarinic acid were dissolved in 1mL of methanol.
Spray-on application of 2µL of samples and standards as 8mm bands. Development in Automatic Developing Chamber (ADC2), 20min chamber saturation, developing distance 70mm, mobile phase ethyl acetate, formic acid, water (15:1:1), plates conditioned to 33% RH. Heating 3min at 100°C and derivatization by immersion into NP reagent (1g of 2-aminoethyl diphenylborinate in 200mL of ethyl acetate) and PEG reagent (10mL of polyethyleneglycol 400 in 200mL of dichloro-methane) (Fig. 1).

Each image of a track on the HPTLC plate after derivatization under UV 366nm defined a fingerprint and for electronic assessment it was described by its retardation factors and the corresponding RGB values (Fig. 2; A). Fingerprints were evaluated by principal component analysis (PCA) and multivariate analysis of variance (MANOVA) (Fig. 2; B).

**Results and discussion**
64 different samples from various species were evaluated. PCA alone was not capable to separate distinct clusters (Fig. 3; A). Therefore, PCA was combined with MANOVA leading to an approach that is known as PCA-MANOVA.
Considering the first ten principal components well separated clusters of species were achieved (Fig. 3; B) and allocation of an unknown sample ("Probe 1", *Thymus vulgaris*) to the corresponding clusters was possible (Fig. 4; A).
It was assumed that the unknown sample can be identified by allocation to the group to which the distance was shortest in a multidimensional space (Fig. 4; B).
By calculating the multivariate distance (md) to the group means, the unknown sample was identified as *Thymus vulgaris* because the distance to the mean of the *Thymus vulgaris* cluster revealed the smallest value (md=16). The mds to all other groups were 118 or higher.

**Literature**
[1] R. Ambühl, Master thesis, Inst. of Pharmaceutical Biology, University of Basel, 2011