Hardware and Software Solution for Planar Chromatography

Cumulative inaugural dissertation

by

Fichou Dimitri

Submitted to the

Faculty of Biology and Chemistry

Prepared at the

Institute of Nutritional Science

Department of Food Sciences

For the degree of

Doctor rerum naturalium (Dr. rer. nat.)

Justus-Liebig-University Giessen, Germany

Giessen 2018

Contents

Thesis reviewers

First Reviewer Prof. Dr. Gertrud E. Morlock

Chair of Food Sciences, Institute of Nutritional Science, Justus-Liebig-University Giessen

Second Reviewer Prof. Dr. Alexander Goesmann

Bioinformatics and Systems Biology, Justus-Liebig-University Giessen

First Examiner Prof. Dr. Bernhard Spengler

Institute of Inorganic and Analytical Chemistry, Justus-Liebig-University Giessen

Second Examiner Prof. Dr. Wolfgang Schwack

Institute of Food Chemistry, University of Hohenheim, Stuttgart, Germany

Declaration

I declare that I have completed this dissertation without the unauthorized help of a second party and only with the assistance acknowledged therein. I have appropriately acknowledged and referenced all text passages that are derived from published literature, and all information that relates to verbal communications. I have abided by the principles of good scientific conduct laid down in the charter of the Justus Liebig University Giessen in carrying out the investigations described in the dissertation.

Giessen, June 2018

Fichou Dimitri

Acknowledgements

To Claudia

Peer-reviewed contributions

1. D. Fichou; G. Morlock; **Anal. Chem. 88 (2017) 12494-12501**. *Open-Source-Based 3D Printing of Thin Silica Gel Layers in Planar Chromatography.*
2. D. Fichou; P. Ristivojevic; G. Morlock; **Anal. Chem. 89 (2017) 2116-2122**. *Proof-of-Principle of rTLC, an Open-Source Software Developed for Image Evaluation and Multivariate Analysis of Planar Chromatograms.*
3. D. Fichou; G. Morlock; **Anal. Chem. 90 (2018) 6984-6991**. *Powerful Artificial Neural Network for Planar Chromatography Image Evaluation Shown for Denoising and Feature Extraction.*
4. D. Fichou; G. Morlock; **J. Chromatogr. A 1560 (2018) 78-81**. *quanTLC, an Online Open-Source Solution for Videodensitometric Quantification.*
5. D. Fichou; I Yuce; G. Morlock; **in submission**. *In Silico and in Situ tools for Signal Highlighting in High-Resolution Mass Spectrometry Hyphenated to Planar Chromatography.*
6. D. Fichou; G. Morlock; **in submission**. *Office Chromatography: Miniaturized All-in-One Open-Source System for Planar Chromatography.*

Oral contributions

1. **11th Balaton Symposium on High-Performance Separation Methods**, Siofok, Hungary, 06.-08.09.2017: G. Morlock, D. Fichou, T. Häbe *Office Chromatography - an open-source based system*
2. **International Symposium for HPTLC 2017**, Berlin, 04.-08.07.2017: D. Fichou, G. Morlock *Office Chromatography*, O-4 **Received a Young Researcher Award**

Poster contributions

1. **3rd International Conference on Natural Products Utilization**, Bansko, Bulgaria, 18.-21.10.2017: G. Morlock, D. Fichou, T. Häbe, I. Yüce *Open-source developments for natural product search*
2. **23rd International Symposium on Separation Science**, Vienna, Austria, 19.-22.09.2017: G. Morlock, D. Fichou, T. Häbe, I. Yüce *Open-source developments for the toolbox of analysts*
3. **International Symposium for HPTLC 2017**, Berlin, 04.-08.07.2017: D. Fichou, P. Ristivojevic, G. Morlock *rTLC: Open source software for multivariate analysis of HPTLC data*, P-4
4. **International Symposium for HPTLC 2017**, Berlin, 04.-08.07.2017: D. Fichou, G. Morlock *Open-source developments for Office Chromatography*, P-3 **Awarded with the Bronze Poster Prize**
5. **International Symposium for HPTLC 2017**, Berlin, 04.-08.07.2017: D. Fichou, G. Morlock *Application of artificial neural network to planar chromatography data*, P-2
6. **40th Symposium for 'Chromatographic methods of investigating the organic compounds'**, Szczyrk (Poland), 23.-26.05.2017: D. Fichou, G. Morlock *Office Chromatography - do it yourself!* **Awarded with the First Poster Prize**
7. **SEP 2017**, Paris, 29.03.2017_: D. Fichou, G. Morlock *Printing of layers, samples and mobile phases in Office Chromatography*, P113
8. **33rd International Symposium on Microscale Separations and Bioanalysis (MSB 2017)**, Noordwijkerhout, 26.-29.03.2017: D. Fichou, G. Morlock *Office Chromatography - do it yourself!*, P95
9. **eRum 2016, european R users meeting**, Poznan, Poland, 12.-14.10.2016: D. Fichou, G. Morlock *Application of artificial neural network to planar chromatography data*

Abstract

Pushed by the citizens and governments, the demand for quality control is increasing in numerous industries. In analytical chemistry especially, innovative apparatus and data analysis solutions are constantly needed to face those challenges.

Using the open-source RepRap 3D printing environment, thin layers of silica gel suitable for planar chromatography were printed. A low-cost 3D printer was modified by replacing the plastic extruder by a slurry doser for production of stationary phases. The final apparatus opened new possibilities for tailored plates, both in terms of layer shape and composition. The development of this device was greatly facilitated by the minimal modifications needed to repurpose the 3D printing electronics and software.

The same strategy was applied to produce an apparatus for office chromatography. In a single miniaturized device, several steps of the planar chromatography pipeline were performed. Liquids were printed via a thermal inkjet cartridge, enabling drop-on-demand of 150 pL with a resolution of 96 dpi. The device was controlled by dedicated software hosted on a Raspberry Pi and available on the local network. Sample application was quantitative with correlation coefficients superior to 0.999. Mobile phase printing for the separation of dyes and parabens was also conducted via inkjet. The Raspberry Pi camera enabled an easy implementation of the documentation step. LEDs were used for the illumination and opened new possibilities in terms of selectivity. The final apparatus was compact, modular and affordable.

In modern analytical chemistry, where there is hardware to produce data, software is needed to analyze it. Using the R programming language and in particular the shiny package, several web applications were developed for data analysis.

To fulfill the need for free, dedicated and fully featured solutions for quantitative evaluation of videodensitograms, quanTLC was created. For an intuitive user experience, the user interface was kept minimalistic, enabling a fast and reproducible analysis.

Multivariate analysis of planar chromatography data is gaining interest among researchers and industry. However, there is no all-in-one solution to perform such analysis. The rTLC software was developed for this purpose. All necessary steps were implemented, *i.e.* data extraction, preprocessing, variable selection, unsupervised and supervised statistics.

Before chromatogram images are evaluated, preprocessing is often necessary. The use of unsupervised learning with artificial neural network was investigated. Inhomogeneous background and noise were removed with this technique. In addition, the new features learned by the network show an improvement in resolution while keeping the quantitative aspect.

Due to the tremendous amount of information produced in high resolution mass spectrometry, data interpretation can be time-consuming. If several data analysis software are available for this task, none of them is focused to the hyphenation of mass spectrometry with planar chromatography. A tailored software, called eicCluster, was created for this technique. After bucketing of the m/z values, the powerful t-Distributed Stochastic Neighbor Embedding algorithm was used to cluster the extracted ion chronograms. The data dimension was reduced to a 2D map where isotopes and fragments from the same molecules were clustered together. The user could then explore the dataset via interactive visualization tools and draw conclusions otherwise hidden.

1 Introduction

In the last years, two fields have attracted a great interest in the media, 3D printing and data sciences. Led especially by the open-source movement, those technologies are now available for free in the case of software, and at a reasonable cost in the case of hardware. This new paradigm has the power to change the way new software and hardware are developed in analytical chemistry. The current challenges are interdisciplinary by nature, and the access to powerful yet affordable tools allows to look for alternative solutions not available before.

1.1 Hardware evolution in planar chromatography

1.1.1 History of planar chromatography

In 1938[1], planar chromatography was invented. In concert with the other techniques like gas chromatography (GC) and high-performance liquid chromatography (HPLC), planar chromatography had evolved considerably during the end of the 20th century to become high-performance thin-layer chromatography (HPTLC) in 1975[2]. The first aspect relevant to this development came from the use of modern instrumentation. Automatized sample application, chromatographic separation and digital cameras allowed a better control of the experiment, a critical point in analytical chemistry. Progress in stationary phase manufacturing was also crucial, with smaller particles in narrower range, typically 5-6 µm, resolution was improved leading to better separations. Finally, the hyphenation with other techniques like mass spectrometry, bioassays for effect directed analysis and densitometry gave new information to the analyst.

1.1.2 UTLC layers

In 2002, Merck commercialized the first UTLC layers based on a monolithic layer of silica gel with 10 µm thickness[3,4]. Performances were demonstrated with the separation of different analytes and compared to HPTLC and TLC layers. Aside from monolithic layers[5], other approaches were explored over the years. These included electrospun polymers[6], glancing angle deposition of silica and metal oxides[7] and low pressure chemical vapor deposition of silicon nitride onto carbon nanotube templates[8].

Those new UTLC layers presented attractive specifications, *e.g.* low solvent consumption, short migration time, smaller migration distance with the potential to lead to smaller carrier and thus smaller apparatus. However, the thinner layer had a smaller capacity and lower sample volumes had to be applied. The spray-on and contact application were not compatible with those fragile layers. The influence of the gas phase during the separation was critical and had to be controlled. The instrumentation footprint was over-sized for the small carrier. Thus, available instrumentation technology was not compatible with those layers.

1.1.3 The office chromatography concept

To overcome those limitations, the office chromatography (OC) concept was coined in 2010[9] (Fig. 1A). The idea was to combine all analytical steps in a single miniaturized device. The idea of a combined system was already patented in 1991[1], but the notion of miniaturization was not incorporated. At this time, the classical format of the chromatographic plate was 20x20 cm and thus an all-in-one apparatus able to handle such format was never commercially available. It must be stressed that the state of the art sample application device, the ATS 4 from CAMAG (Muttenz, Switzerland), was also made for this format. In addition, its minimal reproducible volume of application of 100 nL already overloaded UTLC layers. UTLC layers have a typical format of 10 x 5 cm^2, and an apparatus built around this size could keep a reasonable footprint, and thus, the integration of several analytical steps will be facilitated.

The OC concept used print and media technologies (PMT) to achieve its objectives. The incompatibility of spray-on application with the UTLC layers encouraged to look for other sample application techniques. The use of inkjet printing to print solutions on HPTLC layers was first demonstrated in 2007[10]. The authors modified an inkjet printer, (Canon Pixma iP 3000x) to print a hydro-alcoholic derivatization solution of ninhydrin for the determination of taurine in energy drinks. In 2010 and 2015, the same printer was further modified, this time for sample application on UTLC layers[9,11] (Fig. 1B).

1.1.4 Inkjet printing

Inkjet printing was initially described by Lord Rayleigh in 1878 as a liquid jet of constant radius falling vertically under gravity[12]. It was only in 1951 that Rune Elmqvist patented

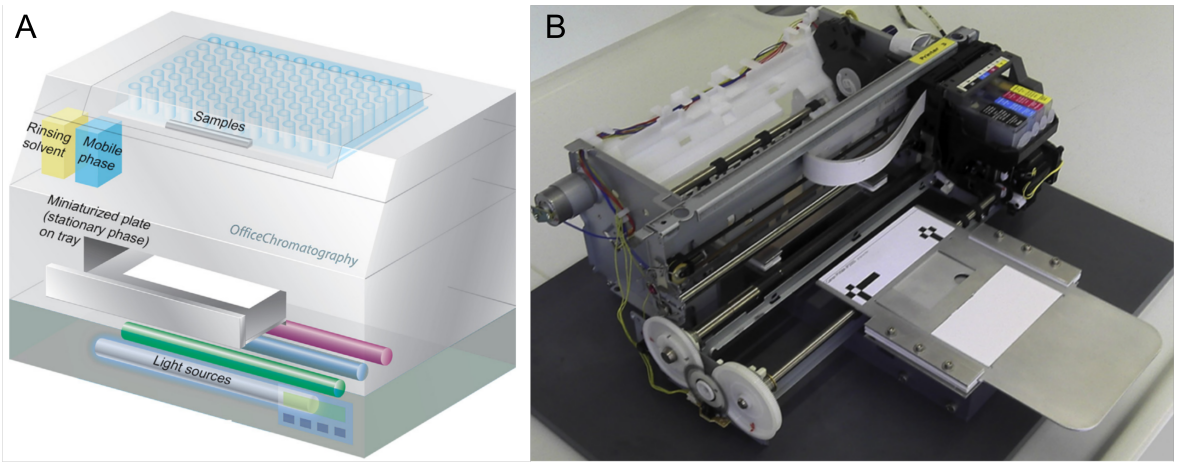


Figure 1: OC design proposal (A) and modified printer (B). Reprinted with permission from Anal. Chem. 82 (2010) 2940 and J. Chromatogr. A 1413 (2015) 127. Copyright © 2010 American Chemical Society and 2015 Elsevier B.V..

the idea of the first commercial inkjet recorder[13]. Inkjet techniques can be divided into two main modes of operation, continuous and drop-on-demands. Continuous inkjet printers rely on the constant creation of ink droplets, ink is supplied by a high pressure pump and droplets are generated through the nozzle by a piezoelectric crystal. This type of printer can print at higher speed but with a lower resolution. Most of the drop-on-demand inkjet printers use either thermal or piezoelectric principles[14,15]. In thermal inkjet (Fig. 2), the ink chamber consists of a heater with a nozzle nearby. A current pulse of a few microseconds is applied to the heater, which transfers heat to the ink. A vapor bubble expands forcing the ink out of the nozzle, once all the heat is transferred, the bubble collapses and the ink droplets break off. In a piezoelectric inkjet printer (Fig. 3), an electric impulse deforms a piezo-ceramic, which generates a pressure wave causing the ink to be ejected from the nozzle.

The modification of commercial printers had been demonstrated previously, for HPTLC[9–11] as well as in the bioprinting field[16]. However, those approaches suffered from the lack of print head control, limiting their use to few laboratory applications. To overcome those limitations without being dependent on PMT companies, several open-source solutions are available. The open-source board InkShield (http://nicholasclewis.com/projects/inkshield/) allows the control of 96 dpi HP C6602 thermal inkjet cartridges from an arduino (http://arduino.cc/). An open-source piezoelectric inkjet print-head also had to be designed and validated for bioprinting[17].

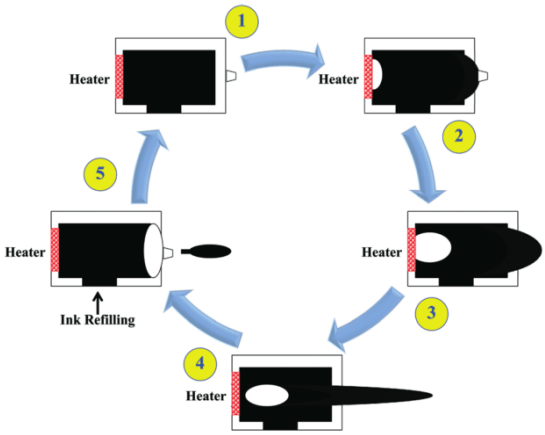


Figure 2: Drop-on-demand thermal inkjet printer. Reprinted with permission from Lab on a Chip 15 (2015) 2538. Copyright © 2015, Royal Society of Chemistry

1.1.5 3D printing

3D printing, also called additive manufacturing was developed by Charles Hull who patented stereolythography in 1986[18] and founded 3D Systems. Several approaches had been proposed to generate 3D objects, always based on the successive application of layers of materials. Stereolythography (SLA) is based on photopolymerisation of a liquid resin by a UV lamp. Between the layers, the stage is translated vertically so the next layer can be polymerized. SLA printers are capable of producing high resolution parts (25 µm layers) but are expensive and the resulting objects are fragile. Inkjet printing can also be used to produce 3D parts by printing a liquid material on layers of powder to bind its particles. Layers are successively distributed by a roller and bound together selectively where the liquid binding material is applied, with the stage dropping vertically between each layer. Selective laser sintering (SLS) is also a powder-based technique with the difference that the polymer powder is sintered together by high power lasers. Layers are distributed on the stage and the laser used to locally increase the temperature to the powder melting point. The advantage of SLS lies in the wide range of materials usable for fabrication. Fused deposition modeling (FDM) is one of the most widespread 3D printing techniques. Thermoplastic filament is driven to the nozzle tip of the extruder, where it is heated to a semi-molten state. The nozzle tip is then moved in the Cartesian space to generate, layer after layer, a 3D object. FDM is compatible with several materials, the most used one being acrylonitrile

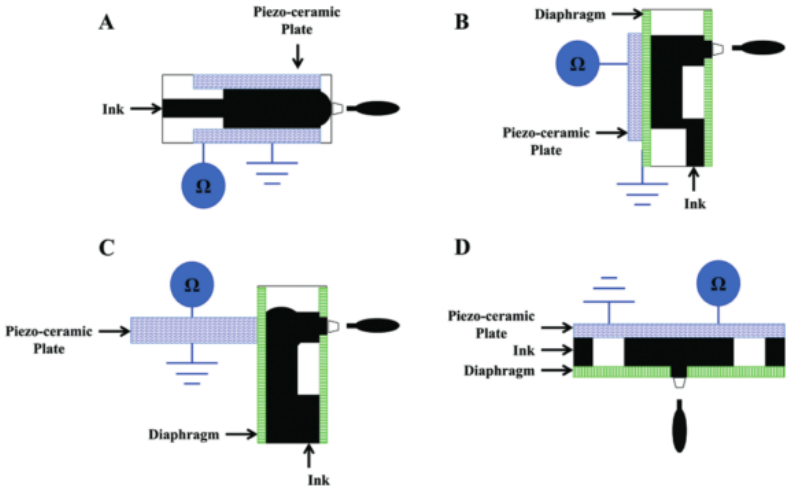


Figure 3: Different design of drop-on-demand piezoelectric inkjet printer; squeeze mode (A), bend mode (B), push mode (C) and shear mode (D). Reprinted with permission from Lab on a Chip 15 (2015) 2538. Copyright © 2015, Royal Society of Chemistry

butadiene styrene (ABS) and polylactic acid (PLA). Others are also available like nylon or polycarbonate, but are more difficult to print with low-cost printers.

3D printing opened new possibilities in a large number of domains including analytical chemistry[19,20]. It was used to build a device for the reproducible preparation of nanospray tips for capillary electrophoresis[21]. Two different devices were designed for paper-assisted spray ionization mass spectrometry[22,23]. The lack of commercial devices for low-temperature plasma ionization was overcome by 3D printing[24]. Reactors were designed and 3D printed for analysis by electrospray[25,26]. SLS was used to design titanium micro-bore column later filled with stationary phase[27,28], which opened new possibilities for column design[29]. Columns of different geometry were also produced in ABS[30]. A radial flow-cell for chemiluminescence detection applicable for ion chromatography was designed and gave an alternative to the conventional design, while increasing signal magnitude and duration[31]. An FDM 3D printer was used to fabricate a low-cost fluorescence detection head for capillary separation[32]. A 3D printed housing with a modified Raspberry Pi camera was used to build a UV-spectrophotometer with a 1-nm resolution[33]. Another 3D printed housing, combined with a mobile phone and a UV lamp were used to document TLC chromatograms[34]. All those diverse examples show the limitless new potential of 3D printing in analytical chemistry.

In 2008, the patents held by 3D Systems and other companies in the sector started

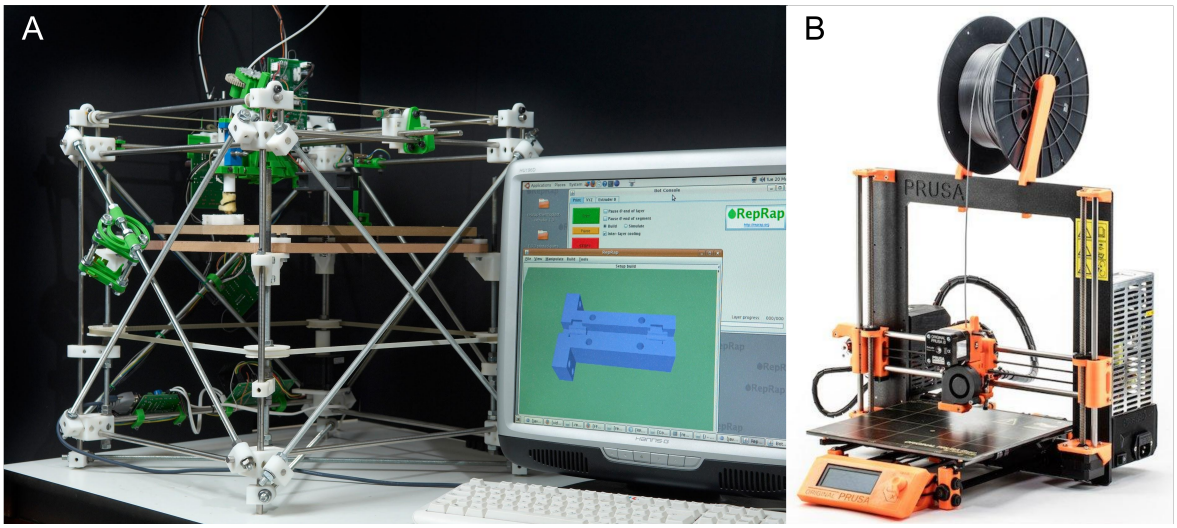


Figure 4: Darwin version 1, the first RepRap 3D printer (A) and the Prusa i3 (B). Reprinted with permission from Robotica 29 (2011) 177. Copyright © 2011 Cambridge University Press.

to expire, opening the way to the RepRap project, initiated by Adrian Bowyer from Bath University[35]. RepRap are 3D printers that are open-source and self-replicating. They are designed to be able to print their own structural parts, while the remaining must be easily available, *e.g.* motors, rods, bearing. By being open-source, anyone can reproduce and modify it, thus encouraging its dissemination and improvement. The first RepRap, called Darwin[35], was FDM based (Fig. 4A). The Prusa i3, designed by Joseph Prusa, is one of the most prevalent designs (Fig. 4B). Those printers can be bought as a kit priced between 300 and 1000 $[36].

Operating those RepRap printers could not be possible without open-source electronics, firmware and software (Fig. 5). The first step in the production of a 3D object takes place in computer assisted design (CAD) software, *e.g.* Blender (https://www.blender.org/) and OpenSCAD (http://www.openscad.org/). The 3D model once ready will be exported in a Standard Tessellation Language or Stereo-Lithography (STL) format. Accepted as standard, the STL format stores the 3D model in a text file via coordinates of each triangulated section forming the 3D model[19]. To convert this shape into machine readable orders, the STL file has to be converted into a GCODE file via a slicer software, *e.g.* Slic3r (http://slic3r.org/). The slicer software will divide the 3D model into a succession of 2D horizontal cross-sections and generate a sequence of GCODE commands that the 3D printer will follow. To send the GCODE to the

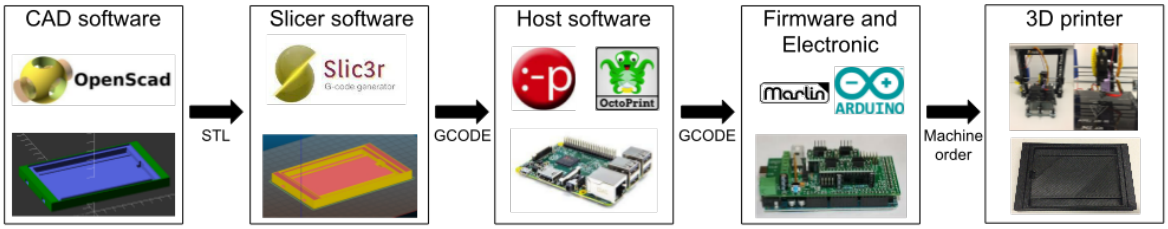


Figure 5: Schematic of the 3D printing process.

printer, a host software is necessary, *e.g.* Printrun (http://www.pronterface.com/) and OctoPrint (https://octoprint.org/). The latter in particular can be hosted on a Raspberry Pi (http://raspberrypi.org/) and allows remote control of the printer via a web application. Interpretation of those GCODE commands is made on the printer side by a firmware. One of the most popular open-source firmware for 3D printing is Marlin (http://marlinfw.org/), which can be easily customized to fit the specific needs of different devices. Finally, electronic boards are necessary to host the firmware and send electronic signals to the different hardware parts of the 3D printer. The Arduino board benefits from a large community and is used in many projects including 3D printing.

The use of 3D printing in the analytical chemistry laboratory to rapidly produce tailored parts for specific tasks is clear. But with the RepRap concept in mind, its use can be extended even more. An apparatus of analytical chemistry is a robot, which will perform tasks to automatize an analytical step, *e.g.* move a motor or set a temperature. RepRap 3D printers are open-source, and thus their electronics and software tool chain are accessible and can be repurposed with minimal knowledge of electronic and programming for usage in a laboratory apparatus. Instead of printing plastic, the same kind of apparatus could be modified to print liquids, *i.e.* sample, mobile phase and derivatization solvents, onto an adsorbent layer. The use of the GCODE standard in particular is appealing, as it allows various and complex actions to be performed easily. In addition, the numerous 3D printer designs and tools developed by the community are a great source of inspiration.

1.2 Data analysis in planar chromatography

Due to its planar aspect, TLC presents interesting particularities compared to GC and HPLC. As samples are applied and separated in parallel, they can be observed and evaluated simultaneously (Fig. 6A). The observation of several samples in GC and HPLC becomes rapidly complex as the number increases. The presence of the totality of the samples on the plate gives a full vision on the difference to HPLC and GC, for which analytes may not elute when remaining in the precolumn and liner. Though a wide range of detectors are available for GC and HPLC, the derivatization in TLC opens up a comparable range, able to detect selectively certain families of compounds with high sensitivity[2]. With the introduction of HPTLC, the results became more reproducible due to the increased control owed to modern instrumentation, and more compounds could be separated with the higher resolution power. The digitalization of the data via a charged coupled device (CCD) camera and flat-bed scanner[37] made the storage of chromatogram images possible for a later evaluation. This digitalization also gave *in silico* access to the signal via a bitmap of pixels, leading to new options for evaluation (Fig. 6B).

1.2.1 Quantitative analysis

The first quantitative evaluation of chromatograms with a digital camera was demonstrated in 1984[1]. For laboratories with limited budget, videodensitometry is a method of choice to perform quantitative analysis on TLC data due to the ever-present access to digital cameras and flat-bed scanners. Different instrumentation for quantification in TLC were compared[38], *i.e.* slit-scanning densitometry, videodensitometry with commercial system and software or digital camera and free software and flat-bed scanning. All approaches gave quantitative results, while the more expensive slit-scanning densitometry was the most sensitive with a limit of detection (10 times lower).

Planar chromatography companies developed software solutions over the years, *i.e.* VideoScan (CAMAG, Muttenz, Switzerland), Sorbfil TLC Videodensitometer (Sorbfil, Krasnodar, Russia)[39], TLSee (Analtech, Newark, DE, USA), Macherey Nagel TLC scanner software (Macherey Nagel, Düren, Germany)[40] and JustTLC (Sweday, Sodra Sandby, Sweden). Other free alternatives are present, but are either limited in capabilities, *i.e.* the R package qtlc available from the command line[41], TLC analyzer written

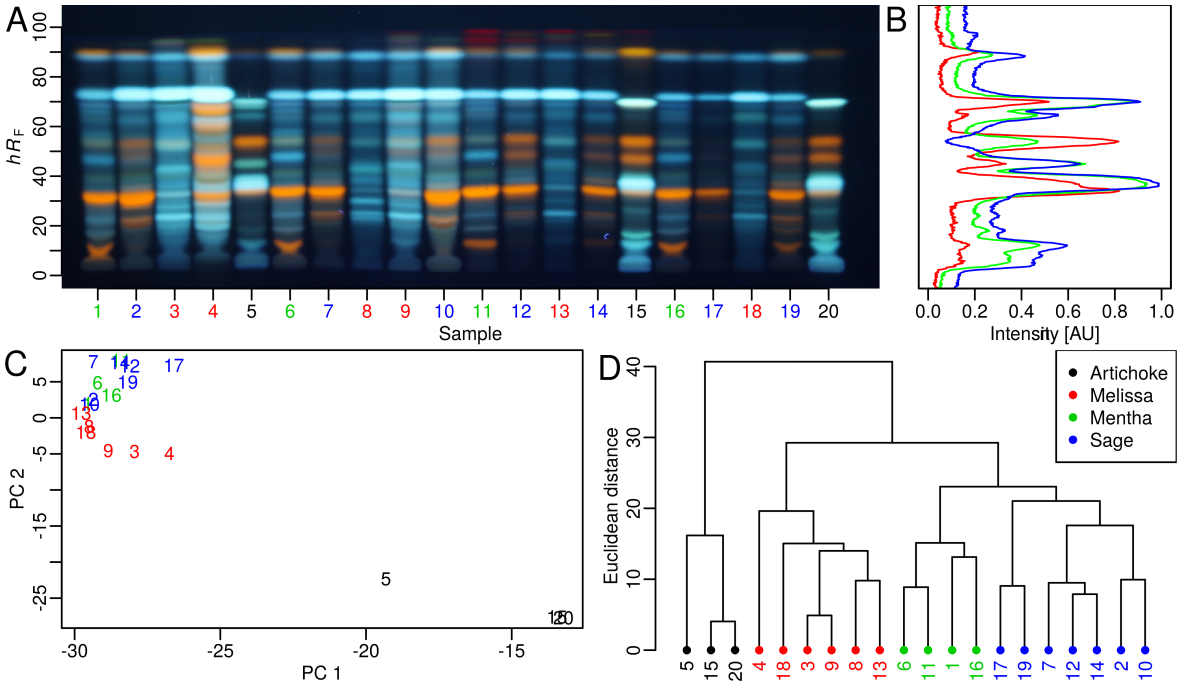


Figure 6: Data analysis with planar chromatography data; chromatogram of medicinal drugs (A), extracted videodensitogram of sample 5 (B), PCA (C) and HCA (D) with the blue channel preprocessed with standard normal variate.

in Matlab[42], or not dedicated to TLC, *i.e.* ImageJ (Research Services Branch, National Institute of Mental Health, Bethesda, MD, USA). None of these solutions was satisfying, *i.e.* freely available and full-featured. Therefore, a new software had to be developed to fulfill this need.

1.2.2 Chemometrics

Apart from quantitative analysis, HPTLC is also used to gather qualitative information, *e.g.* detect adulteration or highlight an active principle. Plate evaluation is classically done visually, comparing different samples in the chromatogram. This process highlights the peaks that will then be proposed as markers and reported in standardized methods. This manual interpretation suffers from its subjectivity. The definition of colors for example can vary from one person to another. In addition, the presence of those markers may not be evident. In those cases, more sophisticated techniques are required.

Table 1: Non exhaustive list of chemometrics studies with HPTLC data

Year	Sample	Unsupervised	Supervised	Reference
2009	Herbs	PCA	K-nn, ANN	43
2009	Herbs	PCA	-	44
2011	Herbs	PCA	-	45
2011	Propolis	PCA, HCA, hierarchical fuzzy clustering	-	46
2011	Pharmaceutical formulations	PCA, HCA	-	47
2012	Herbs	PCA	PLS-DA, OPLS-DA	48
2013	Herbs	-	ANN	49
2014	Microalgea	PCA	-	50
2014	Propolis	PCA, HCA	LDA	51
2014	Herbs	-	K-nn, CART, SPA-LDA, PCA-DA, SVM, PLS-DA	52
2014	Propolis	PCA, HCA	PLS-DA	53
2014	Herbs	PCA	-	54
2014	Propolis	Similarity analysis, HCA, K-means clustering,	ANN, SVM	55
2014	Herbs	PCA	-	56
2014	Herbs	PCA	-	57
2016	Algea	PCA	-	58
2016	Propolis	PCA	-	59
2016	Herbs	PCA	-	60
2016	Propolis	PCA	-	61
2016	Propolis	PCA	-	62
2017	Herbs	-	PLS	63
2017	Herbs	PCA, HCA	-	64
2017	Herbs	PCA, HCA	-	65
2017	Biopolymers	PCA, HCA	-	66
2017	Propolis	PCA	-	67
2017	Herbs	PCA, HCA, heatmap	-	68
2018	Propolis	PCA	-	69
2018	Herbs	PCA	-	70
2018	Herbs	PCA	OPLS	71

The term chemometrics, coined by Swante Wold in the early 70's[72], refers to the science of applying data-driven means to extract information from chemical systems. By using mathematical and statistical procedures, a maximum of information can be extracted from the data. Compared with other analytical chemistry techniques, the use of chemometrics in HPTLC is new and mainly focused on herbal products and propolis (Table 1). Three particularities of HPTLC make it interesting for this application: (1) the low cost per sample, (2) parallel analysis, and (3) the image format, the analysis of which is a growing field in data science.

1.2.2.1 Preprocessing

Even if improved with HPTLC, the difference in experimental conditions from one plate to another can be critical in chromatographic systems and it highlights the need for carefully chosen preprocessing steps[62]. Smoothing allows to remove noise that comes for example from the CCD camera, this can be applied directly on the image, *e.g.* median filter or wavelet transform[45,73,74] or on the videodensitograms, *e.g.* Savitky Golay algorithm. Baseline correction is important and can help for example to reduce the differences of the derivatization reagent. Shift in retention must be corrected as most algorithms are not able to cope with it and can be done via different alignment algorithms[50,62]. Normalization depends on the algorithm and on the data, it is common for example to scale and center the data with principal component analysis (PCA) or support vector machine (SVM), while this step is not necessary with random forest (RF). For the variable selection, most authors chose to use one of the red, green and blue channels or grayscale with an hR[F] between 0 and 100. Others choose to use peak area or height after a step of peak integration made in a quantitative videodensitometry software.

1.2.2.2 Unsupervised statistics

The task for the mathematical model in unsupervised techniques is to describe datasets without prior knowledge to highlight structures and the presence of clusters. The output of those models can generally be plotted and lead to highly visual and informative figures (Fig. 6C and 6D). Even if labels are available and the final objective of the study is to apply supervised statistics, unsupervised techniques can provide insight and are generally applied first.

PCA is the most popular chemometric tools used in HPTLC (Table 1). This is a technique that allows the reduction of the dimensionality of the dataset, while retaining a maximum of information for further visualizations (Fig. 6C). The algorithm has two main outputs, *i.e.* the scores, which represent the samples in the new dimensional space, and the loadings, which are the correlation coefficients between the principal components and the original variables. Hierarchical cluster analysis (HCA) is the second most popular technique. Samples are grouped based on their similarity (Fig. 6D). A similarity metric is used to calculate the distance between samples, *e.g.* Euclidean distance. A clustering algorithm is then used to group the samples together. Other unsupervised algorithms

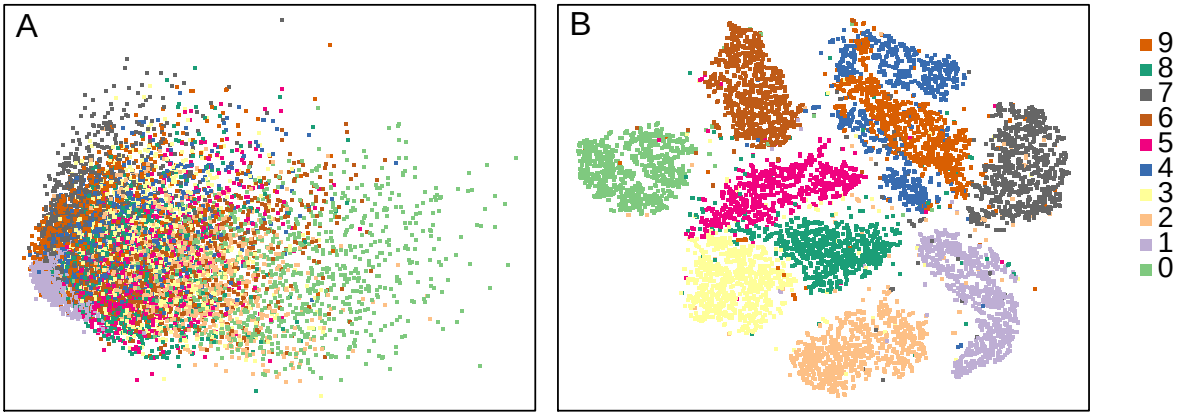


Figure 7: Dimension reduction on 10000 digits from the MNIST dataset with PCA (A) and t-SNE (B).

have been developed recently. t-Distributed Stochastic Neighbor Embedding (t-SNE) is a dimension reduction technique for high dimension datasets[75]. Figure 7 shows the visualization map obtained with PCA and t-SNE on 10,000 digits of the MNIST dataset and shows the promising capabilities of this algorithm.

1.2.2.3 Supervised statistics

In supervised techniques, the task is to generate a model to predict discrete (classification) or continuous properties (regression). It consists of the following steps: (1) split the dataset into training, validation and test set, (2) select variables and apply preprocessing, (3) build a model with the training set, while evaluating it with the validation set, and (4) perform a final evaluation with the test set.

The use of supervised techniques is less spread in HPTLC than unsupervised techniques (Table 1). Algorithms include linear discriminant analysis (LDA) based on linear discriminate functions able to maximize between-class variance and minimize within-class variance. Partial least square (PLS)[76] had been used for regression and classification, *i.e.* PLS discriminant analysis (PLS-DA) as well as orthogonal PLS discriminant analysis (OPLS-DA). PLS is similar to PCA but in a predictive fashion, the algorithm finds the components in the input matrix, which described it best and are most correlated with the target value. Classification and regression tree (CART) is based on a decision tree, for which the samples are split repeatedly between each branch depending of variable values. At each node, a decision on class belonging or quantity can be made.

A more modern algorithm called random forest (RF)[77] builds a high number of trees, typically 500, based on a subset of the data, both sample subset and variable subset, and can avoid the over-fitting often observed with CART. A support vector machine (SVM)[78] finds optimal hyperplane boundary in the high dimensional space to separate the different classes.

Artificial neural networks (ANNs) are structures consisting of interconnected processing elements called artificial neurons. Those artificial neurons, organized in layers, are able to process data and represent knowledge. ANNs are trained iteratively to calculate an output layer given an input layer, while crossing one or more hiddens layers. With the increasing processing power and data access, more complex ANN can be trained. By stacking several hidden layers, so called deep learning[79], the network is able to learn more abstract features and give state of the art results on tasks like image classification[80] and machine translation[81]. ANNs can also be used for unsupervised learning by training them to reproduce their input[82,83], *e.g.* auto encoder[84] and restricted Botzmann machine[85].

Validation of the resulting model is highly critical in supervised analysis to avoid over-fitting and is often overlooked. During the development phase, the validation set is used to select the best preprocessing, model algorithm and parameters. Cross validation can also be used, either k-fold, leave-one-out or bootstraped cross validation[86]. Final judgment of the performance of the model is made with the test set. It is mandatory and must not be evaluated until the end of the development.

Application of chemometrics in HPTLC is a growing field. However, there is no common software platform for this analysis. The step of videodensitogram extraction in particular is performed most of the time in a non-dedicated software, *e.g.* ImageJ before being exported to be further analyzed, *e.g.* in Matlab (MathWorks, Natick, MA, USA). This leads to time-consuming and error prone workflows, incompatible with the ambition of a streamlined, reproducible analysis.

1.2.3 Hyphenation with mass spectrometry

Unlike GC and HPLC, the elution head based hyphenation of mass spectrometry and HPTLC made a big step forward in 2004[87] with the direct elution and transfer of TLC zones into an electrospray mass spectrometer (Fig. 8A). This new approach replaced

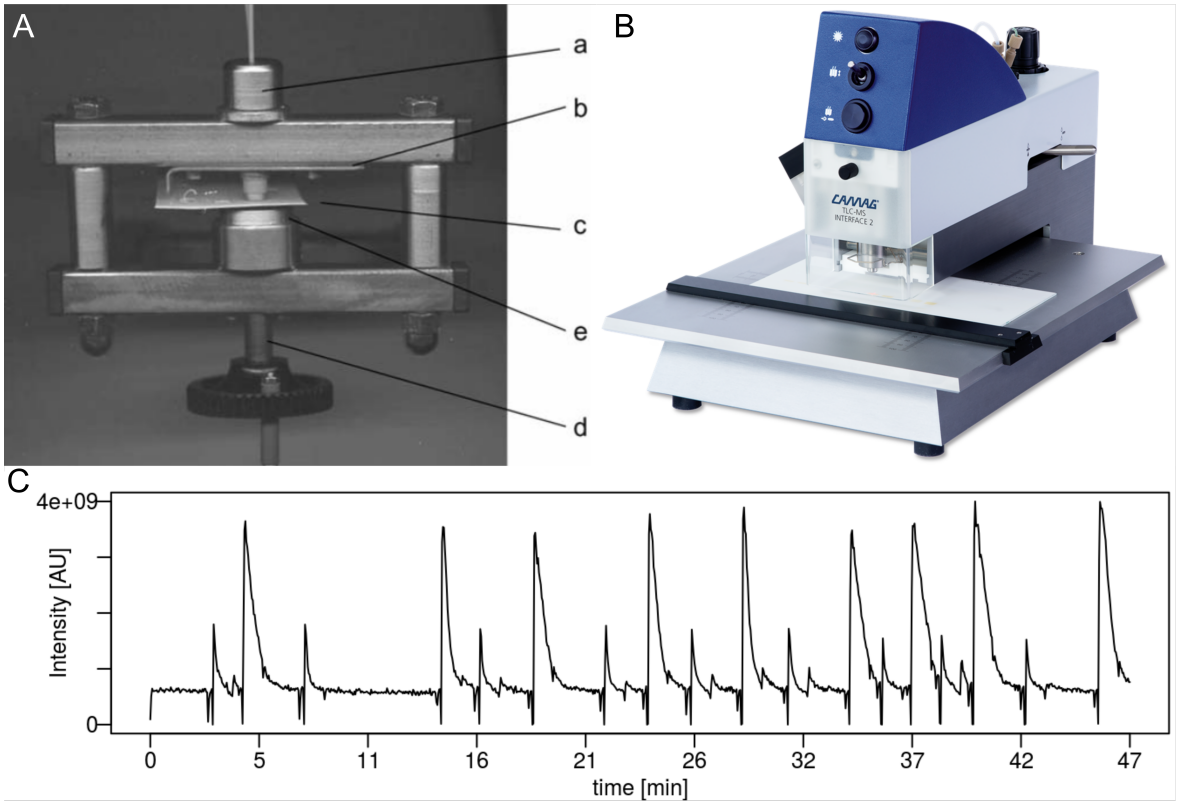


Figure 8: First hyphenation apparatus (A: Plunger (a) is held in place by the locking device (b). The TLC plate (c) is positioned between the upper plunger (a) and the lower plunger (e), which can be moved by the screw (d)), state of the art commercial apparatus (B) and typical total ion current of an HPTLC-MS experiment (C). Reprinted with permission from Anal. Bioanal. Chem. 378 (2004) 964, Copyright © 2004, Springer Nature.

the manual scrapping off the adsorbent of the plates. This apparatus was modified for usage on glass plates[88–91]. A commercial instrument was first available in 2009 (Fig. 8B).

The hyphenation of HPTLC with high-resolution mass spectrometry (HRMS) produced a large amount of data during the analysis, making data analysis tools of primary importance. With GC and HPLC, those tools were introduced as soon as the hyphenation was possible[92]. The need for software solutions in mass spectrometry went together with the evolution of the instrumentation. New solutions were presented frequently, often dedicated to a specific use case. In 2017, the THEMIS software[93] was focused on petroleum products analyzed by ultra HRMS to separate noise from signal in big data sets. The same year, the R package AssayR was released to process data from a metabolic sample

analyzed via high resolution widescan liquid chromatography hyphenated to HRMS. Other solutions automatized a full analytical pipeline, *e.g.* enviMass[94], XCMS[95] or Thermo TraceFinder (Thermo Fisher Scientific, Schwerte, Germany). However, those open-source or commercial softwares were either dedicated to a specific type of samples or only applicable to HPLC or GC and therefore not directly usable for HPTLC-HRMS.

HPTLC-HRMS presents several specificities. The nature of the signal itself is different with a chemical background specific to the open planar format[96] which can cause ion suppression and a reduction of the signal intensity. A few of those background signals can be eliminated by plate washing[97] or the use of special plates for mass spectrometry commercialized by Merck. The use of a column between the elution head and the mass spectrometer was also tested[91,98,99]. Other specificities of HPTLC-HRMS are intrinsic to the technique and required special handling of the data, *e.g.* sample and rinsing elution from a same plate are present in a single data file, desorption of the analyte leads to non Gaussian peaks and coeluted compounds of a single zone are eluted together (Fig. 8C). Those particularities of HPTLC-HRMS showed that there is room for a dedicated solution to analyze the data produced by this technique.

1.2.4 The R programming language

R is a programming language focused on statistical computing. It first appeared in 1993[100] and is now widely used in academia and industry to develop statistical software and data analysis tools. One of the strengths of R comes from the code collection created by its community in the form of packages. Several R packages are dedicated to chemometrics, *e.g.* MS data reading[101], preprocessing[102,103], data analysis[104], peak alignment[105,106] and peak detection[107]. Outside of chemistry, several packages are worth mentioning. The caret package is a warper for the supervised learning algorithms available in R[108]. All of those algorithms can then be used from a common interface along with useful tools for optimization and visualization. The shiny package permits to create web applications without knowledge of web development[109], making it easy to create data analysis solutions accessible from every computer connected to the Internet.

1.3 Aim of this work

The objective of this work was to exploit recent advances in 3D printing and data science to develop tools for planar chromatography. Those open-source tools should be an important source of inspiration and should permit to unlock previously encountered problems as well as opening new avenues for the analyst.

It was shown that instrumentation in planar chromatography did not evolve in a similar pace as for HPLC and GC. An answer to this need was given by the OC concept. However, the modification of commercial and close-source systems did not allow the full control necessary and significantly impaired its spreading in the laboratory. Repurposing the low-cost environment of 3D printers, it should give this control. This strategy should first be validated for printing of silica gel layers. Then, a more ambitious full system should be developed, able to perform several steps of the analysis.

In analytical chemistry, data analysis can be a time consuming, nevertheless indispensable step. The analyst is always in need for tailored, fast and easy to use software. The R programing language had grown notably in the past years, and with it, the data science field in general. Numerous tools and algorithms had been developed and are yet to be integrated in the analyst's toolbox.

1.4 References

(1) Sherma, J.; Morlock, G. *J. Planar Chromatogr. - Mod. TLC* **2008**, *21* (6), 471–477.

(2) Kaiser, R.; Zlatkis, A. *High Performance Thin-layer Chromatography*; Elsevier Inc.: Amsterdam, 1977.

(3) Hauck, H. E.; Schulz, M. *J. Chromatogr. Sci.* **2002**, *40* (December), 1–3.

(4) Hauck, H. E.; Schulz, M. *Chromatographia* **2003**, *57* (SUPPL.), S/313–S/315.

(5) Frolova, A. M.; Chukhlieb, M. A.; Drobot, A. V.; Kryshtal, A. P.; Loginova, L. P.; Boichenko, A. P. *Open Surf. Sci. J.* **2009**, *1*, 40–45.

(6) Clark, J. E.; Olesik, S. V. *Anal. Chem.* **2009**, *81* (10), 4121–4129.

(7) Bezuidenhout, L. W.; Brett, M. J. *J. Chromatogr. A* **2008**, *1183* (1-2), 179–185.

(8) Song, J.; Jensen, D. S.; Hutchison, D. N.; Turner, B.; Wood, T.; Dadson, A.; Vail, M. A.; Linford, M. R.; Vanfleet, R. R.; Davis, R. C. *Adv. Funct. Mater.* **2011**, *21* (6), 1132–1139.

(9) Morlock, G. E.; Oellig, C.; Bezuldenhout, L. W.; Brett, M. J.; Schwack, W.; Powell, S. C.; Morlock, G. E.; Oellig, C.; Bezuidenhout, L. W.; Brett, M. J.; Schwack, W.; Bezuldenhout, L. W.; Brett, M. J.; Schwack, W. *Anal. Chem.* **2010**, *82* (7), 2940–2946.

(10) Morlock, G.; Stiefel, C.; Schwack, W. *J. Liq. Chromatogr. Rel. Technol.* **2007**, *30* (15), 2171–2184.

(11) Häbe, T. T.; Morlock, G. E. *J. Chromatogr. A* **2015**, *1413*, 127–134.

(12) Rayleigh, F. *Proc. London Math. Soc.* **1878**, *S1-10*, 4–13.

(13) Rune, E. Measuring instrument of the recording type, 1951.

(14) Le, H. P. *J. Imaging Sci. Technol.* **1998**, *42*, 49–62.

(15) Li, J.; Rossignol, F.; Macdonald, J. *Lab Chip* **2015**, *15* (12), 2538–2558.

(16) Boland, T.; Tao, X.; Damon, B. J.; Manley, B.; Kesari, P.; Jalota, S.; Bhaduri, S. *Mater. Sci. Eng. C* **2007**, *27* (3), 372–376.

(17) De Maria, C.; Ferrari, L.; Montemurro, F.; Vozzi, F.; Guerrazzi, I.; Boland, T.;

Vozzi, G. *Procedia Eng.* **2015**, *110*, 98–105.

(18) Hull, C. W. Apparatus for production of three-dimensional objects by stereolithography, 1986.

(19) Gross, B. C.; Erkal, J. L.; Lockwood, S. Y.; Chen, C.; Spence, D. M. *Anal. Chem.* **2014**, *86* (7), 3240–3253.

(20) Gross, B.; Lockwood, S. Y.; Spence, D. M. *Anal. Chem.* **2017**, *89* (1), 57–70.

(21) Tycova, A.; Prikryl, J.; Foret, F. *Electrophoresis* **2016**, *37* (7-8), 924–930.

(22) Salentijn, G. I.; Permentier, H. P.; Verpoorte, E. *Anal. Chem.* **2014**, *86* (23), 11657–11665.

(23) Duarte, L. C.; Colletes de Carvalho, T.; Lobo-Júnior, E. O.; Abdelnur, P. V. V.; Vaz, B. G.; Coltro, W. K. T. *Anal. Methods* **2015**, *8* (3), 496–503.

(24) Martínez-Jarquín, S.; Moreno-Pedraza, A.; Guillén-Alonso, H.; Winkler, R. *Anal. Chem.* **2016**, *88* (14), 6976–6980.

(25) Mathieson, J. S.; Rosnes, M. H.; Sans, V.; Kitson, P. J.; Cronin, L. *Beilstein J. Nanotechnol.* **2013**, *4* (1), 285–291.

(26) Scotti, G.; Nilsson, S. M. E.; Haapala, M.; Pöhö, P.; Boije, G.; Yli-kauhaluoma, J.; Kotiaho, T. *React. Chem. Eng.* **2017**, *2*, 299–303.

(27) Sandron, S.; Heery, B.; Gupta, V.; Collins, D. a; Nesterenko, E. P.; Nesterenko, P. N.; Talebi, M.; Beirne, S.; Thompson, F.; Wallace, G. G.; Brabazon, D.; Regan, F.; Paull, B. *The Analyst* **2014**, *139* (24), 6343.

(28) Gupta, V.; Talebi, M.; Deverell, J.; Sandron, S.; Nesterenko, P. N.; Heery, B.; Thompson, F.; Beirne, S.; Wallace, G. G.; Paull, B. *Anal. Chim. Acta* **2016**, *910*, 84–94.

(29) Gupta, V.; Beirne, S.; Nesterenko, P. N.; Paull, B. *Anal. Chem.* **2018**, *90*, 1186–1194.

(30) Fee, C.; Nawada, S.; Dimartino, S. *J. Chromatogr. A* **2014**, *1333*, 18–24.

(31) Gupta, V.; Mahbub, P.; Nesterenko, P. N.; Paull, B. *Anal. Chim. Acta* **2018**,

1005, 81–92.

(32) Prikryl, J.; Foret, F. *Anal. Chem.* **2014**, *86* (24), 11951–11956.

(33) Wilkes, T. C.; McGonigle, A. J. S.; Willmott, J. R.; Pering, T. D.; Cook, J. M. *Opt. Lett.* **2017**, *42* (21), 4323.

(34) Yu, H.; Le, H. M.; Kaale, E.; Long, K. D.; Layloff, T.; Lumetta, S. S.; Cunningham, B. T. *J. Pharma. Biomed. Anal.* **2016**, *125*, 85–93.

(35) Jones, R.; Haufe, P.; Sells, E.; Iravani, P.; Olliver, V.; Palmer, C.; Bowyer, A. *Robotica* **2011**, *29* (January 2011), 177–191.

(36) Wittbrodt, B. T.; Glover, A. G.; Laureto, J.; Anzalone, G. C.; Oppliger, D.; Irwin, J. L.; Pearce, J. M. *Mechatronics* **2013**, *23* (6), 713–726.

(37) Bernard-Savary, P.; Poole, C. F. *J. Chromatogr. A* **2015**, *1421*, 184–202.

(38) Popovic, N.; Sherma, J. *Trends Chromatogr.* **2014**, *9*, 21–28.

(39) Kowalska, I.; Cieśla, Ł.; Oniszczuk, T.; Waksmundzka-hajnos, M.; Oleszek, W.; Stochmal, A. *J. Liq. Chromatogr. Rel. Technol.* **2013**, *36* (17), 2387–2394.

(40) Soponar, F.; Moţ, A. C.; Sârbu, C. *J. Chromatogr. A* **2008**, *1188* (2), 295–300.

(41) Pavicevic, I. D. Package 'qtlc', 2016.

(42) Victoria, A.; Hess, I. *J. Chem. Educ.* **2007**, *84* (5), 842–847.

(43) Tian, R.; Xie, P.; Liu, H. *J. Chromatogr. A* **2009**, *1216* (11), 2150–2155.

(44) Wang, X. H.; Xie, P. S.; Lam, C. W. K.; Yan, Y. Z.; Yu, Q. X. *J. Pharma. Biomed. Anal.* **2009**, *49* (5), 1221–1225.

(45) Komsta, Ł.; Cieślab, Ł.; Bogucka-Kockac, A.; Józefczykd, A.; Kryszeńe, J.; Waksmundzka-Hajnos, M. *J. Chromatogr. A* **2011**, *1218* (19), 2820–2825.

(46) Sârbu, C.; Moţ, A. C. *Talanta* **2011**, *85* (2), 1112–1117.

(47) Zarzycki, P. K.; Zarzycka, M. B.; Clifton, V. L.; Adamski, J.; Glod, B. K. *J. Chromatogr. A* **2011**, *1218* (33), 5694–5704.

(48) Ogegbo, O. L.; Eyob, S.; Parmar, S.; Wanga, Z.-T.; Bligh, S. W. A. *Anal. Methods*

2012, *4* (8), 2522–2527.

(49) Agatonovic-Kustrin, S.; Loescher, C. M. *Anal. Chim. Acta* **2013**, *798*, 103–108.

(50) Audoin, C.; Holderith, S.; Romari, K.; Thomas, O. P.; Genta-Jouve, G. *J. Planar Chromatogr. - Mod. TLC* **2014**, *27* (5), 328–332.

(51) Morlock, G.; Ristivojević, P.; Chernetsova, E. S. *J. Chromatogr. A* **2014**, *1328*, 104–112.

(52) Wong, K. H.; Razmovski-Naumovski, V.; Li, K. M.; Li, G. Q.; Chan, K. *J. Pharma. Biomed. Anal.* **2014**, *95*, 11–19.

(53) Ristivojević, P.; Andrić, F.; Trifković, J.; Vovk, I.; Stanisavljević, L. Ž.; Tešić, Ž. L.; Dušanka M. Milojković-Opsenica. *J. Chemom.* **2014**, *28* (4), 301–310.

(54) Sagi, S.; Avula, B.; Wang, Y. H.; Zhao, J.; Khan, I. A. *J. Sep. Sci.* **2014**, *37* (19), 2797–2804.

(55) Tang, T. X.; Guo, W. Y.; Xu, Y.; Zhang, S. M.; Xu, X. J.; Wang, D. M.; Zhao, Z. M.; Zhu, L. P.; Yang, D. P. *Phytochem. Anal.* **2014**, *25* (3), 266–272.

(56) Pei-Shan, X.; Shuai, S.; Shunjun, X.; Longgang, G. *Chromatogr. Sep. Tech.* **2014**, *5* (6).

(57) Mavimbela, T.; Viljoen, A.; Vermaak, I. *J. Appl. Res. Med. Aromat. Plants* **2014**, *1* (1), 8–14.

(58) Agatonovic-Kustrin, S.; Morton, D. W.; Ristivojević, P. *J. Chromatogr. A* **2016**, *1468*, 228–235.

(59) Chasset, T.; Häbe, T. T.; Ristivojevic, P.; Morlock, G. E. *J. Chromatogr. A* **2016**, *1465*, 197–204.

(60) Maldini, M.; Montoro, P.; Addis, R.; Toniolo, C.; Petretto, G. L.; Foddai, M.; Nicoletti, M.; Pintore, G. *Ind. Crops Prod.* **2016**, *94*, 665–672.

(61) Milojković Opsenica, D.; Ristivojević, P.; Trifković, J.; Vovk, I.; Lušić, D.; Tešić, Ž. *J. Chromatogr. Sci.* **2016**, No. 4, 1–7.

(62) Ristivojević, P. M.; Morlock, G. E. *J. Planar Chromatogr. - Mod. TLC* **2016**, *29*

(4), 310–317.

(63) Boka, V. I.; Stathopoulou, K.; Benaki, D.; Gikas, E.; Aligiannis, N.; Mikros, E.; Skaltsounis, A. L. *Phytochem. Lett.* **2017**, *20*, 379–385.

(64) Guzelmeric, E.; Ristivojević, P.; Vovk, I.; Milojković-Opsenica, D.; Yesilada, E. *J. Pharma. Biomed. Anal.* **2017**, *132*, 35–45.

(65) Hage, S.; Morlock, G. E. *J. Chromatogr. A* **2017**, *1490*, 201–211.

(66) Ristivojević, P.; Morlock, G. E. *Food Hydrocolloids* **2017**, *64*, 78–84.

(67) Ristivojević, P.; Trifković, J.; Vovk, I.; Milojković-Opsenica, D. *Talanta* **2017**, *162* (October 2016), 72–79.

(68) Shawky, E.; Selim, D. A. *J. Chromatogr. B* **2017**, *1061-1062* (May), 134–138.

(69) Guzelmeric, E.; Ristivojević, P.; Trifković, J.; Dastan, T.; Yilmaz, O.; Cengiz, O.; Yesilada, E. *LWT–Food Sci. Technol.* **2018**, *87*, 23–32.

(70) Ibrahim, R. S.; Fathy, H. *J. Chromatogr. B* **2018**, *1080* (February), 59–63.

(71) Liu, X.; Ahlgren, S.; Korthout, H. A.; Salomé-Abarca, L. F.; Bayona, L. M.; Verpoorte, R.; Choi, Y. H. *J. Chromatogr. A* **2018**, *1532*, 198–207.

(72) Wold, S. *Chemom. Intell. Lab. Syst.* **1995**, *30* (1), 109–115.

(73) Komsta, Ł. *Anal. Chim. Acta* **2009**, *641* (1-2), 52–58.

(74) Komsta, Ł. *J. Chromatogr. A* **2009**, *1216* (12), 2548–2553.

(75) Van Der Maaten, L. J. P.; Hinton, G. E. *J. Mach. Learn. Res.* **2008**, *9*, 2579–2605.

(76) Wold, S.; Sjöström, M.; Eriksson, L. *Chemom. Intell. Lab. Syst.* **2001**, *58* (2), 109–130.

(77) Breiman, L. *Mach. Learn.* **2001**, *45* (1), 5–32.

(78) Cortes, C.; Vapnik, V. *Machine Learning* **1995**, *20* (3), 273–297.

(79) LeCun, Y.; Bengio, Y.; Hinton, G. *Nature* **2015**, *521* (7553), 436–444.

(80) Krizhevsky, A.; Sutskever, I.; Hinton, G. E. *Adv. In Neur. Inf. Proc. Syst.* **2012**, *25* (1), 1097–1105.

(81) Cho, K.; Merrienboer, B. van; Gulcehre, C.; Bahdanau, D.; Bougares, F.; Schwenk,

H.; Bengio, Y. **2014**.

(82) Freund, Y.; Haussler, D. *Adv. In Neur. Inf. Proc. Syst.* **1992**, *4* (1), 912–919.

(83) Hinton, G. E.; Salakhutdinov, R. R.; Salakhutdinov, G. E. H.; R., R. *Sciences* **2006**, *339* (March), 1095–1099.

(84) Burger, H. C.; Schuler, C. J.; Harmeling, S. *IEEE Conf. on Comp. Vis. and Pat. Recog.* **2012**, 2392–2399.

(85) Tang, Y.; Salakhutdinov, R.; Hinton, G. *IEEE Conf. on Comp. Vis. and Pat. Recog.* **2012**, 2264–2271.

(86) Berrueta, L. A.; Alonso-Salces, R. M.; Heberger, K. Supervised pattern recognition in food analysis, 2007, *1158*, 196–214.

(87) Luftmann, H. *Anal. Bioanal. Chem.* **2004**, *378* (4), 964–968.

(88) Alpmann, A.; Morlock, G. *Anal. Bioanal. Chem.* **2006**, *386*, 1543–1551.

(89) Luftmann, H.; Aranda, M.; Morlock, G. *Rapid Commun. Mass Spectrom.* **2007**, *21*, 3772–3776.

(90) Morlock, G.; Jautz, U. *J. Planar Chromatogr. - Mod. TLC* **2008**, *21* (5), 367–371.

(91) Morlock, G.; Schwack, W. *J. Chromatogr. A* **2010**, *1217* (43), 6600–6609.

(92) Hilario, M.; Kalousis, A.; Pellegrini, C.; Müller, M. *Mass Spectrom. Rev.* **2006**, *25* (3), 409–449.

(93) Gavard, R.; Rossell, D.; Spencer, S. E.; Barrow, M. P. *Anal. Chem.* **2017**, *89*, 11383–11390.

(94) Martin Loos. blosloos/enviMass: Utilities to Process Mass Spectrometry (LC-MS) Data for Environmental Trend Analysis, 2016.

(95) Smith, C. a; Want, E. J.; O'Maille, G.; Abagyan, R.; Siuzdak, G. *Anal. Chem.*

2006, *78* (3), 779–787.

(96) Morlock, G. *J. Liq. Chromatogr. Rel. Technol.* **2014**, *37* (20), 2892–2914.

(97) Glavnik, V.; Vovk, I.; Albreht, A. *J. Chromatogr. A* **2017**, *1482*, 97–108.

(98) Mroczek, T. *J. Pharma. Biomed. Anal.* **2016**, *129*, 155–162.

(99) Schwack, W.; Pellissier, E.; Morlock, G. *Anal. Bioanal. Chem.* **2018**.

(100) Ihaka, R. *Proceedings of the 30th Symposium on the Interface* **1998**, 392–396.

(101) Gibb, S. readMzXmlData: Reads Mass Spectrometry Data in mzXML Format, 2015.

(102) Stevens, A.; Ramirez Lopez, L. Package 'prospectr', 2014, 1–22.

(103) Kristian Hovde Liland; Mevik, B.-H.; Canteri, R. Package 'baseline', 2015.

(104) Wehrens, R. Chemometrics with R - Multivariate Data Analysis in the Natural Sciences and Life Sciences, 2017, 17.

(105) Gerretzen, J.; Eilers, P.; Wouters, H.; Bloemberg, T.; Wehrens, R. Package 'ptw', 2017.

(106) Giorgino, T. *J. Stat. Software* **2009**, *31* (7), 1–24.

(107) Borchers, H. W. Package 'pracma', 2018.

(108) Kuhn, M. *J. Stat.Softw.* **2008**, *28* (5), 1–26.

(109) Chang, W.; Cheng, J.; Allaire, J.; Xie, Y.; McPherson, J. Shiny, web application framework for r, 2016.

Publication 1

Proof-of-Principle of rTLC, an Open-Source Software Developed for Image Evaluation and Multivariate Analysis of Planar Chromatograms

Dimitri Fichou[1], Petar Ristivojevic[2] and Gertrud E. Morlock[1]

1. Chair of Food Science, Institute of Nutritional Science, and Interdisciplinary Research Center (IFZ), Justus Liebig University Giessen, Heinrich-Buff-Ring 26-32, 35392 Giessen, Germany

2. On leave from the Innovation Center of the Faculty of Chemistry, University of Belgrade, P.O. Box 51, 11158 Belgrade, Serbia

Published in:

Analytical chemistry (2016) 88:12494-12501

Submitted: October 12, 2016; Accepted: November 15, 2016; Published: November 15, 2016

analytical chemistry

Article

pubs.acs.org/ac

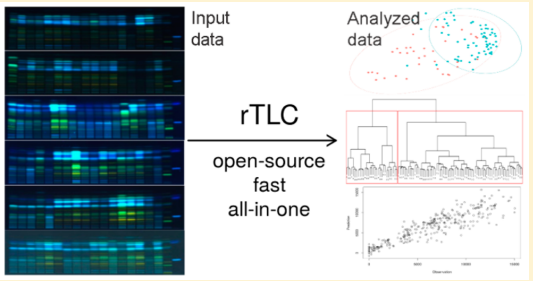
Proof-of-Principle of rTLC, an Open-Source Software Developed for Image Evaluation and Multivariate Analysis of Planar Chromatograms

Dimitri Fichou,[†] Petar Ristivojević,[†,‡] and Gertrud E. Morlock*[,†]

[†]Chair of Food Science, Institute of Nutritional Science, and Interdisciplinary Research Center (IFZ), Justus Liebig University Giessen, Heinrich-Buff-Ring 26-32, 35392 Giessen, Germany

[‡]On leave from the Innovation Center of the Faculty of Chemistry, University of Belgrade, P.O. Box 51, 11158 Belgrade, Serbia

ABSTRACT: High-performance thin-layer chromatography (HPTLC) is an advantageous analytical technique for analysis of complex samples. Combined with multivariate data analysis, it turns out to be a powerful tool for profiling of many samples in parallel. So far, chromatogram analysis has been time-consuming and required the application of at least two software packages to convert HPTLC chromatograms into a numerical data matrix. Hence, this study aimed to develop a powerful, all in one open-source software for user-friendly image processing and multivariate analysis of HPTLC chromatograms. Using the caret package for machine learning, the software was set up in the R programming language with an HTML−user interface created by the shiny package. The newly developed software, called rTLC, is deployed online, and instructions for direct use as a web application and for local installation, if required, are available on GitHub. rTLC was created especially for routine use in planar chromatography. It provides the necessary tools to guide the user in a fast protocol to the statistical data output (e.g., data extraction, preprocessing techniques, variable selection, and data analysis). rTLC offers a standardized procedure and informative visualization tools that allow the user to explore the data in a reproducible and comprehensive way. As proof-of-principle of rTLC, German propolis samples were analyzed using pattern recognition techniques, principal component analysis, hierarchic cluster analysis, and predictive techniques, such as random forest and support vector machines.



Natural extracts may contain thousands of individual compounds, and the majority of these are present in low concentrations down to the trace level. Though it is challenging, it is important to obtain reliable fingerprints that represent sound profiles of physiologically active compounds.[1] Its simplicity, cost-effective operation, and the possibility of simultaneous analysis of up to 20 samples in parallel makes high-performance thin-layer chromatography (HPTLC) a technique of choice in herbal and food analysis.[2,3] For evaluation, the HPTLC fingerprint of a complex sample is visually compared to that of a certified reference sample or to marker compounds being characteristic for the respective sample. The main disadvantage of such a manual pattern recognition technique and its visual comparison is its subjectivity, and it highly depends on the analyst's perception. Hence, hyphenation of HPTLC with highly sophisticated multivariate techniques provides objective fingerprints, mainly based on mathematical models.[4,5] As HPTLC chromatograms contain hundreds of pixels, this multidimensionality is used to extract a maximum of information out of the chromatograms.[4] For example, pattern recognition techniques can recognize chemical compound patterns, identify characteristic marker compounds as well as classify unknown samples according to their biological activity.

Though increasing, there are still a limited number of research papers on the combination of HPTLC with multivariate data analysis. Most of these are based on the investigation of propolis, herbal samples, biopolymers, and microalgae.[6−16] Although propolis is one of the most investigated honeybee products, the separation of its complex phenolic compound composition is still challenging analysts. After derivatization with Neu's reagent, PEG and detection at UV 366 nm, phenolic components showed differently colored bands. Such colorful HPTLC chromatograms are highly appropriate input data for evaluation by multivariate data analysis. There exists a wide range of derivatization reagents with different specificity and capability of detection. The resulting characteristically colored bands generate different profiles on the red, green, and blue (RGB) channels. Thus, derivatization reagents can substantially influence the separation performance and data evaluation.[6]

Received: October 12, 2016
Accepted: November 15, 2016
Published: November 15, 2016

DOI: 10.1021/acs.analchem.6b04017
Anal. Chem. 2016, 88, 12494−12501

Table 1. Overview of Publications Related to HPTLC and Multivariate Analysis

no.	samples	multivariate techniques	mainly statistical software (in several cases, not all the software used was reported)	ref
1	herbs	K-nearest neighbors, artificial neural network (ANN)	Matlab R2007 (MathWorks, Natick, MA, USA); Matlab toolbox for multivariate analysis was not reported	10
2	propolis	PCA, HCA, partial least-squares-discriminant analysis (PLS-DA)	Matlab R2011a, PLS toolbox version 6.2.1 (Eigenvector Research Incorporated, Manson, WA, USA) Image J1.48c version (Research Services Branch, National Institute of Mental Health, Bethesda, MD, USA)	6
3	herbs	PCA, PLS-DA orthogonal PLS-DA (O-PLS DA)	SIMCA-P+ Version 12 (Umetrics AB, Umea, Sweden), VideoScan (CAMAG, Muttenz, Switzerland)	12
4	propolis	PCA, HCA, LDA	Matlab R2011a, PLS toolbox version 6.2.1, SPSS Version 21 (IBM Corporation, Armonk, NY, USA), LIBSVM Version 3.16[29]	7
5	herbs	K-nearest neighbors classification and regression tree (CART), successive projection algorithm-linear discriminant analysis (SPA-LDA), PCA-discriminant analysis (PCA-DA), support vector machine-discriminant analysis (SVM-DA), PLS-DA	Matlab R2012b, PLS toolbox version 7.3.1, SPA toolbox 1.0 (self-developed programs written in Matlab), Classification toolbox version 2.0 (Milano Chemometrics and QSAR Research Group, Milano, Italy)	21
6	herbs	PCA	XLSTAT (Addinsoft, New York, NY, USA); densitogram extraction from winCATS (CAMAG)	9
7	herbs	PCA	Origin pro (OriginLab, Northampton, MA, USA); self-developed software for video densitogram extraction was not reported	13
8	propolis	PCA, HCA	TLC Analyzer;[30] software for HCA and PCA was not reported	8
9	propolis	similarity analysis, HCA K-means clustering, ANN, SVM	Self-programmed software Xe2 IDE (Embarcadero, San Francisco, CA, USA), SPSS Version 21, LIBSVM version 3.16[29]	11
10	biopolymers	PCA, HCA	Matlab R2011a, PLS toolbox version 6.2.1, Image J1.48c version	16
11	microalgae	PCA	Matlab, PLS toolbox	31

Contrary to other chromatography techniques, such as high-performance liquid chromatography (HPLC) and gas chromatography (GC), which offer a direct export of data for further multivariate analysis, images of HPTLC chromatograms must first be converted to a numerical data matrix. Various software, toolboxes, and algorithms have been applied for image processing and multivariate analysis of HPTLC chromatograms so far (Table 1). Such packages lack in domain-specific functionality, which results in a manual, lumbering, and time-consuming pipeline of the data handling. The user is forced to open, process, and save the data through different software packages and toolboxes to perform the analysis.[6−16]

For the first time, we describe and introduce rTLC in this study. It is a newly developed open-source web application for image processing and multivariate analysis of HPTLC chromatograms. The focus is laid on the different possibilities and advantages of the application, such as a fast and simple image processing workflow and application of a range of chemometric techniques suited for planar chromatography. One driving force for developing rTLC was to provide users with a unique solution to analyze HPTLC data. The access to a simple and accurate open-source web application, instead of purchasing a number of licenses, was another impetus. Many useful features for the analysis of HPTLC data, such as line profile of target compounds, band comparison, signal preprocessing as well as comma separated value (CSV) export for analysis on other platforms were integrated. Pattern recognition techniques such as principal component analysis (PCA), hierarchical cluster analysis (HCA), and heat map are applicable on separate channels (RGB and grayscale) or in combination. Prediction techniques such as random forest (RF), linear discriminant analysis (LDA), support vector machine (SVM), partial least-squares (PLS), and classification and regression tree (CART) analysis were integrated as well. The increasing number of publications in the field of planar chromatography hyphenated with multivariate analysis motivated to redesign software and add many new tools. This makes rTLC suitable for a wide range of applications in herbal, food, and environmental science.

EXPERIMENTAL SECTION

Setup of the Open-Source Web Application. The rTLC application is written with the R programming language.[17] R is an open-source language and environment for statistical computing and graphics. A key feature of R lies in its community of sharing users, who contribute to the extension of the language via packages, allowing others to use their work. rTLC uses in particular the shiny package to create an HTML based user interface[18] and the caret package for machine learning.[19] This way, the application was deployed online and is directly accessible via a modern Internet browser with an Internet connection. Because it is a web application, the user is not required to install software. Direct use of rTLC (http://shinyapps.ernaehrung.uni-giessen.de/rtlc) and, in cases required, instructions for local installation are available on GitHub: https://github.com/DimitriF/rTLC-apps.

Example Data Set. A given sample set was used as proof-of-principle of the newly developed software. Samples (n = 106) of German propolis obtained from the Apicultural State Institute (Stuttgart, Germany) were analyzed in a previous study.[20,21] The resulting 7 chromatograms in the JPEG format were manually labeled before the statistical analysis, leading to the assignment of 37 blue-type and 69 orange-type samples of German propolis. The rTLC parameters set are discussed subsequently.

RESULTS AND DISCUSSION

rTLC, the newly developed open-source web application for image processing and multivariate analysis of HPTLC chromatograms, is introduced for the first time. The simple and streamlined workflow (Figure 1) provides the necessary tools to reproducibly guide the user in a fast protocol to the statistical data output. For regular cases, the evaluation of HPTLC chromatograms took only a few minutes. The proof-of-principle was demonstrated via a German propolis data set, which was also made available as demonstration file in the rTLC software. Thus, the user is able to follow and reproduce the results reported below.

DOI: 10.1021/acs.analchem.6b04017
Anal. Chem. 2016, 88, 12494−12501

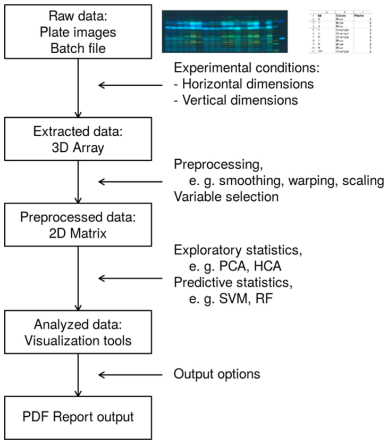


Figure 1. Workflow of the newly developed rTLC software performed within a few minutes for regular cases.

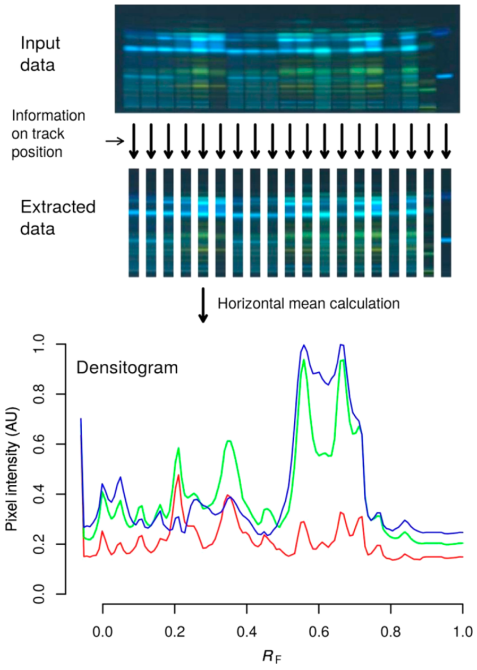


Figure 2. Processing of the experimental parameters for extraction of the HPTLC chromatograms to obtain the HPTLC densitograms.

Data Input. The user had to upload two files in rTLC to provide an appropriate data set for image evaluation and multivariate data analysis: (1) HPTLC chromatograms which contain the independent variables and (2) a batch file which contains the dependent variables about each sample on the plates, such as classes, botanical and geographical origin. Information on the experimental conditions is necessary to automatically extract each chromatogram from the HPTLC plate (e.g., the distances used during sample application and chromatography). rTLC supports the upload of commonly used image formats such as JPEG, TIFF, and PNG. The software computes the horizontal mean for each pixel of the chromatogram on the RGB channels as well as the grayscale, which is the mean of those three channels. At the end of this step, the data are in the form of a 3D array with samples as rows, R_F as columns and channels as layers (Figure 2). rTLC provides tools for line profiles of target compounds, comparison between tracks, pattern identification as well as identification of characteristic chemical and biological markers. The profile comparison of RGB channels as well as grayscale helps to find similarities and dissimilarities between samples before and after signal preprocessing.

Data Preprocessing. Recently, preprocessing methods used in HPTLC fingerprinting were discussed.[22] Among others, the appearance of a nonhomogeneous background after derivatization, an increased noise level, and band shifts are caused by variation in mobile and vapor phase composition, humidity, temperature, operator handling, and instrumental instability. Thus, warping techniques are recommended to mitigate such experimental drawbacks.[4,23] Two peak alignment procedures were integrated into the rTLC software and are available to correct inter- and intraplate band shifts:[24] (1) parametric time warping and (2) dynamic time warping. Further integrated options for data preprocessing such as denoising, normalization, and baseline removal aimed at improving the quality of the data set. The software provides the Savitzky−Golay and median filter, which are denoising/filtering methods commonly used in preprocessing of HPTLC chromatograms.[4,5] The baseline removal process was found to be mandatory in almost all cases,[4] whereas good statistical models were also obtained without baseline correction.[7] Hence, it is recommended to compare results with and without baseline correction. Also, a normalization step is not mandatory, and there is no consensus when it is obligatory—sometimes it makes the results better, sometimes even worse. The preferred method of signal normalization is the standard normal variate (SNV) method. Finally, rTLC provides autoscaling and mean centering to transform variables in the same unit.[5,6,24] The selection and need for preprocessing tools depends on the project and may be chosen by the users to obtain ready-to-use data for statistical analysis.

Variable Selection. HPTLC chromatograms provide a high number of variables for the given, often limited number of available samples. There are several approaches regarding the nature of used variables for multivariate analysis. Important variables that contain information for the aimed classification should be kept, whereas variables encoding the noise and/or with no discriminating power should be removed.[25] For this purpose, rTLC provides options for careful selection of variables for a specific channel or all channels together. The statistical analysis part also informs on this selection, which can be optimized to keep the important information only.

Exploratory Statistics. The user is only working with a data matrix, i.e. with samples as rows and variables as columns; with this, it is possible to compute pattern recognition techniques such as PCA, HCA and heatmap. For each of these techniques, informative visualization tools are available that illustrate the data in various perspectives and allow the user to highlight patterns by comparing the results with a chosen column of the batch file. For both, beginners and experienced R-users, an editor is available and can be used for other types of techniques or custom-made plots.

DOI: 10.1021/acs.analchem.6b04017
Anal. Chem. 2016, 88, 12494−12501

Predictive Statistics. With the same matrix as mentioned before, this feature allows the user to train a predictive model, used for the subsequent prediction of the properties of new samples. There are two main techniques in predictive statistics (i.e., classification and regression); both are available in the software. Before the training, the data set is split into training and test set to produce a true validation set and avoid overfitting. The application uses the caret package[19] of the R language to tune a model and choose the optimal parameters for a given algorithm. The available predictive techniques are LDA, PCA (regression only), PLS (regression only), RF, CART as well as SVM with linear and polynomial kernel.

A model will be trained for each value of a grid, automatically created but editable, and the parameters which give the best validation result will be kept for the final model. The choice of the best set of parameters is made according to a cross-validation procedure; available procedures are k-fold cross validation, bootstrapping, and leave-one-out-cross-validation. A summary metric must be chosen to select the best model. For regression, the statistical parameters can be expressed by root mean squared error (RMSE) or R^2. For classification, accuracy, kappa, sensitivity, or sensibility are available as summary metrics.

Different output tools are available to explore the result, such as confusion matrix of the test set, prediction table, and model summary. Also here, an editor is available to produce custom-made plots. At the end of this step, a model file can be downloaded and used in other sessions to predict the properties of new samples.

Proof-of-Principle of rTLC. HPTLC chromatograms contain comprehensive information regarding the polarity, chemical, and spectral properties of individual compounds in a sample. As a case study, HPTLC chromatograms of German propolis samples were used to illustrate the practical application of the rTLC software. The HPTLC chromatograms of propolis showed a complex mixture of phenolic compounds and, thus, were highly appropriate input data to demonstrate the performance and power of rTLC. Visual comparison of the respective HPTLC chromatograms and RGB channels (Figure 3A) revealed a difference in the chemical composition of the two types of German propolis. The blue type of propolis had several blue bands at R_F values around 0.2, 0.3, and 0.6 (Figure 3B). The orange type of propolis showed a rich phenolic profile and contained characteristic orange and yellow bands in the R_F range of 0.1−0.5 and high fluorescent blue bands in the R_F range of 0.5−0.8 (Figure 3C). Next, two unsupervised techniques (PCA and HCA) and two supervised techniques (RF and SVM) were selected to illustrate the capabilities of multivariate analysis by rTLC. Parametric time warping (aligned to the first sample), SNV, and mean centering were used as preprocessing steps.

Unsupervised Techniques. Commonly used pattern recognition techniques,[5,6] such as PCA and HCA, are performed by rTLC in a fast and simple way. PCA was applied on the data set for the RGB channels as well as on the grayscale image. The variable of interest was class as color assignment (labeling and different symbols were not chosen). The blue channel (Figure 4A) and grayscale data (Figure 4D) with a R_F range of 0−1 as variable selection showed the best discrimination between the two sorts of German propolis samples and their statistical performances were discussed subsequently.

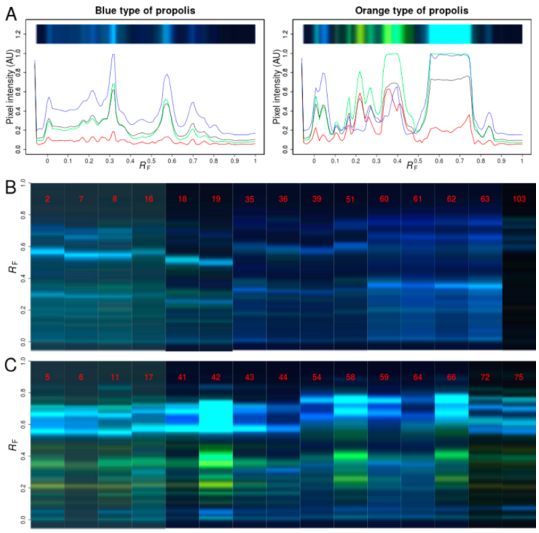


Figure 3. RGB channels (A) and HPTLC chromatograms of the phenolic profiles of the blue-type (B) and orange-type (C) German propolis samples.

In case of the blue channel data, PCA resulted in a five-component model, explaining 78.41% of the total variance. PC1 described 40.99%, while PC2 explained 15.34% of the total variance (Figure 4A). The most influential phenolic compounds were identified using the loading plots. For PC1, the compounds at R_F 0.04, 0.38, 0.53, 0.66, and 0.98 had positive contributions while the compounds at R_F 0.29, 0.58, and 0.77 had negative contributions (Figure 4B). For PC2, the compounds at R_F 0.27, 0.52, 0.63, and 0.82 had positive contributions, while the compounds at R_F 0.06, 0.30, 0.36, 0.56, and 0.72 had negative contributions (Figure 4C).

In the case of the grayscale image, the total variance explained by the first three PCs was 59.66% (PC1:32.58%, PC2:15.61%, and PC3:11.45%) (Figure 4D). The discrimination between the two types of propolis samples is mainly driven by the first component. For PC1, positive influences were found at R_F 0.06, 0.34, 0.39, 0.53, and 0.66 and negative ones at R_F 0.28, 0.58, and 0.79 (Figure 4E). For PC2, positive influences were observed at R_F 0.04, 0.36, 0.57, and 0.73 and negative ones at R_F 0.27, 0.52, 0.65, and 0.84 (Figure 4F). Once those influential R_F values are known, the researcher can apply structure elucidating techniques or refer to the literature to identify such discriminatory compounds.

Cluster analysis is an often used classification technique. This algorithm performs a hierarchical cluster analysis using the distance between samples. At the beginning, each sample is assigned to its own cluster, iteratively, the closest clusters are joined together, and the distances between clusters are recomputed, continuing until there is only one cluster. The simplest and most intuitive way to mathematically define the similarity between objects is based on the Euclidean distance. rTLC provides several routes to define the similarity between objects. According to the blue channel and grayscale, there was a good discrimination between the orange- and blue-type propolis samples, which was in agreement with PCA.[5−7] For the blue channel, "class bis" (x-labeling and color), Euclidean

DOI: 10.1021/acs.analchem.6b04017
Anal. Chem. 2016, 88, 12494−12501

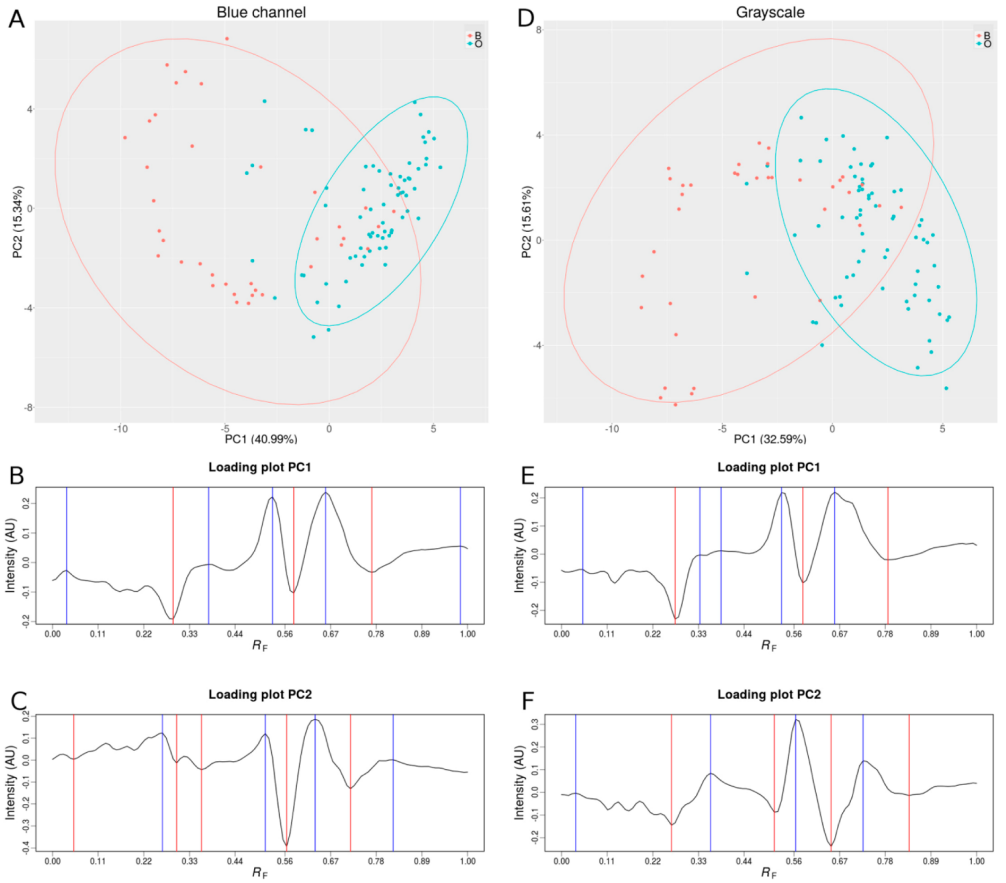


Figure 4. PC scores (A and D) and loading plots according to the blue channel (B and C) and grayscale image (E and F) evaluation.

distance, ward method, and a cluster number of 3 (blue channel) or 2 (grayscale) were chosen.

In the dendrogram of the blue channel data (Figure 5A), the first cluster out of three had a distance of 49 and was mainly composed of orange samples, whereas the second cluster had a distance of 25 and was dominated of blue-type propolis samples. The third cluster had a distance of 28 and consisted mainly of the latter samples. The dendrogram obtained for the grayscale data showed two clusters (Figure 5B). The first cluster contained almost all blue samples, and the second cluster consisted mostly of orange samples, with distances of 57 and 62, respectively. The few blue-type propolis samples grouped into the orange-type cluster differed in their patterns compared to the other blue-type propolis samples. These samples can be considered as a mixture of both types of propolis due to the natural variation in the chemical composition. For such cases, it has to be proven that the variations in the experimental condition had been removed during the preprocessing step, as far as possible.

Supervised Techniques. In supervised techniques, a set of data describing objects of known features is used to construct a training set that is used to predict those features for new samples then. Supervised techniques were applied in a wide range of chromatographic, spectrophotometric, and sensorial data, for quantification, fingerprinting, authentication, and detection of adulteration of food and herbal products.[25] The feature can be discrete, like the geographical or botanical origin, or continuous like the concentration of a target molecule in the investigated samples.

As a first step of the supervised procedure, the data were split between training and test set. Second, preprocessing techniques were applied on the training and test set. Note that for normalization, the mean centering and standard deviation of the training set is used to standardize the test set to avoid overfitting.[25] After the following variable selection, prediction models were built using the training set for each row of the tuning grid and each step of the cross-validation procedure. Afterward, the best parameters were selected, and the final model was trained with those parameters on the entire training set. Lastly, the reliability of the model was evaluated using the test set. Two powerful supervised algorithms were selected to present this feature: RF and SVM with linear kernel. Like for PCA and HCA, the following preprocessing was used: parametric time warping, SNV, and mean centering. In each case, the ratio of training to test set was 3:1 and the cross-validation method was 5-fold cross validation with total accuracy as summary metric of choice for the selection of the best model. The outcome was studied for each of the three

DOI: 10.1021/acs.analchem.6b04017
Anal. Chem. 2016, 88, 12494−12501

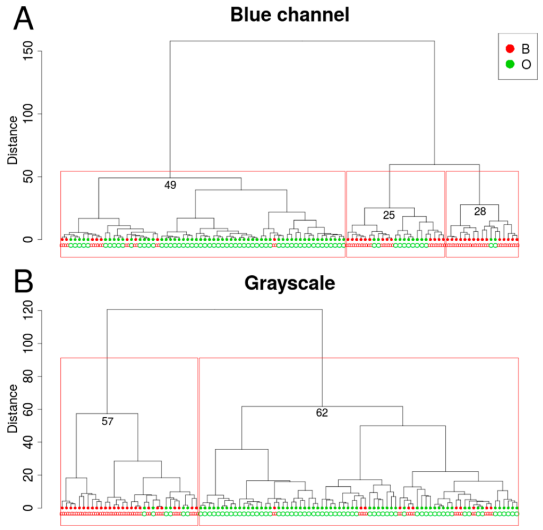


Figure 5. Dendrograms for blue channel (A) and grayscale (B) image evaluation of the German propolis samples.

channels and the grayscale image. In all cases, the prediction efficiency was high and demonstrated the power of the technique to reproduce human decisions.

Though RF has rarely been used as multivariate tool in food and herbal research so far, there are several benefits that could make the RF algorithm an appropriate supervised tool in HPTLC analysis: it can be used (1) when there are much more variables than observations, (2) for two- or multiclassification, and (3) for a good predictive performance, even when most of the predictive variables are noise, and thus, a preselection of variables is not required. As another benefit, this algorithm does not need standardization. The RF classifier needs optimization for two parameters to generate a prediction model: the number of classification trees desired (ntree) and the number of variables (mtry) which are used for tree growing in each tree. The accepted default values for those two parameters are 500 for ntree and sqrt (mall) for mtry, whereby mall is the total number of variables in the original data set. The most important parameter, mtry, can be optimized with the caret package, in contrast to the ntree parameter. This optimization led to more accurate models.[26−28] By the way, the option PLS resulted in an equivalent outcome to RF and SVM.

For all channels, the accuracy of classification of the training set was 100%. Those models were clearly overfitted, and this outcome must not be taken into account to judge a model. The confusion matrix was obtained for each channel on the test set and during the cross validation (Table 2 A). The green channel showed a good accuracy for cross validation and for the test set. For the blue channel data and grayscale image, the comparison between cross validation and test set showed more consistency, which was in accordance to PCA and HCA. Detailed statistical parameters for the blue channel showed the performance of the model according to different metrics (Table 2 B). The importance of the variables for the RF algorithm trained on the blue channel is evident (Figure 6). In contrast to the variables highlighted in the loading plots of the PCA, the model resulted in other variables to discriminate the two types of propolis.

The SVM algorithm separates the classes by an optimal hyperplane that maximizes the distances between classes by defining boundaries for the closest classes (support vectors) from the margins of the class. This way, SVM minimizes the training error with regard to the separation of the considered classes by using the least complex boundaries out of all possible ones. The optimal hyperplane is obtained by an interactive algorithm that minimizes an error function that contains a parameter (penalty error) to control the complexity of the model and to avoid overfitting.[24,25] Even if the results for each channel and the grayscale image (Table 3A) were comparable with the RF results, this algorithm performed slightly worse, in particular, on the cross-validation data set. The tuning step chose values of 0.25 for cost and 2 for gamma, except for the blue channel where the optimum cost was 0.5. Detailed statistical parameters for the grayscale showed the performance of the model according to different metrics (Table 3B).

Table 2. RF Algorithm Model: Confusion Matrix for the Test Set of Blue- and Orange-Type Propolis Samples as well as Cross Validation Set for Each Channel (A) and Detailed Summary Metrics for the Blue Channel on the Three Data Sets (Accuracy, Sensitivity, and Specificity; B)

A

			cross validation			test set		
channel	optimum mtry		blue-type	orange-type	accuracy	blue-type	orange-type	accuracy
red	2	blue	6	6	0.7857	14	11	0.8077
		orange	0	16		4	49	
green	15	blue	9	3	0.8571	16	9	0.8590
		orange	1	15		2	51	
blue	2	blue	9	3	0.8929	17	8	0.8333
		orange	0	16		5	48	
gray	2	blue	9	3	0.8929	16	9	0.8590
		orange	0	16		2	51	

B

channel	RF model parameters	accuracy	sensitivity	specificity
blue	training set	1.0000	1.0000	1.0000
	test set	0.8929	1.0000	0.8421
	cross validation	0.8333	0.7727	0.8571

DOI: 10.1021/acs.analchem.6b04017
Anal. Chem. 2016, 88, 12494−12501

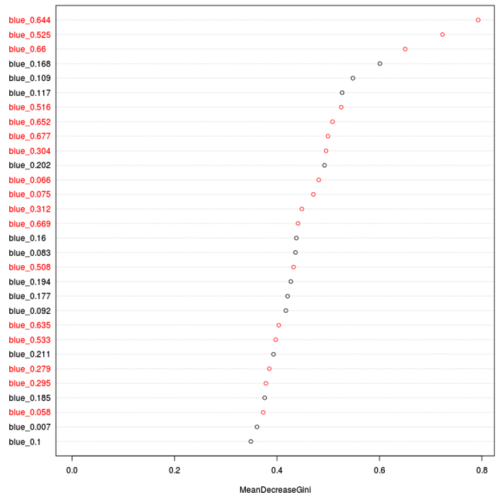


Figure 6. Variable importance for the RF algorithm model trained with the blue channel in the discrimination of orange- and blue-type propolis samples (red: variables of PCA loading plots).

CONCLUSIONS

According to our knowledge, there was no dedicated all-in-one software for a streamlined image evaluation and multivariate analysis of HPTLC chromatograms. The newly developed rTLC application was designed as user-friendly open-source software to ease fingerprint comparisons. New perspectives and conclusions on the data set are supported by a wide range of visualization tools, owed to high plotting capabilities of the R software. A great step forward was achieved by a substantial reduction of the analysis time. rTLC solved the supervised and unsupervised data handling within few minutes, whereas the current practice needs several hours using at least two different software packages. To the best of our knowledge, rTLC is the most concise tool available for application of different pattern recognition and prediction techniques for HPTLC chromatograms. On the one hand, the open-source asset of this application may attract users for the powerful combination of HPTLC and multivariate analysis. On the other hand, it may encourage the users to contribute to this technology through feedback, discussing ideas, and adding new functionalities to the software.

AUTHOR INFORMATION

Corresponding Author

*(G.E.M.) E-mail: Gertrud.Morlock@uni-giessen.de. Tel.: +49-641-99-39141. Fax: +49-641-99-39149.

ORCID

Gertrud E. Morlock: 0000-0001-9406-0351

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

Thank is owed to the Ministry of Education, Science and Technological Development of the Republic of Serbia, grant No. 172017 for financial support of P.R.

REFERENCES

(1) Joshi, D. D. *Herbal Drugs and Fingerprints*; Springer: Berlin, 2012.
(2) Krüger, S.; Morlock, G. E. In *Instrumental Thin-Layer Chromatography, Handbooks in Separation Science*, 1st ed.; Poole, C., Ed; Elsevier Science: Amsterdam, 2014; pp 409−429.
(3) Morlock, G. E.; Schwack, W. *J. Chromatogr. A* **2010**, *1217*, 6600−6609.
(4) Komsta, Ł. *Chromatogr. Res. Int.* **2012**, *2012*, ArticleNo. 893246.
(5) Milojković-Opsenica, D.; Ristivojević, P.; Andrić, F.; Trifković, J. *Chromatographia* **2013**, *76*, 1239−1247.
(6) Ristivojević, P.; Andrić, F.; Trifković, J.; Vovk, I.; Stanisavljević, L.Ž.; Tešić, Ž.L.; Milojković-Opsenica, D. *J. Chemom.* **2014**, *28*, 301−310.
(7) Morlock, G. E.; Ristivojević, P.; Chernetsova, E. S. *J. Chromatogr. A* **2014**, *1328*, 104−112.
(8) Sârbu, C.; Moţ, A. C. *Talanta* **2011**, *85*, 1112−1117.
(9) Sagi, S.; Avula, B.; Wang, Y. H.; Zhao, J.; Khan, I. A. *J. Sep. Sci.* **2014**, *37*, 2797−2804.
(10) Tian, R. T.; Xie, P. S.; Liu, H. P. *J. Chromatogr. A* **2009**, *1216*, 2150−2155.
(11) Tang, T. X.; Guo, W. Y.; Xu, Y.; Zhang, S. M.; Xu, X. J.; Wang, D. M.; Zhao, Z. M.; Zhu, L. P.; Yang, D. P. *Phytochem. Anal.* **2014**, *25*, 266−272.

Table 3. SVM Model with Linear Kernel: Confusion Matrix for the Test Set of Blue- and Orange-Type Propolis Samples and Cross-Validation Set for Each Channel (A) and Summary Metrics for the Grayscale Image Data on the Three Data Sets (Accuracy, Sensitivity, and Specificity; B)

A

	optimum			cross validation			test set		
channel	cost	gamma	type	blue-type	orange-type	accuracy	blue-type	orange-type	accuracy
red	0.25	2	blue	6	6	0.7143	19	6	0.7692
			orange	2	14		12	41	
green	0.25	2	blue	8	4	0.8214	18	7	0.8077
			orange	1	15		8	45	
blue	0.5	2	blue	9	3	0.8214	17	8	0.7564
			orange	2	14		11	42	
gray	0.25	2	blue	8	4	0.8517	18	7	0.8590
			orange	0	16		4	49	

B

channel	SVM model parameters	accuracy	sensitivity	specificity
gray	training set	1.0000	1.0000	1.0000
	test set	0.8571	1.0000	0.8000
	cross validation	0.8590	0.8182	0.8750

DOI: 10.1021/acs.analchem.6b04017
Anal. Chem. 2016, 88, 12494−12501

(12) Ogegbo, O. L.; Eyob, S.; Parmar, S.; Wang, Z. T.; Annie Bligh, S. W. *Anal. Methods* **2012**, *4*, 2522−2527.
(13) Xie, P. S.; Sun, S.; Xu, S.; Guo, L. *J. Chromatogr. Sep. Tech.* **2014**, *5*, 249−258.
(14) Zarzycki, P. K.; Zarzycka, M. B.; Clifton, V. L.; Adamski, J.; Głód, B. K. *J. Chromatogr. A* **2011**, *1218*, 5694−5704.
(15) Milojković Opsenica, D.; Ristivojević, P.; Trifković, J.; Vovk, I.; Lušić, D.; Tešić, Ž. *J. Chromatogr. Sci.* **2016**, *54*, 1077−1083.
(16) Morlock, G. E.; Ristivojević, P. *Food Hydrocolloids* **2016**, *64*, 78−84.
(17) R Core Team. *R*, a language and environment for statistical computing; R Foundation for Statistical Computing: Vienna, Austria, 2016. https://www.R-project.org/.
(18) Chang, W.; Cheng, J.; Allaire, J.; Xie, Y.; McPherson, J. *Shiny*, web application framework for r; 2016. https://CRAN.R-project.org/package=shiny.
(19) Kuhn, M. *J. Stat. Softw.* **2008**, *28*, 1−26.
(20) Kunz, N.; Scholl, I.; Schroeder, A.; Morlock, G. E. Presented at the 58th Annual convention of the Association of the German Bee Research Institutes, Berlin, Germany, March 29−31, 2011; Poster P 21.
(21) Morlock, G. E.; Scholl, I.; Kunz, N.; Schroeder, A. *CAMAG Bibliogr. Service CBS* **2013**, *111*, 13−15.
(22) Ristivojević, P.; Morlock, G. E. *J. Planar Chromatogr.–Mod. TLC* **2016**, *29*, 310−317.
(23) Wong, K. H.; Razmovski-Naumovski, V.; Li, K. M.; Li, G. Q.; Chan, K. *J. Pharm. Biomed. Anal.* **2014**, *95*, 11−19.
(24) Mohd, K. S.; Azemin, A.; Hamil, M. S. R.; Bakar, A. R. A.; Dharmaraj, S.; Hamdan, M. R.; Mohamad, H.; Mat, N.; Ismail, Z. *Asian J. Pharm. Clin. Res.* **2014**, *7*, 110−116.
(25) Berrueta, L. A.; Alonso-Salces, R. M.; Heberger, K. *J. Chromatogr. A* **2007**, *1158*, 196−214.
(26) Breiman, L. *Mach. Learn.* **2001**, *45*, 5−32.
(27) Ai, F. F.; Bin, J.; Zhang, Z. M.; Huang, J. H.; Wang, J. B.; Liang, Y. Z.; Yu, L.; Yang, Z. Y. *Food Chem.* **2014**, *143*, 472−478.
(28) Svetnik, V.; Liaw, A.; Tong, C.; Culberson, J. C.; Sheridan, R. P.; Feuston, B. P. *J. Chem. Inf. Comp. Sci.* **2003**, *43*, 1947−1958.
(29) Chang, C.-C.; Lin, C.-J. *ACM Trans. Intell. Syst. Technol.* **2011**, *2*, 1−27.
(30) Hess, A. V. I. *J. Chem. Educ.* **2007**, *84*, 842−847.
(31) Audoin, C.; Holderith, S.; Romari, K.; Thomas, P. O.; Genta-Jouve, G. J. *J. Planar Chromatogr.–Mod. TLC* **2014**, *27*, 328−332.

DOI: 10.1021/acs.analchem.6b04017
Anal. Chem. 2016, 88, 12494−12501

Publication 2

Open-Source-Based 3D Printing of Thin Silica Gel Layers in Planar Chromatography

Dimitri Fichou[1] and Gertrud E. Morlock[1]

1. Chair of Food Science, Institute of Nutritional Science, and Interdisciplinary Research Center (IFZ), Justus Liebig University Giessen, Heinrich-Buff-Ring 26-32, 35392 Giessen, Germany

Published in:

Analytical chemistry (2017) 89:2116-2122

Submitted: December 3, 2016; Accepted: January 10, 2017; Published: January 10, 2017

analytical chemistry

Article

pubs.acs.org/ac

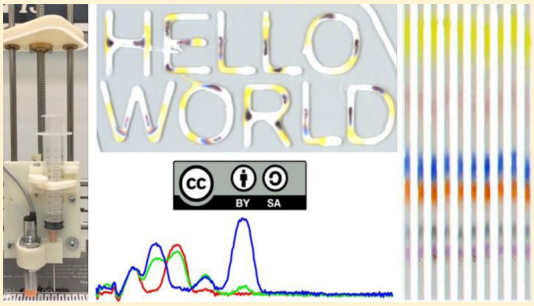
Open-Source-Based 3D Printing of Thin Silica Gel Layers in Planar Chromatography

Dimitri Fichou and Gertrud E. Morlock*

Chair of Food Sciences, Institute of Nutritional Science, and Interdisciplinary Research Center (IFZ), Justus Liebig University Giessen, Heinrich-Buff-Ring 26-32, 35392 Giessen, Germany

Supporting Information

ABSTRACT: On the basis of open-source packages, 3D printing of thin silica gel layers is demonstrated as proof-of-principle for use in planar chromatography. A slurry doser was designed to replace the plastic extruder of an open-source Prusa i3 printer. The optimal parameters for 3D printing of layers were studied, and the planar chromatographic separations on these printed layers were successfully demonstrated with a mixture of dyes. The layer printing process was fast. For printing a 0.2 mm layer on a 10 cm × 10 cm format, it took less than 5 min. It was affordable, i.e., the running costs for producing such a plate were less than 0.25 Euro and the investment costs for the modified hardware were 630 Euro. This approach demonstrated not only the potential of the 3D printing environment in planar chromatography but also opened new avenues and new perspectives for tailor-made plates, not only with regard to layer materials and their combinations (gradient plates) but also with regard to different layer shapes and patterns. As such an example, separations on a printed plane layer were compared with those obtained from a printed channeled layer. For the latter, 40 channels were printed in parallel on a 10 cm × 10 cm format for the separation of 40 samples. For producing such a channeled plate, the running costs were below 0.04 Euro and the printing process took only 2 min. All modifications of the device and software were released open-source to encourage reuse and improvements and to stimulate the users to contribute to this technology. By this proof-of-principle, another asset was demonstrated to be integrated into the Office Chromatography concept, in which all relevant steps for online miniaturized planar chromatography are performed by a single device.



Planar chromatography is an analytical technique widely used in quality control and research.[1] This planar technique substantially benefitted from improvements over the past decade,[2,3] though thin-layer chromatography (TLC) dates back to 1938.[4] Standardized layer materials were launched in the 1960s, and first, industrially manufactured precoated layers were made of irregular particles of mainly 10−15 μm with a relatively broad range of the particle size (5−20 μm).[5] With the simultaneous automation of the required devices and smaller particles (5−7 μm) with a narrower particle size distribution, high-performance thin-layer chromatography (HPTLC) was introduced in 1975.[6] With regard to adsorbent modifications, first silanized (RP2) TLC layers were introduced in 1970, RP2, RP8, and RP18 HPTLC layers in 1978, and amino, cyano, and diol phases between 1982 and 1987; layers for ion exchange and size exclusion were reported in the beginning of the 1970s.[5]

In 2001, Merck introduced ultrathin-layer chromatography (UTLC) plates with a binder-free layer made of monolithic silica gel.[7,8] In the past decade, further types of UTLC layer materials and fabrication were introduced.[9−11] Besides further monolithic silica gel layers,[12,13] these developments included, for example, electrospun polymers,[14−18] glancing angle deposition (GLAD) of silica and metal oxides,[19−24] and low-pressure chemical vapor deposition of silicon nitride onto carbon nanotube templates.[25−29]

In UTLC, the ultrathin layer (thickness mostly ≤50 μm) can easily be damaged by prevailing spray-on sample application techniques, whereas contact application has to cope with a reduced surface activity of UTLC layers, and thus, broadened start zones.[11] New sample application concepts are needed to investigate and optimize such layers, as demonstrated for the accurate printing of sample solutions via a low-cost office printer.[30,31] Not only sample application but also other required steps and their available instrumentation are oversized for miniaturized planar chromatography.

To overcome these challenges, the concept of Office Chromatography was reported in 2010.[30,11] The idea was the synergistic exploitation of innovations in print and media technology (PMT), e.g., inkjet printer or flatbed scanner, for its application in miniaturized planar chromatography. The final

Received: December 3, 2016
Accepted: January 10, 2017
Published: January 10, 2017

DOI: 10.1021/acs.analchem.6b04813
Anal. Chem. 2017, 89, 2116−2122

apparatus, following this size-reduced concept and using the ultrathin layer as core, could be a fully online system capable of performing all the required steps of the analysis, from sample application to documentation, including development and derivatization. Beyond that, a fast printing of the ultrathin layer was also imagined and included in the concept.[11] Keeping pace with all these innovations (not only in PMT) to be exploited interdisciplinarily, we consider as new potential and challenge for the planar chromatographic open system.

3D printing, also called additive manufacturing, is based on application of layers of materials one after another to obtain a 3D object.[32] Several printing methods were explored. Fused deposit modeling (FDM) was one of the mostly used techniques. Thermoplastic materials were extruded and deposed layer after layer on a stage. Two principally used plastics were polylactic acid and acrylonitrile butadiene styrene (ABS), but also others were utilized. Stereolithography was another technique based on photopolymerization, in which an UV lamp was used to polymerize a liquid resin layer after layer. With inkjet printing, a liquid binder was added on a powder bed; thereafter the stage was lowered and another layer of powder was added allowing growing the 3D shape. Selective laser sintering was similar to inkjet printing except that the powder was sintered with a laser instead of being bound with a liquid.

There are several examples of 3D printing in analytical chemistry.[32,33] It was applied for detection, i.e., in the field of mass spectrometry[34−38] and fluorescence detection.[39] It was used in the field of chromatography to print metal columns, later filled with the stationary phase,[40,41] to print a porous media column[42] and to build a cartridge for paper chromatography mass spectrometry.[36] Though 3D printing was introduced in the early 1980s,[32] it was only recently that the public realized its power.[43] Its latest success could be explained by the introduction of RepRap 3D printers in 2008.[44,45]

The RepRap project is based on self-replicating, open-source devices (http://reprap.org). As the printer without any intellectual property can reproduce its own parts to a major extent, the cost of the device was reduced to 300 Euro (up to 2000 Euro). Its open-source aspect triggered tremendous progress on both the hardware and software part of the system. Although the performance of RepRap printers sold in kits does not compete with respective industry leaders, the RepRap concept targets the users who mount the printer and configure the software by themselves. This self-made approach builds up highly valuable skills for troubleshooting and device modification.

The RepRap 3D printer is composed of several parts: (1) the computer aided design software, which designs the parts needed (exported into .STL files), (2) the slicer which transforms the .STL files into .GCODE files, controlling, for example, the layer height and dosage speed, (3) the firmware which controls the board interpreting g-code commands into machine order, (4) the electronics, including in particular the motors, the temperature sensors, the heated part, and the electronic board, and (5) the mechanical body where the movements take place.

For the first time, this study describes the modification of a low-cost open-source 3D printer into a 3D silica gel layer printer. The hardware and software parts were modified to achieve this new purpose. Advantage was taken of the existing hardware and several operational features. The proof-of-principle of the open-source 3D printing of silica gel layers was demonstrated on the example of the separation of dyes. Results on a printed plane layer were compared with those obtained from a printed channeled layer. The potential and a benchmarking of the novel technique were also discussed.

EXPERIMENTAL SECTION

Chemicals and Materials. Dye mixture III composed of 6 lipophilic dyes, i.e., Dimethyl Yellow, Oracet Red G, Solvent Blue 35, Sudan Red G, Solvent Blue 22, and Oracet Violet 2R (in descending hR_F order),[46] was purchased from CAMAG, Muttenz, Switzerland. Toluene (gradient grade), TLC and HPTLC plates silica gel 60 were obtained from Merck, Darmstadt, Germany. Bidistilled water was produced by a Heraeus Destamat Bi-18E (Thermo Fisher Scientific, Schwerte, Germany). Silica gel of a particle size of 5−25 μm, and calcium sulfate, 99%, were purchased from Sigma-Aldrich, Steinheim, Germany.

Configuration of the Firmware. The 3D printer was controlled by an Arduino Mega 2560 (https://arduino.cc) and a RAMPS 1.4 shield (http://reprap.org/wiki/RAMPS_1.4). The Arduino received signals from the computer and sent and received signals to and from the RAMPS. The RAMPS allowed controlling 5 stepper motors, 3 temperature sensors, 6 end-stop sensors, 1 heated stage, 2 extruder resistors, and 1 fan. Arduino was loaded with the Marlin firmware (http://marlinfw.org).

Printer Modification. All parts required for the printer modification are summarized (Table S-1). An open-source designed 3D printer (Prusa i3 rework, Emotiontech, Toulouse, France) was self-mounted and modified by substituting the FDM extruder by a self-made slurry doser printed using ABS plastic material (Emotiontech, Toulouse, France). Therein, a 10 mL polypropylene syringe was installed with a 1.2 × 50.0 mm needle cut horizontally at 25 mm (B. Braun, Melsungen, Germany). The syringe piston was pushed by a nema 17 stepper motor (Emotiontech, Toulouse, France) using a 8 mm trapezoidal lead screw and nut (amazon, Munich, Germany), guided by two 6 mm thick and 250 mm long smooth shafts (Emotiontech, Toulouse, France). The smooth shafts used 6 mm brass steering bushing for linear motion retrieved from an old paper printer, a modified design compatible with more common LM6UU bearing was made available. The Z-level of the slurry doser (constant distance toward the glass plate) was controlled via an inductive sensor (Emotiontech, Toulouse, France).

Pattern Design and G-code File. Layer patterns were designed with OpenSCAD (http://openscad.org) and exported into the .STL format. Several layer patterns were used, e.g., 100 mm × 100 mm × 0.2 mm and 95 mm × 95 mm × 0.2 mm (l × b × h) (Figure S-1A). Slic3r (http://slic3r.org) was used to convert them into the .GCODE file format. Important parameters set in Slic3r were (1) 60 mm/s plate movement, (2) 0.2 mm layer thickness, (3) 100% infill with rectilinear pattern and 90° filling angle to print straight lines instead of the 45° angle of the oblique ones, (4) 730% first layer extrusion width to apply 4 mL of slurry per dm^2 and 375% to apply 2 mL per dm^2 (in Slic3r, calculation for the first layer was wrong and this value was used to control it), (5) 16 mm inner diameter of the syringe, (6) infill perimeter overlap 0%, (7) 0 perimeter and 0 horizontal layer top and bottom, (8) 0.3 mm nozzle diameter of the print head, (9) adjustment of the stage shape for a maximum 100 mm × 100 mm plate format, and (10) stage temperature at 110 °C. Channel patterns were created in Slic3r

DOI: 10.1021/acs.analchem.6b04813
Anal. Chem. 2017, 89, 2116−2122

(with 30% infill and 375% first layer extrusion width) and the 95 mm × 95 mm × 0.2 mm .STL file (Figure S-1B), leading to 40 single path channels with 16 μL per track. Though many layers can be printed, only one layer was printed, as preparative layers were not intended. After computer aided design, the .STL file does not contain information on the number of layers. This number is defined in the .GCODE file by the layer thickness chosen during the slicing, and the height of the 3D object is encoded in the .STL file.

Preparation of the Slurry. The slurry was prepared immediately before use by mixing 0.88 g of silica gel, 0.12 g of calcium sulfate, and 3.00 g of water with a vortex for 1 min until the mixture was homogeneous. For printing with the above-mentioned parameters on a 10 cm × 10 cm glass plate, 4 mL of this slurry were needed per dm^2. Another approach was made for HPTLC particles. Some layers of commercially available HPTLC plates silica gel 60 were scraped off. Thereof, 1.5 g were dissolved in 4.5 g of bidistilled water. A volume of 2 mL of this slurry were needed per dm^2 (10 cm × 10 cm glass plate). Vigorous agitation with a vortex was sufficient to obtain a working HPTLC silica gel slurry.

Printing of the Layer. Layers were printed on 10 cm × 10 cm glass plates. The bare glass plates were obtained by removal of the previously used layer (Merck) and cutting to the right dimension using the smartCut Plate Cutter (CAMAG). These glass plates were cleaned with soap water, pure water, and acetone, followed by drying in a clean oven at 105 °C for 20 min. The .GCODE files were sent to the printer with Pronterface (http://pronterface.com). The freshly prepared slurry was filled into the syringe and immediately, the printing process was started. The glass plate to be printed was positioned on a stage, which was a printed circuit board (PCB) heatbed (http://reprap.org/wiki/PCB_Heatbed) supplied with the 3D printer. Simultaneously with the printing, the plate was heated at 110 °C and even 5 min longer, which enabled the drying of the printed layer. Infrared temperature measurements showed a temperature variation of ±5 °C, if compared to the temperature set. Cracking of the layer never occurred with this process. A video was made available to demonstrate the layer print (Video V-1).

TLC/HPTLC Analysis. The dye solution III was diluted 1:10 in toluene and applied as spots or 6 mm bands on the printed layer plate using the Automated TLC Sampler 4 (ATS4, CAMAG). The volume range applied was 0.5−5 μL. The distance between the tracks was 9 mm. The distance from the lower edge was 8 mm and from the side edge 25 mm. For the channeled pattern experiment, 1 μL applications were manually applied using a 5 μL glass capillary (Hirschmann, Eberstadt, Germany). Chromatography was performed in a twin-trough chamber (CAMAG) with 5 mL of toluene up to a migration distance of 70 mm. Chromatograms were documented under white light illumination (reflection and transmission mode) with an exposure time of 10 ms using the DigiStore 2 Documentation System (CAMAG). Densitograms (absorption measurement) were recorded at 550 nm (a compromise wavelength; not in the absorbance maximum of all individual dyes) with the TLC Scanner 3 (CAMAG). The measuring slit dimension was 4.0 mm × 0.3 mm and the scanning speed 20 mm/s.

RESULTS AND DISCUSSION

In a previous review on Office Chromatography, the printing of layers and even gradient layers (e.g., varying in thickness or/and layer material) was outlined.[11] In this first study on the printing of a chromatographic layer, a first proof-of-principle is given exploiting open-source hardware and software. Specific terms of 3D printing used in this study are explained in a glossary (Glossary G-1).

Building the 3D Printer for Plate Production. As a commercially available kit, an open-source designed 3D printer (Prusa i3 rework) was self-mounted and modified by substituting the FDM extruder with a self-made slurry doser. This slurry doser was self-designed in OpenSCAD and printed using ABS plastic material. The printer was switched from the FDM setup to the slurry doser setup within 10 min. The slurry doser, in which a syringe was installed for supply of the slurry (Figure 1), was connected to the X-carriage of the self-mounted printer in the same way as the FDM extruder (Figure S-2) was connected. Starting from this first modified printer and via several iterations, a working prototype was obtained (Figure 2). The designed hardware, controlled by the software, allowed an automation of the slurry application.

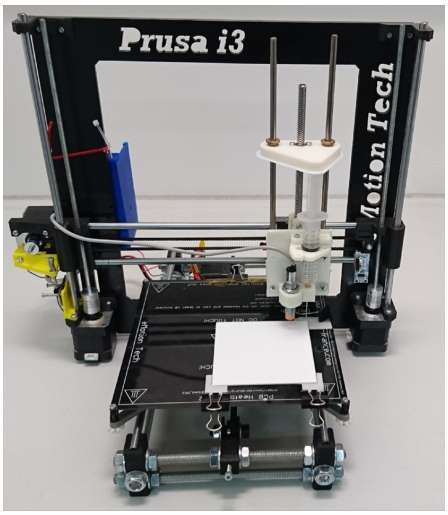


Figure 1. Modified open-source 3D printer for 3D printing of thin layers suited for planar chromatographic separation.

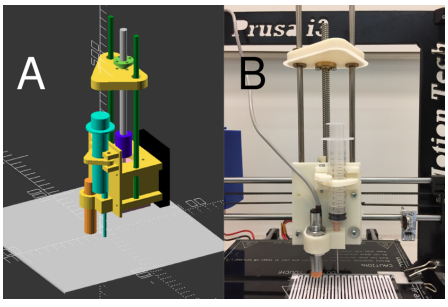


Figure 2. Self-made slurry doser as (A) computer-aided design (orange: inductive sensor; turquoise: syringe) and (B) homemade, 3D printed prototype.

DOI: 10.1021/acs.analchem.6b04813
Anal. Chem. 2017, 89, 2116−2122

Several modifications were made to configure the firmware: (1) modification of the printing volume, (2) disabling the extruder resistance and the associate security, (3) configuration of the *auto bed leveling* feature to correct the inhomogeneous level of the stage, and (4) modification of the step per millimeter for the accurate control of the amount of printed slurry. All modifications, materials, .SCAD files, .STL files, Slic3r configuration files, and the Marlin firmware files are available on GitHub (https://github.com/DimitriF/Slurry-doser). Especially such features were treasured like the possibility to correct the stage Z-level during the print, the stage heating to remove the water in the slurry during printing (without any need for an oven), the accurate Cartesian movement, the numerous options of the software and firmware for optimization. For example, an inductive sensor was self-mounted (Figure 2A, orange part). This sensor controlled the Z-level of the slurry doser during the printing process to ensure a constant distance toward the glass plate. This was found to be crucial for printing a homogeneous layer. At the beginning of each printing session, the sensor checked the position of the glass plate at three different spots of the printing area (Video S-1) and the Z-level of the slurry doser was automatically corrected by the firmware during printing. Prototyping of the slurry doser, including syringe and sensor, was fast allowing discussing, designing, printing, mounting, and testing new iterations in a few days.

Challenges of 3D Layer Printing. Several challenges were overcome during the development of this prototype for 3D printing of thin silica gel layers. Interestingly, the difficulties were similar to the ones encountered with FDM printers. For example, the plastic had to be extruded with a constant and adequate temperature (250 °C), the input filament had to be of a constant diameter (2.85 mm ±0.10 mm) and be free of contamination. Otherwise, the extruder did clog and the printing result was flawed. An inhomogeneous Z-level caused an uneven layer distribution (Figure 3A). An inhomogeneous leveling impaired the slurry printing, as the needle tip was too close to or far away from the glass surface, which resulted in an uneven silica gel surface. The accuracy of the optimal distance of needle tip to glass surface was also crucial to avoid either needle scratching on the glass plate or an inappropriate layer quality (Figure 3C). Too viscous slurries (solid-to-liquid ratio of below 1:2.5) did not achieve a good application, whereas too less concentrated slurries (solid-to-liquid ratio of above 1:3.5) resulted in a breakage of the continuous line and, thus, a drop by drop application (Figure 3B).

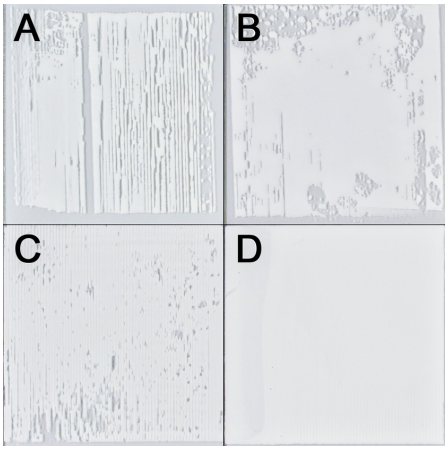


Figure 3. Illustration of some challenges of 3D printing of silica gel layers: impact on the layer quality of (A) inhomogeneous Z-level, (B) sedimentation in a too less viscous slurry and a homogeneous but too high Z-level toward the plate, (C) homogeneous Z-level but too close to the plate, and (D) homogeneous and well calibrated Z-level but sedimentation of the slurry at the end of the process (stripe on left side).

During printing, a sedimentation of the silica gel in the slurry led to an inhomogeneous distribution of the particles on the plate and, thus, an uneven silica gel surface. This effect was observed particularly for the final dosing volume of the syringe filling (Figure 3D, stripe at the left side). The sedimentation over time was overcome by a fast protocol and the instant use of the freshly prepared slurry of an optimized concentration. The impact of different binder or other additives and their optimized concentration is the focus of a further study. Further, the 3D printer had to be calibrated accurately. Motor steps per millimeter were mandatory to achieve an accurate printing process for the Cartesian movement as well as for the accuracy of the slurry volume dosed. Also, the glass plate had to be positioned evenly on the stage and at the optimal distance to the needle tip. Hence, the sensor in the self-designed slurry doser was essential. Most crucial was this precise alignment for application of the first layer to obtain a sufficient adhesion on the glass surface.

The parameters chosen during the .GCODE file creation with Slic3r were important, too. In particular, the amount of the applied slurry controlled the final layer thickness. Those were calculated on the printed plates by fitting a quadratic model on different commercially available plates of known thickness (glass plates with 50, 100, and 250 μm layers) with the mean values of the pixel intensities in the center of the plate (Figure S-3). A volume of 4 mL of slurry per 10 cm × 10 cm plate led to a layer of 0.2 mm, whereas 2 mL of slurry led to a layer of 0.1 mm (Video V-1). With the chosen setup with regard to slurry concentration, needle diameter, dosage speed, and distance, the layers printed with 4 mL of slurry were well suited for chromatography. Layers printed with 2 mL of slurry resulted in an inappropriate layer quality and showed a less robust printing process (Figure 3C).

The *auto bed leveling* feature of Marlin was precondition for a successful printing and proved the adequacy of 3D printing for layer printing. Printing of silica gel layers was more delicate than printing of the FDM plastic parts needed for modification of the printer. Layer printing aimed at the production of as thin as possible layers starting from the very first layer, whereas for FDM, the first layer could be printed thicker. For printing ultrathin layers through a more accurate positioning, a further modification of the *Z*-axis of the printer with trapezoidal screws and the use of a needle with a smaller inner diameter are the focus of another study. It should be noted that the current 1.5 version of the Prusa i3 rework is modified in this sense and compatible with the proposed layer design. Also, the slurry doser is compatible and even adaptable to other 3D printer. For the successful printing of homogeneous TLC/HPTLC layers, the *Z*-axis of the 3D printer was an essential part, as it compensated small differences in the leveling of the stage/glass plate with the auto bed leveling feature. This study was performed using a low-cost 3D printer. An apparatus with an adequately even stage level had an at least 2−3-fold price.

DOI: 10.1021/acs.analchem.6b04813
Anal. Chem. 2017, 89, 2116−2122

Additionally, the use of the 3D printing firmware, software, and electronics and their open-source aspects facilitated the implementation of the proposed process which would had taken years to develop otherwise, even though many features of the original system were not used.

Proof-of-Principle of Chromatography on Printed TLC Layers. The printed silica gel particles bound with gypsum were too loose for spray-on application and thus, the samples were applied by contact application using the ATS 4. The printed layers with the applied samples were developed with toluene. A first proof-of-principle of a successful chromatography on 3D printed layers is presented (Figure 4). The

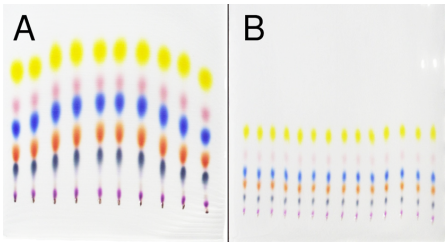


Figure 4. Effect of the development direction on the separation: (A) the diluted dye mixture (5 μL/spot applied) was developed perpendicular to the printing path direction and (B) in the direction of the printing path (1 μL/spot applied).

substance migration was more homogeneous (better repeatability of hR_F values) when the development was performed in the same direction of the printing, if compared to the perpendicular development path. This was explained by the instant heating of the layer during printing that pronounced a minor rill-like surface structure. Development parallel to the printing path was preferred, as it substantially reduced the hR_F deviation (Figure 4B).

New Potential, Shown for Channeled Patterns. The printing of layers offered new potential with regard to shape and geometry. As examples to trigger ideas for novel planar chromatographic separations, "hello world" was printed (Figure S-4) as well as a channeled pattern of 40 tracks per plate (10 cm × 10 cm, Figure 5). The latter layer shape was produced from the 95 × 95 × 0.2 mm .STL file with Slic3r and an infill of 30%, but many others can be designed. This channeled pattern already demonstrated a high sample throughput, though much higher throughputs are possible. Each track had its dedicated channel, made by a single path of the syringe needle. On such tiny paths, the separation was successful, though an optimization was obvious. Hence, new patterns were created and successfully demonstrated that were different from the classically layered surfaces in TLC/HPTLC/UTLC.

These printed channeled patterns presented several advantages over the printed flat surface pattern: (1) each sample had its own dedicated track allowing a high sample throughput and close positioning to adjacent tracks on the plate, (2) flexibility in the number of paths, pattern, and amount of slurry applied via the .STL file and the Slic3r parameters, (3) reduction of the sample costs as the slurry does not need to be applied on the whole surface, (4) a potential improvement in separation once the conditions are optimized, (6) once integrated in the Office Chromatography concept, the possibility for studying different mobile phases, migration distance, and derivatization reagents

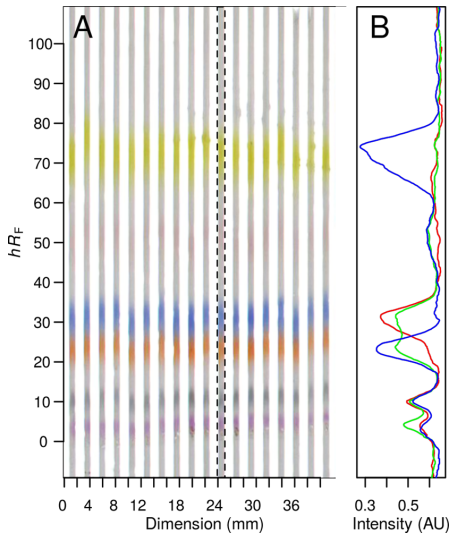


Figure 5. New potential with regard to shape and geometry: (A) As an example, chromatograms of a channeled plate and (B) video densitograms of the dashed track.

on the same plate, allowing an extensive exploration and study of a sample on one plate. This outlines that the advantages of 3D printed layers offer great potential for novel approaches and fast solutions for analytical challenges.

Chromatographic Performance on 3D Printed TLC Layers. For evaluation of the 3D printing processes, and thus, all the differently printed TLC layers, different sample volumes were applied and developed (Figure 6A). After documentation, the extracted video densitograms and recorded densitograms at 550 nm were compared. The use of different channels for extractions of the video densitograms showed a specific detection for Solvent Blue 35 and Sudan Red G using the Red and Blue channels, respectively (Figure 6B). The absorbance measurement at 550 nm showed a sufficient separation of the six dyes (resolution $R > 0.94$ between Solvent Blue 35 and Sudan Red G for the 0.5 μL application) and a flat baseline, though minor rills were evident (Figure 6C). Nevertheless, video densitometry data after baseline correction showed for example a R^2 of 0.993 for the calibration curve of Solvent blue 35. Thus, a first proof for a satisfying homogeneity of the TLC layer print was given. The TLC chromatograms and densitograms obtained on 3D printed layers gave evidence for its successful use in planar chromatography.

3D Printing of HPTLC Silica Gel Particles. A first feasibility study was made for printing of HPTLC particles. The slurry made of silica gel, which was a priori scraped off a HPTLC plate, was compatible with the proposed printing process (Figure 7). The presence of a polymer-based binder in the silica gel allowed the use of the spray-on application via the ATS 4. Defects due to few punctually aggregated particles in the slurry were visible on the chromatograms, but did not impair chromatography and, thus, the separation. At the given relative humidity of the air of below 30% that had an influence on the separation (Figure S-5), the chromatographic performance of the printed HPTLC plate (Figure 7A,B) was comparable to that of a commercially available HPTLC plate

DOI: 10.1021/acs.analchem.6b04813
Anal. Chem. 2017, 89, 2116−2122

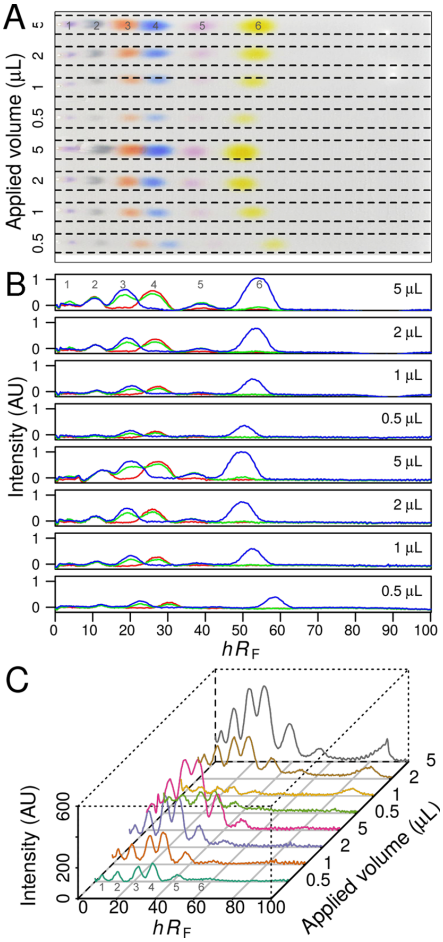


Figure 6. Example for a successful separation and semi-quantitative TLC analysis of Oracet Violet 2R (1), Solvent Blue 22 (2), Sudan Red G (3), Solvent Blue 35 (4), Oracet Red G (5), and Dimethyl Yellow (6) on a channeled plate developed in the direction of the printing path: (A) chromatograms, (B) video densitograms extracted between the dashed lines, and (C) densitograms recorded at 550 nm.

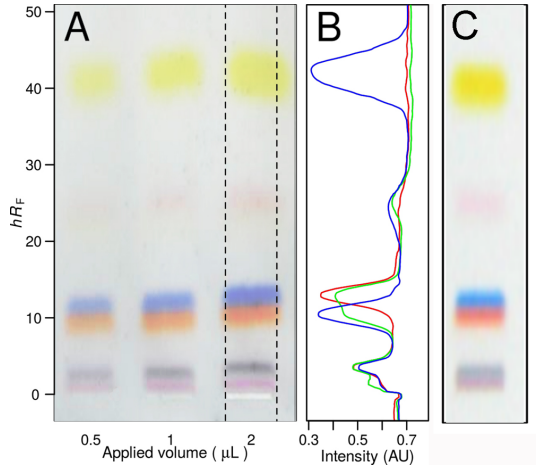


Figure 7. Comparison of chromatograms obtained from a (A) 3D printed HPTLC plate (B: respective video densitograms of the dashed track) versus (C) commercially available HPTLC plate.

silica gel 60 (Figure 7C). As HPTLC is used for quantitative studies in general, detailed figures of merits of printing HPTLC slurries are the focus of another study. There is immense potential for chromatographic improvements in the slurry materials used, a field which is at its infancy and still has to be exploited and demonstrated.

Benchmarking. In addition to the investment costs of 550 Euro for the 3D printer, the cost for the modifications were ~80 Euro. Thus, the total material costs of building the modified 3D printer for 3D printing of layers were 630 Euro (Table S-1). Material costs of producing a 0.2 mm thick plate were below 0.25 Euro and for a channeled plate of 40 tracks below 0.04 Euro (Table S-2). The printing of the plane layer last 5 min for a 100 mm × 100 mm × 0.2 mm pattern at 60 mm/s, whereas the printing of the channeled pattern (40 tracks) took only 2 min. To conclude, the fast and cost-effective 3D printing of layers based on open-source software and hardware was judged to be an exciting option.

Outlook. After this first proof-of-principle that focused the design concept and important technical aspects of the 3D printing apparatus and its modifications, there is abundant room left for optimization. The rigidity of the frame is critical to obtain ultrathin layers. This could be achieved with a smaller, stronger hardware or the use of a smaller needle diameter of 0.8 mm instead of 1.2 mm. The use of scanning electron microscopy for a more accurate measurement of the thickness and surfaces is recommended then. The proposed slurry recipe is a first start for further optimization. The impact of different polymer-based binders or other additives for layer robustness and their optimized slurry concentration are worth studying and will allow the use of spray-on application. The use of inkjet printing of sample solutions on the demonstrated gypsum-bound layers is focus of another study. The G-code preparation will also benefit from special software that is focused on the chromatographer's perspective with a better control of the applied slurry per paths and the distance in between.

CONCLUSION

The modification of a 3D printer for printing of thin silica gel layers demonstrated the potential of 3D printing of adsorbents for the analytical chemist. As the original printer and the software used were open-source, these modifications were facilitated and the resulting design was also released open-source. The printed layers were inexpensive, fast to produce and proven via the separation of a dye mixture. Different layer shapes and patterns were demonstrated and the inherent potential is still at its infancy. Although the use of silica gel will facilitate the integration of this technology in the laboratory, new plate materials, gradient plates (e.g., gradual material switch and thickness) and adsorbent combinations (e.g., multiphases and mixed phases) can be printed. This new potential will be exploited and trigger ideas for novel planar chromatographic separations. By 3D printing of thin layers, another asset was demonstrated to be integrated into the Office

DOI: 10.1021/acs.analchem.6b04813
Anal. Chem. 2017, 89, 2116−2122

Chromatography concept for online miniaturized planar chromatography.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.analchem.6b04813.

Table S-1, compilation of materials for 3D printer and slurry doser, including costs; Table S-2, calculation of production costs; Figure S-1, images of slicing process as .STL file; Figure S-2, images of Prusa i3 rework with FDM extruder installed; Figure S-3, correlation of layer thicknesses to mean pixel values to estimate the thicknesses of the printed layers; Figure S-4, 3D printing of different layer shapes and patterns; Figure S-5, influence of relative humidity of the air on separation; and Glossary G-1 with terms used (PDF)

5 min video showing the 3D print of a layer (AVI)

■ AUTHOR INFORMATION

Corresponding Author

*Phone: +49-641-99-39141. Fax: +49-641-99-39149. E-mail: Gertrud.Morlock@uni-giessen.de.

ORCID

Gertrud E. Morlock: 0000-0001-9406-0351

Notes

The authors declare no competing financial interest.

■ REFERENCES

(1) Sherma, J. *J. AOAC Int.* **2016**, *99* (2), 323−331.
(2) Morlock, G.; Schwack, W. *TrAC, Trends Anal. Chem.* **2010**, *29* (10), 1157−1171.
(3) Morlock, G.; Schwack, W. *J. Chromatogr. A* **2010**, *1217* (43), 6600−6609.
(4) Izmailov, N.; Shraiber, M. *Farmatsiya* **1939**, *6*, 1−7.
(5) Sherma, J.; Morlock, G. *J. Planar Chromatogr.–Mod. TLC* **2008**, *21*, 471−477.
(6) Kaiser, R.; Zlatkis, A. *High Performance Thin-Layer Chromatography*; Elsevier Inc.: Amsterdam, The Netherlands, 1977.
(7) Hauck, H. E.; Schulz, M. *J. Chromatogr. Sci.* **2002**, *40*, 550−552.
(8) Hauck, H. E.; Schulz, M. *Chromatographia* **2003**, *57*, S313−S315.
(9) Poole, S. K.; Poole, C. F. *J. Chromatogr. A* **2011**, *1218* (19), 2648−2660.
(10) Jim, S. R.; Brett, M. J. In *Instrumental Thin-Layer Chromatography*; Elsevier Inc.: Amsterdam, The Netherlands, 2015; pp 53−72.
(11) Morlock, G. *J. Chromatogr. A* **2015**, *1382*, 87−96.
(12) Lv, Y.; Lin, Z.; Tan, T.; Svec, F. *J.Chromatogr. A* **2013**, *1316*, 154−159.
(13) Frolova, A. M.; Konovalova, O. Y.; Loginova, L. P.; Bulgakova, A. V.; Boichenko, A. P. *J. Sep. Sci.* **2011**, *34*, 2352−2361.
(14) Clark, J. E.; Olesik, S. V. *Anal. Chem.* **2009**, *81* (10), 4121−4129.
(15) Beilke, M. C.; Zewe, J. W.; Clark, J. E.; Olesik, S. V. *Anal. Chim. Acta* **2013**, *761*, 201−208.
(16) Clark, J. E.; Olesik, S. V. *J. Chromatogr. A* **2010**, *1217* (27), 4655−4662.
(17) Kampalanonwat, P.; Supaphol, P.; Morlock, G. *J. Chromatogr. A* **2013**, *1299*, 110−117.
(18) Lu, T.; Olesik, S. V. *J. Chromatogr. B: Anal. Technol. Biomed. Life Sci.* **2013**, *912*, 98−104.
(19) Bezuidenhout, L. W.; Brett, M. J. *J. Chromatogr. A* **2008**, *1183*, 179−185.
(20) Jim, S. R.; Taschuk, M. T.; Morlock, G.; Bezuidenhout, L. W.; Schwack, W.; Brett, M. J. *Anal. Chem.* **2010**, *82* (12), 5349−5356.
(21) Oko, A. J.; Jim, S. R.; Taschuk, M. T.; Brett, M. J. *J. Chromatogr. A* **2011**, *1218* (19), 2661−2667.
(22) Jim, S. R.; Oko, A. J.; Taschuk, M. T.; Brett, M. J. *J. Chromatogr. A* **2011**, *1218* (40), 7203−7210.
(23) Jim, S. R.; Foroughi-Abari, A.; Krause, K. M.; Li, P.; Kupsta, M.; Taschuk, M. T.; Cadien, K. C.; Brett, M. J. *J. Chromatogr. A* **2013**, *1299*, 118−125.
(24) Wannenmacher, J.; Jim, S. R.; Taschuk, M. T.; Brett, M. J.; Morlock, G. *J. Chromatogr. A* **2013**, *1318*, 234−243.
(25) Song, J.; Jensen, D. S.; Hutchison, D. N.; Turner, B.; Wood, T.; Dadson, A.; Vail, M. A.; Linford, M. R.; Vanfleet, R. R.; Davis, R. C. *Adv. Funct. Mater.* **2011**, *21*, 1132−1139.
(26) Jensen, D. S.; Kanyal, S. S.; Gupta, V.; Vail, M. A.; Dadson, A. E.; Engelhard, M.; Vanfleet, R.; Davis, R. C.; Linford, M. R. *J. Chromatogr. A* **2012**, *1257*, 195−203.
(27) Kanyal, S. S.; Jensen, D. S.; Miles, A. J.; Dadson, A. E.; Vail, M. A.; Olsen, R.; Scorza, F.; Nichols, J.; Vanfleet, R. R.; Davis, R. C.; Linford, M. R. *J. Vac. Sci. Technol., B: Nanotechnol. Microelectron.: Mater., Process., Meas., Phenom.* **2013**, *31*, 031203.
(28) Jensen, D. S.; Kanyal, S. S.; Madaan, N.; Miles, A. J.; Davis, R. C.; Vanfleet, R.; Vail, M. A.; Dadson, A. E.; Linford, M. R. *J. Vac. Sci. Technol., B: Nanotechnol. Microelectron.: Mater., Process., Meas., Phenom.* **2013**, *31*, 031803.
(29) Kanyal, S. S.; Häbe, T. T.; Cushman, C. V.; Dhunna, M.; Roychowdhury, T.; Farnsworth, P. B.; Morlock, G.; Linford, M. R. *J. Chromatogr. A* **2015**, *1404*, 115−123.
(30) Morlock, G.; Oellig, C.; Bezuidenhout, L. W.; Brett, M. J.; Schwack, W. *Anal. Chem.* **2010**, *82*, 2940−2946.
(31) Häbe, T. T.; Morlock, G. *J. Chromatogr. A* **2015**, *1413*, 127−134.
(32) Gross, B. C.; Erkal, J. L.; Lockwood, S. Y.; Chen, C.; Spence, D. M. *Anal. Chem.* **2014**, *86*, 3240−3253.
(33) Gross, B. C.; Lockwood, S. Y.; Spence, D. M. *Anal. Chem.* **2017**, *89* (1), 57−70.
(34) Martínez-Jarquín, S.; Moreno-Pedraza, A.; Guillén-Alonso, H.; Winkler, R. *Anal. Chem.* **2016**, *88* (14), 6976−6980.
(35) Tycova, A.; Prikryl, J.; Foret, F. *Electrophoresis* **2016**, *37* (7−8), 924−930.
(36) Salentijn, G. I.; Permentier, H. P.; Verpoorte, E. *Anal. Chem.* **2014**, *86*, 11657−11665.
(37) Mathieson, J. S.; Rosnes, M. H.; Sans, V.; Kitson, P. J.; Cronin, L. *Beilstein J. Nanotechnol.* **2013**, *4* (1), 285−291.
(38) Duarte, L. C.; Colletes de Carvalho, T.; Lobo-Júnior, E. O.; Abdelnur, P. V. P. V.; Vaz, B. G.; Coltro, W. K. T. *Anal. Methods* **2016**, *8* (3), 496−503.
(39) Prikryl, J.; Foret, F. *Anal. Chem.* **2014**, *86*, 11951−11956.
(40) Sandron, S.; Heery, B.; Gupta, V.; Collins, D. A.; Nesterenko, E. P.; Nesterenko, P. N.; Talebi, M.; Beirne, S.; Thompson, F.; Wallace, G. G.; Brabazon, D.; Regan, F.; Paull, B. *Analyst* **2014**, *139* (24), 6343−6347.
(41) Gupta, V.; Talebi, M.; Deverell, J.; Sandron, S.; Nesterenko, P. N.; Heery, B.; Thompson, F.; Beirne, S.; Wallace, G. G.; Paull, B. *Anal. Chim. Acta* **2016**, *910*, 84−94.
(42) Fee, C.; Nawada, S.; Dimartino, S. *Journal of Chromatography A* **2014**, *1333*, 18−24.
(43) Google trend on 3D printing. https://google.com/trends/explore?date=all&q=3d+printing (accessed Nov 27, 2016).
(44) Wittbrodt, B. T.; Glover, A. G.; Laureto, J.; Anzalone, G. C.; Oppliger, D.; Irwin, J. L.; Pearce, J. M. *Mechatronics* **2013**, *23* (6), 713−726.
(45) Jones, R.; Haufe, P.; Sells, E.; Iravani, P.; Olliver, V.; Palmer, C.; Bowyer, A. *Robotica* **2011**, *29*, 177−191.
(46) Morlock, G.; Brett, N. *J. Chromatogr. A* **2015**, *1390*, 103−111.

DOI: 10.1021/acs.analchem.6b04813
Anal. Chem. 2017, 89, 2116−2122

Publication 2 - Supporting information

Supporting Information

Open-source-based 3D printing of thin silica gel layers in planar chromatography

Dimitri Fichou and Gertrud E. Morlock*

Chair of Food Sciences, Institute of Nutritional Science, and Interdisciplinary Research Center (IFZ), Justus Liebig University Giessen, Heinrich-Buff-Ring 26-32, 35392 Giessen, Germany

*Corresponding author. Tel.: +49-641-99-39141; fax: +49-641-99-39149; E-mail address: Gertrud.Morlock@uni-giessen.de (G. E. Morlock).

Table of contents

Table S-1 Compilation of materials needed for the 3D printer and slurry doser, including prices and total costs.

Name	Reference	Link manufacturer	Price per unit (Euro VAT)	Amount	Price total (Euro VAT)
Prusa i3 rework	1234568399	www.reprap-3d-printer.com	550	1	550
Syringe moving	Printed	-	0	1	0
Syringe holder	Printed	-	0	1	0
Syringe caps	Printed	-	0	1	0
Motor holder	Printed	-	0	1	0
Coupler	Printed	-	0	1	0
Smooth shafts 6 mm	1234568567	www.reprap-3d-printer.com	18	1	18
Linear bearing lm6uu	1234568275	www.reprap-3d-printer.com	2	2	4
Trapezoidal lead screw 8 mm and nut	B01E5GS33U	www.amazon.de	3	1	3
Nema 17	312	www.reprap-3d-printer.com	16	1	16
Inductive sensor	1234568401	www.reprap-3d-printer.com	15	1	15
ABS, 1kg	1234568224	www.reprap-3d-printer.com	21	1	21
Syringe, 10 mL	4616103V	us.bbraunoem.com	1	1	1
Syringe needle, 1.20 x 50 mm	4667123	us.bbraunoem.com	1	1	1
Screw M3 x 30	-	local hardware store	0	5	0
Screw M4 x 20	-	local hardware store	0	4	0
Nut M4	-	local hardware store	0	4	0
				Total slurry doser (Euro)	79
				Total modified 3D printer (Euro)	629

Table S-2 Calculation of production costs for the 3D layer print.

Materials	**Cost per gram (Euro)**	**Weight (g) for flat layer, 100% of 4 mL**	**Weight (g) for channeled layer, 30% of 2 mL**
Silica gel	0.25	0.88	0.13
Calcium sulfate	0.23	0.12	0.02
	Total price	**0.25 Euro**	**0.04 Euro**

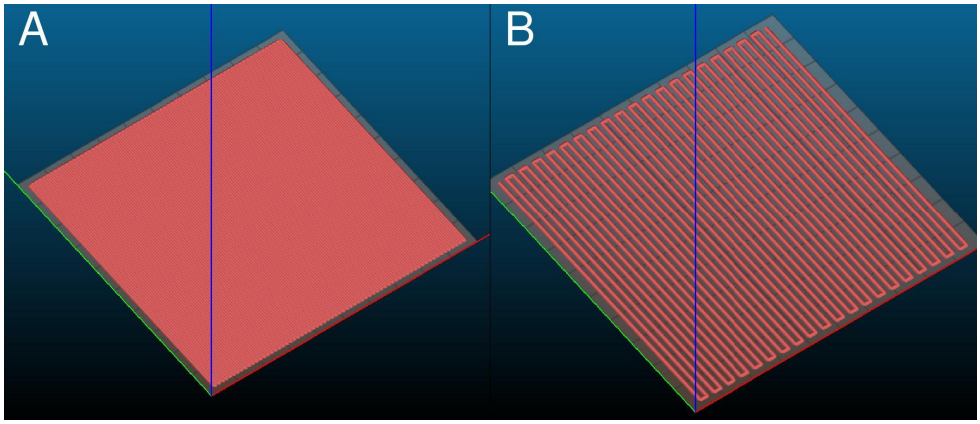


Figure S-1. Images of the slicing process for a 95 x 95 x 0.2 mm format as .STL file with (A) 100% infill for a plane layer and (B) 30% infill for a pattern of 40 channels.

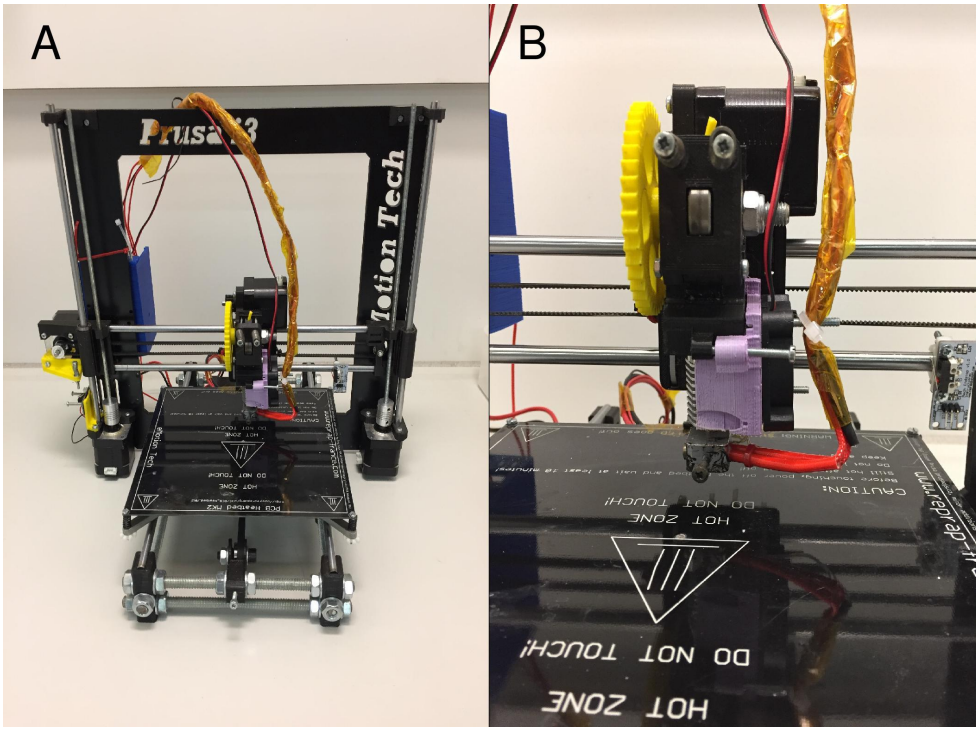


Figure S-2. Images of the (A) Prusa i3 rework with the FDM extruder installed (B, zoom).

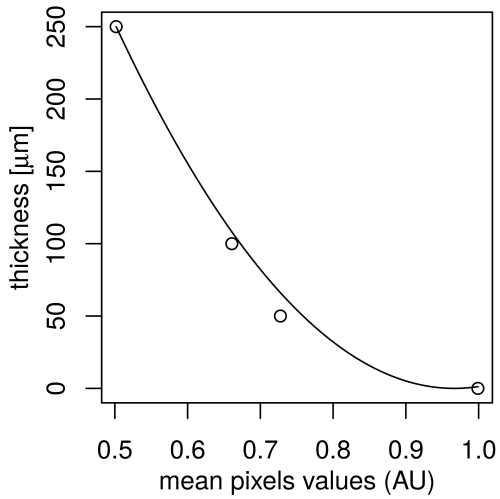


Figure S-3. Correlation (quadratic fitting) of the layer thicknesses of 50 µm, 100 µm and 250 µm to the mean pixel values in the center of the plate, used to estimate the layer thickness of the printed layers.

Figure S-4. 3D printing of different layer shapes and patterns: example to trigger ideas for novel planar chromatographic separations.

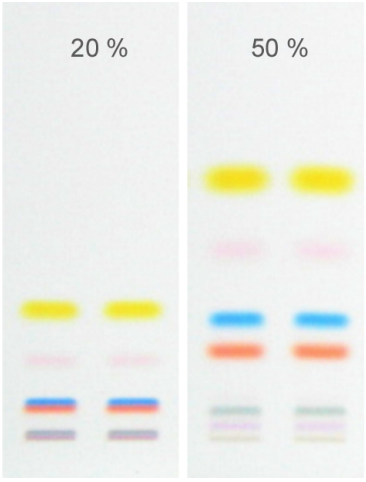


Figure S-5. Influence of the relative humidity of the ambient air (20% and 50%) on the separation of six dyes on a commercially available HPTLC plate silica gel 60 (Merck).

Glossary G-1

ABS: Acrylonitrile Butadiene Styrene: along with PLA: PolyLactic Acid, they were the 2 main plastics used with FDM.

Arduino: Open source micro controller, coupled with the RAMPS, it allowed to send and receive signal from the computer to the 3D printer.

Auto bed leveling: Feature allowing the correction of the Z level of the plastic extruder or slurry doser during the print, homogeneous level being primordial for a successful print.

CAD: Computer assisted design: Software based design used to design 3D object.

End stop: Sensor used by the printer at start up to define the origin of the Cartesian space.

FDM: Fused deposit manufacturing, the 3D printing technique used by the Prusa i3 and most RepRap 3D printer.

Firmware: Loaded in the Arduino, it allowed the interpretation of human readable information, the GCODE, into machine readable information.

First layer extrusion width: slic3r parameter, parameter allowing finer control of the applied material on the first layer.

GCODE: file containing command to be interpreted by the firmware, for example: *G1 X100 Y5 E10* means moved the extruder in position X100 and Y 5 and extruded 10 mm of material on the way; *M109 S110* set the stage temperature to 110°C.

Inductive sensor: Sensor used to detect the presence of metal in auto bed leveling feature.

Infill pattern: slic3r parameters, pattern chosen for the manufacturing of the object.

Infill perimeter overlap: slic3r parameter, set at a different value by default.

Infill: slic3r parameter defining the percentage of material to be extruded to form the object typically set at 20-50% for FDM.

Layer thickness: slic3r parameter defining the Z intervals of the additive manufacturing process.

Nozzle diameter: slic3r parameter, the inner diameter of the extruder output.

PCB heat bed: Printed circuit board: important part of a FDM 3D printer, mandatory to print ABS, optional for PLA, it allowed a better adherence of the plastic to the stage.

Plastic extruder: Main part of an FDM 3D printer, where the plastic was heated at high temperature, typically 250°C for ABS, and applied on the stage

RAMPS: Open source electronic board linked the Arduino board

RepRap: Self-replicating, open source 3D printer.

Slic3r: software used to convert .STL files into .GCODE files, contained several options allowing fine control of the process.

Step per millimeter: Firmware setting used to calculate the number of stepper motor steps necessary to produce a motion of 1 mm

Stepper motor: Highly accurate motor allowing precise movement.

STL: STereoLitography: File format widely used in 3D printing.

Trapezoidal lead screw and nut: Along with the smooth shaft and the brass steering bushing or lm6uu, they transformed the rotation of the stepper motor into translation allowing linear motion.

X carriage: in the Prusa i3, the plastic part supporting the extruder moved in the X and Z directions.

Publication 3

Powerful Artificial Neural Network for Planar Chromatographic Image Evaluation, Shown for Denoising and Feature Extraction

Dimitri Fichou[1] and Gertrud E. Morlock[1]

1. Chair of Food Science, Institute of Nutritional Science, and Interdisciplinary Research Center (IFZ), Justus Liebig University Giessen, Heinrich-Buff-Ring 26-32, 35392 Giessen, Germany

Published in:

Analytical chemistry (2018) 90:6984-6991

Submitted: March 22, 2018; Accepted: April 30, 2018; Published: April 30, 2018

analytical chemistry

Article

Cite This: Anal. Chem. 2018, 90, 6984−6991

pubs.acs.org/ac

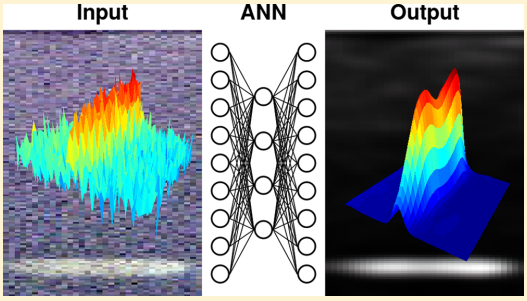
Powerful Artificial Neural Network for Planar Chromatographic Image Evaluation, Shown for Denoising and Feature Extraction

Dimitri Fichou and Gertrud E. Morlock*

Chair of Food Science, Institute of Nutritional Science, and Interdisciplinary Research Center (IFZ), Justus Liebig University Giessen, Heinrich-Buff-Ring 26-32, 35392 Giessen, Germany

Supporting Information

ABSTRACT: An artificial neural network (ANN) is presented as a new and superior technique for processing planar chromatography images. Though several algorithms are available for image processing in planar chromatography, the use of ANN has not been explored so far. It simulates how the human brain interprets images, and the intrinsic features of the image were captured on patches of pixels and successfully reconstructed afterward. The obtained high number of observations was a perfect basis for using ANN. As examples, three quite different data sets were processed with this new algorithm to demonstrate its versatility and benefits. Powerful features, which the ANN learned from the image data set, improved the quality of the analytical data. Thus, noise or inhomogeneous background of bioautograms was removed as demonstrated for salvia extracts, improving their bioquantifications. On colorful fluorescence chromatograms of further botanical extracts, the power and benefit of the feature extraction were demonstrated. Using ANN, videodensitometric results were improved. If compared to conventional digital processing, the resolution between two adjacent blue fluorescent bands increased from 0.95 to 1.18 or between two orange fluorescent bands from 0.77 to 1.57. The trueness of the new ANN was successfully verified by comparison with conventional densitometric results of the absorbance of separated tea extracts. The correlation coefficients of epigallocatechin gallate therein improved from 0.9889 with median filter to 0.9959 using this new ANN algorithm. The code was released open-source to the scientific community as a ready-to-use tool to exploit this potential, spread its usage, and boost improvements in planar chromatographic image evaluation.



In the field of computer vision, artificial neural networks (ANNs) are becoming the norm. These biologically inspired tools take advantage of the growing computing power, the use of more sophisticated algorithms, and the use of bigger data sets.[1] The best results on numerous tasks such as handwritten digit recognition are now produced with ANN, and no other technique could challenge those superior results.[2,3] The use of deep learning, which consists of multiple layers of hidden units, allowed the algorithm to learn more complicated features and solve challenging tasks.[4,5] Unsupervised techniques, like the restricted Boltzmann machine (RBM) algorithm[6−8] and denoising autoencoder,[9] are already state of the art in image denoising.[10−12]

Images also play an important role in high-performance thin-layer chromatography (HPTLC), which is a technique of analytical chemistry widely used in analytical laboratories.[13] Once the separation is done, images of the resulting chromatogram are captured via different light sources, or it is preceded by a derivatization step for selective or even specific detection. The output data are in form of an image, and traditionally, the conclusions are drawn on the basis of visual interpretation. To facilitate this image evaluation, preprocessing is needed to align and highlight the zone signals and diminish the influence of background or noise.[14] Simple techniques, like gamma and contrast corrections, are sufficient in the case of a visual inspection of the chromatogram.

More sophisticated techniques are required for the fast image-based quantification (videodensitometry), which is an upcoming trend. Before integration, preprocessing was needed either directly of the chromatogram (image) or of the extracted videodensitograms (single dimension vectors of data). For extraction, a third party software was used, e.g., ImageJ (U. S. National Institute of Health, Bethesda, MD, USA), JustTLC (Sweday, Sodra Sandby, Sweden), and Sorbfil TLC Videodensitometer (Jsc Sorbpolymer, Krasnodar, Russia).[15] For chromatogram preprocessing, the median filter was chosen as the most powerful one among several filters as well as the wavelet shrinkage algorithm.[16] For preprocessing of the extracted videodensitograms, baseline removal, Savitzky-Golay smoothing, standardization, and peak alignment were used.[17,18] Such preprocessing tools transform raw data (input) into ready-

Received: March 22, 2018
Accepted: April 30, 2018
Published: April 30, 2018

DOI: 10.1021/acs.analchem.8b01298
Anal. Chem. 2018, 90, 6984−6991

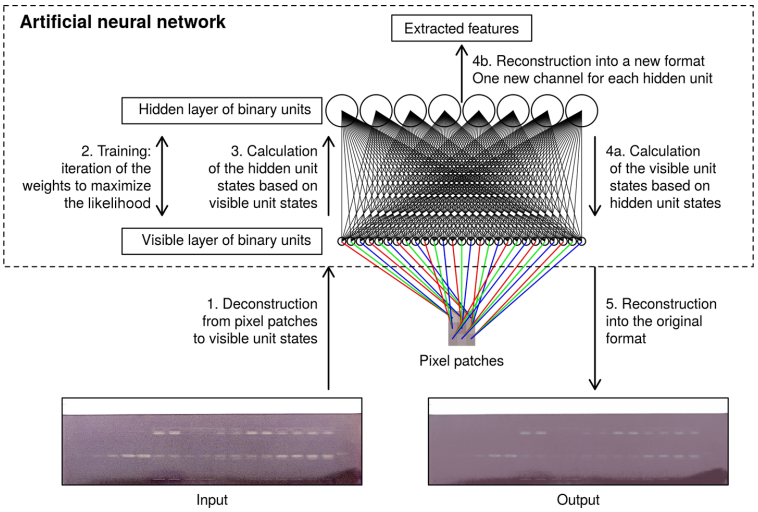


Figure 1. Workflow of the image reconstruction and processing by the ANN algorithm, depicted for a *Bacillus subtilis* bioautogram.

to-use data (output) for the next step of the analysis. They mitigate the influence of potential experimental errors and facilitate the chromatogram evaluation. In this respect, the power of ANN has not been exploited so far. For example, in direct bioautography, interpretation of the bioautogram can be compromised by an inhomogeneous background caused by the applied bioassay. Initial evaluations of such bioautograms were semiquantitative using a planimeter.[19,20] After improvement of the workflow and the use of scanning densitometry, bioquantifications were shown.[21−23] The question arises of how ANN may improve such especially challenging quantifications.

Videodensitometry in combination with chemometrics (Table S-1) was applied in two ways: unsupervised for pattern recognition or supervised to predict properties on new samples.[24−27] In both approaches, preprocessing of the signal was important and involved multiple steps, thoroughly chosen for a specific project. Generally, grayscale or specific channels of the red, green, and blue (RGB) color space were selected. The same preprocessing tools[14] were used before machine learning model calculations. ANN was used in this context to predict properties on samples (predictive statistics);[25,27] however, it should be pointed out that the number of samples was limited. Thus, ANN could not show its full potential. Recently, the proof-of-principle of effect-directed classifications was demonstrated.[28]

In this study, the potential of the combination of planar chromatograms or bioautograms with ANN was investigated using the RBM algorithm. This special ANN algorithm iteratively learns to reconstruct its input features by encoding them into a hidden layer of features. Weight parameters link each visible unit to each hidden unit. Crossing the ANN is done by a matrix multiplication and an operation to transform the result of the matrix multiplication into a standardized range. During the training of the ANN, the weights are iteratively modified to maximize the likelihood of the reconstructed visible layer. The principle of this ANN algorithm is discussed as well as the effect of the different options available. Three quite different data sets were processed with the algorithm, i.e., images taken at white light illumination (vis, Gram-positive antibiotics in the bioautogram), UV 254 nm (absorbing compounds), and UV 366 nm (fluorescent compounds). The capabilities of ANN and its advantages over established state of the art methods should be figured out as well as its potential to boost scientific progress.

EXPERIMENTAL SECTION

Samples and Data Sets. All samples were purchased from a local pharmacy in Giessen. Botanical extracts of melissa, mentha, salvia, and artichoke were analyzed according to the European Pharmacopoeia method of artichoke.[29] Tea extracts were analyzed according to a previous method.[30] A 3.5 $\mu g/\mu L$ methanolic epigallocatechin gallate (EGCG) solution was used. *Salvia miltiorrhiza* extracts were analyzed according to a direct bioautography method.[21] The resulting three images were used as data sets.

Hardware and Software. All code was written in the R programming language.[31] In particular, the deepnet package for the RBM algorithm[32] was used. All computations were run on a Lenovo G510 laptop with 4 CPU Intel® Core i5-4200M CPU, 2.50 GHz, and 16 GB RAM. Computers with a lower capacity were not suited due to the high computing power needed to run these analyses. The newly programmed code was made available on GitHub in the form of an R package (https://github.com/DimitriF/DLC), including instructions for installing and for running locally a simple program to apply the process on other data sets.

Input Data and Deconstruction. Preprocessing tools were avoided, and images were read as captured. The presented algorithm considered the pixels as observations and not the samples. Before applying the deconstruction process, it was necessary to access the RGB bitmap and apply normalization. The dimension of the plate image was reduced from the classic 2.000 × 1.000 pixels (width × height) to a respective smaller dimension of 256 pixels in height for the bioautogram and UV 366 nm chromatogram and 512 pixels in height for the UV 254 nm chromatogram. The ratio of pixels in height and width was fixed during this resizing. The deconstruction process consisted

DOI: 10.1021/acs.analchem.8b01298
Anal. Chem. 2018, 90, 6984−6991

Figure 2. Selected examples showing the ANN effect on the selected number of hidden units and epochs for the *Bacillus subtilis* bioautogram (A) and HPTLC chromatogram at UV 366 nm containing fluorescent zones of four different botanical extracts (B).

of flattening the 3D data array in a 2D matrix with pixel patches as rows. The pixel patch size was set to be 5 × 5 pixels in all experiments.

RBM Algorithm Training. Most of the selectable parameters were fixed to be 1000 for the mini-batch size, 0.1 for the learning rate, 2 for the contrastive divergence, and 0.5 for the momentum. Only two hyperparameters were changed from one experience to another: the number of hidden units and the number of epochs (Table S-2).

The number of hidden units corresponded to the number of features, for which the ANN was able to encode the information before trying to retrieve it. The number of epochs represented the number of times the entire data set was presented to the ANN during the training.

Reconstruction and Feature Extraction. To be suitable for interpretation, the state of the input and hidden units had to be converted back into a 3D array. For the reconstruction with the input unit, the data set size was multiplied during deconstruction, and thus, a mean RGB value of each pixel was applied to retrieve the original dimension. For the feature extraction with the hidden layer, a 3D array was created with one new channel per hidden unit.

Quantitative Evaluation. The chromatogram at UV 254 nm was available as a data set, showing EGCG in 18 tea samples. The classical method used a median filter of 3, followed by videodensitogram extraction, baseline correction with the rolling ball algorithm, and finally peak integration via the green channel.[16] The rolling ball algorithm was used for preprocessing of 2D HPTLC data.[33] It simulates the rolling of a ball of a defined diameter along the videodensitogram. The ball could not fall inside the peak, as the diameter was larger. Therefore, only noise and background were removed, whereas the signal was sustained. The baseline package from R had been used with 30 as width of the local window for minimization/maximization and 5 as width of the local window for smoothing.[34] The new ANN technique used the same process on the green channel after preprocessing and on one of the extracted features from one of the hidden units.

DOI: 10.1021/acs.analchem.8b01298
Anal. Chem. 2018, 90, 6984−6991

RESULTS AND DISCUSSION

Input Data and Deconstruction. As input data sets, three quite different images were used as captured. (1) As the first example for the detection in the visible range, direct bioautography was used for effect-directed analysis (HPTLC-EDA-Vis) of antibiotics active against Gram-positive bacteria in *Salvia miltiorrhiza* extracts. The resulting *Bacillus subtilis* bioautogram was recorded under white light illumination. A drawback of such a biological detection may be a bioautogram with an irregular background and noisy baseline. This was considered as an appropriate challenge for ANN with regard to any improvements. (2) In the second example for fluorescence detection (HPTLC-FLD), four different botanical extracts were separated. The resulting UV 366 nm chromatogram with colorfully fluorescent zones challenged the ANN to learn its reconstruction. (3) In the third example for detection of absorbing compounds (HPTLC-UV), tea extracts were separated, and the resulting UV 254 nm chromatogram contained quantitative information to compare the performance of quantitative image evaluation techniques (ANN classical videodensitometry).

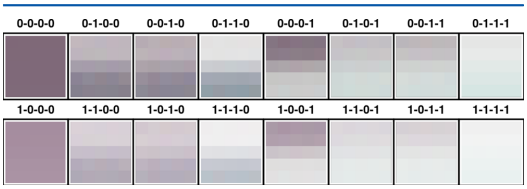


Figure 3. Synthetic pixel patches for all possible binary states of a 4 hidden unit ANN, trained on the *Bacillus subtilis* bioautogram for a 50 epochs processing.

Most of the machine learning algorithms took a matrix as input, i.e., two dimensions with observations as rows and variables as columns. Chemometrics is a specific field of data science because the number of variables is usually much higher than the number of observations. In contrast, the presented ANN algorithm considered the pixels as observations and not the samples as is usual for most chemometric approaches. Thus, the number of observations was substantially increased. Another difference to classical chemometrics applied in HPTLC was the avoidance of preprocessing tools, except for normalization between 0 and 1, which worked well using a sigmoidal function. There was no variable selection step, and thus, all available data in the RGB image were used. For example, by considering only the grayscale image or by using only a specific color channel, important information would be lost for ANN. This comprehensive use-all-data approach was also important for ANN to reconstruct the image afterward and to sustain the high visual power of HPTLC images.

The dimension of the image was reduced (from 2000 to 256 or 512 pixels in height) to keep the computing time affordable and to cope with the data size. Once the 3D array was loaded into the RAM, the processing itself took place. The first step was to deconstruct it into a 2D matrix with 1 row per pixel (Figure 1, step 1). The trick here was to integrate also the surrounding pixels. Such pixel patches let the system capture a smaller or bigger window of information. For example, a patch of 3 × 3 pixels or 5 × 5 pixels led, respectively, to 27 or 75 variables in the deconstructed matrix. Thus, a 512 × 1024 pixel image produced around 500 000 of those pixel patches. This high number allowed the use of higher level algorithms than the ones used previously in HPTLC combined with chemometrics.

Training the RBM Algorithm. This deconstructed matrix was the input for the RBM algorithm. This ANN algorithm had the task to capture the intrinsic features of the data set without the noise and background. There were multiple parameters to select when applying this algorithm. Only two hyperparameters (hidden units and epochs) were kept flexible from one experience to another. Most selectable parameters were kept fixed. Compromises had to be made between speed of training, trueness of the representation, and removal of the noise and background. The mini-batch size fixed at 1000 was the number of samples used between weight iterations. The learning rate fixed at the factor of 0.1 was applied when iterating the weight. For example, a smaller factor value led to a slower learning while a too big value missed local minima, and thus, it was the optimal data representation. The contrastive divergence fixed at 2 was an inherent parameter to the training of the RBM algorithm. It represented the number of times the mini-batch crossed the ANN before evaluation of errors and iteration of weights. The momentum fixed at a factor of 0.5 helped the ANN to learn faster by remembering the last weight iteration and proceeding in the same direction, resulting in less oscillation and allowing a faster equilibrium. Finally, the activation function for both visible and hidden units was a sigmoid function.

Reconstruction and Feature Extraction. Such ANN was trained for each new data set (image). The number of hidden units corresponded to the number of features, for which the ANN was able to encode the information before trying to retrieve it. The number of epochs represented the number of times the entire data set was presented to the ANN during the training. As this operation was computing intensive (up to 15 min), it was possible to save the weight parameter matrix to use it on another data set (Figure 1, step 2). Once a trained ANN was available, the data set crossed it for the last time (Figure 1, steps 3 and 4a). The final step was the reconstruction into the

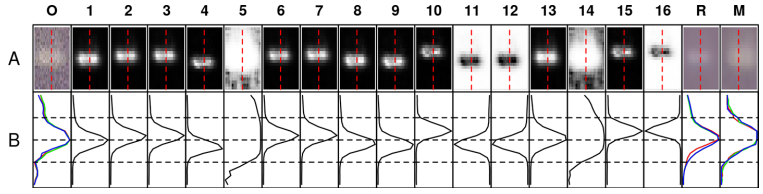


Figure 4. Extracted ANN features (A) of a zoomed *Bacillus subtilis* bioautogram band and respective normalized videodensitograms (B) via an ANN with 16 hidden units trained for 50 epochs: original data (O), normalized hidden unit states data (1−16), reconstructed data (R), and data after median filter processing (M).

DOI: 10.1021/acs.analchem.8b01298
Anal. Chem. 2018, 90, 6984−6991

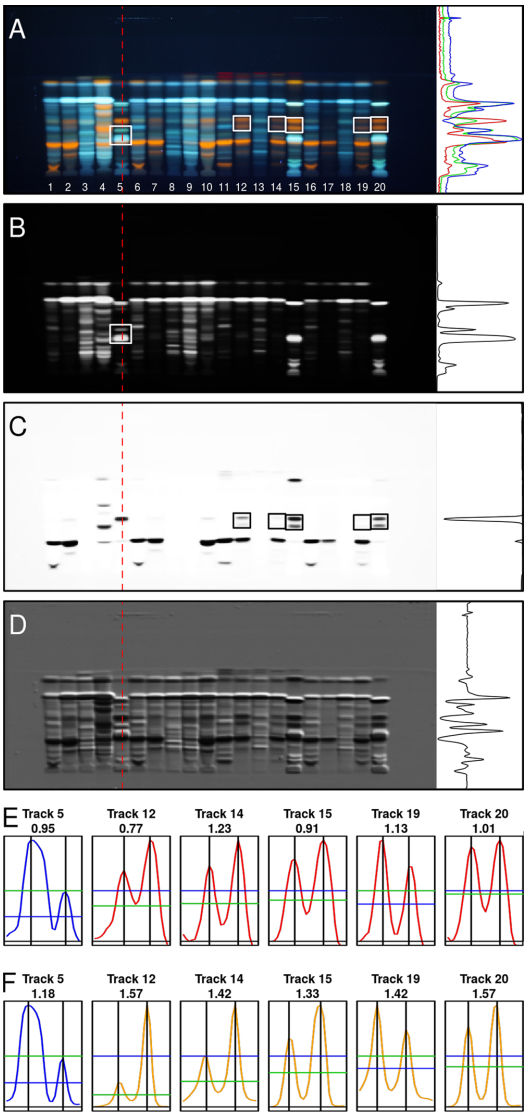


Figure 5. Exploration of an ANN with 128 units trained for 50 epochs for the HPTLC chromatogram at UV 366 nm (videodensitogram of red-dashed track): original image (A) and three units responsible for discrimination of orange (B) or blue bands (C) or band start and band end (D); videodensitograms for comparison of the resolution (listed below track no.) obtained by the blue and red channels after baseline correction (E) versus those obtained by extracted ANN features without baseline correction (F) of two blue fluorescent bands framed in B (track 5) and two orange fluorescent bands framed in C (tracks 12, 14, 15, 19, and 20).

original format, meaning original array dimension (Figure 1, step 5).

Originally, we only intended this image reconstruction to result in an improved quality after crossing the ANN. However, during the training phase, the system learned to represent

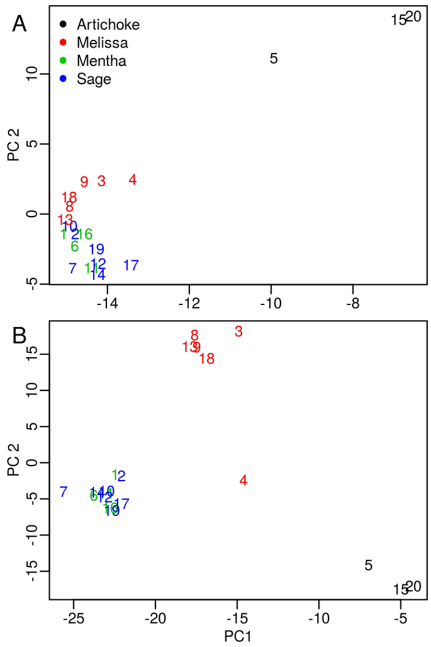


Figure 6. Principal component analysis for data obtained after standard normal variate preprocessing using the blue channel (A) versus those obtained by the three extracted ANN features of Figure 5B–D (B); numbers correspond to track numbers.

important features of the data into the hidden layers. By using the hidden unit states as new channels, it was possible to expand the color space and to create new information (Figure 1, step 4b). Those new channels were dependent on the RGB values and on the spatial distribution of the intensity in a specific pixel patch and, thus, focused on special features. Highly interesting new features were discovered as discussed later.

Tuning of the Hyperparameters. The choice of hyperparameters depended on the chromatogram type and the objective. For example, the objective for the *Bacillus subtilis* bioautogram of separated *Salvia miltiorrhiza* extracts was to remove the CCD noise and the darkened background at the lower plate edge (Figure 2A). Considering the effect of the number of epochs and hidden units on the denoising and trueness of the bioautogram, 50 epochs and 16 hidden units were considered to be optimal. A higher number of epochs was crucial for the ANN weights to stabilize, which means that fewer differences between iterations were given. In contrast, a small number of epochs was sufficient, if the objective was to produce a background-corrected image. A high number of hidden units allowed the system to capture the background, while a small number (e.g., 4 or 8) forced the system to focus on the signal. Interestingly, the ANN algorithm used the white part on the plate top as a starting point. At the beginning of the training, pixel patches from this region activated units in the same way as pixel patches of bright bands. However, at the end of the training, those two types of bright data were successfully discriminated by the ANN.

As the second example, the chromatogram at UV 366 nm showing the separation of further botanicals (Figure 2B) was

DOI: 10.1021/acs.analchem.8b01298
Anal. Chem. 2018, 90, 6984–6991

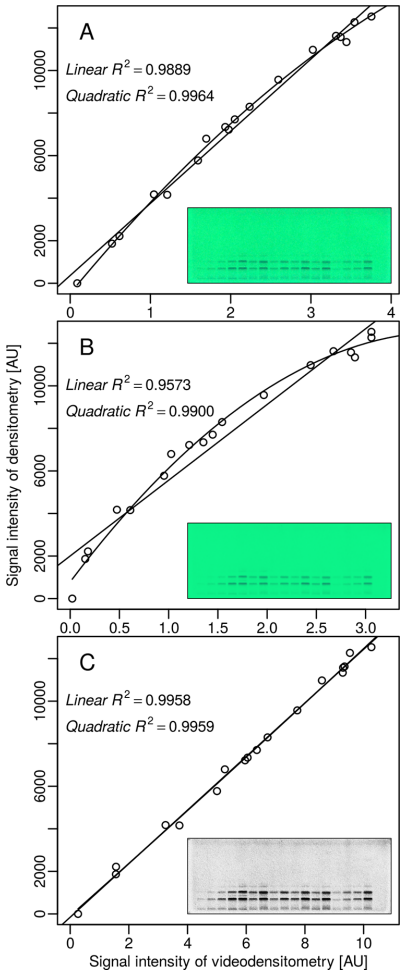


Figure 7. Correlation of the EGCG signal responses in tea extracts obtained by conventional densitometry versus three different processings for videodensitometry, i.e., median filtering (A), ANN denoising (B, green channel), and ANN feature extraction (C; ANN trained with 8 hidden units for 50 epochs).

investigated with regard to the impact of the number of epochs and hidden units needed to reconstruct the image. In this case, the objective for the ANN was to encode the intrinsic features of the data set inside the hidden layers. After 100 epochs, the ANN managed to represent the image, independent of the number of hidden units. However, for higher numbers of hidden units, the ANN learned to rapidly represent some more pronounced patterns, for example, the orange band pattern. Also along with a higher number of hidden units, the evolution of the reconstruction slowed down earlier. This indicated that the ANN was able to reach a stable state more rapidly, and the evolution of the weight parameters stopped earlier, linking the visible and hidden units. In the opposite direction, forcing the ANN to encode in a smaller number of hidden units led to a blurry chromatogram. Indeed, the information storage capacity of the system was smaller; so, important details were removed during the reconstruction, and the chromatogram was smoothened. This effect was similar to the one observed with a Savitzky-Golay smoothening or median filter processing with a too big window size: noise was removed but so was the signal. To conclude, a high number of hidden units and epochs had to be used for a good reconstruction of the image. For example, 128 hidden units were used for this data set with 50 epochs.

Investigation of Other Libraries. Other libraries of machine learning, i.e., the darch library[35] in R and the theano library[36] in Python, were used to test whether a similar processing technique and result can be achieved on the given data sets. Though another set of parameters could be optimal for those libraries, it is important to note that the parameters were kept static and just transferred. For optimization of the hyperparameters, a good start was important. Complex ANNs were difficult to train, and it could happen that it failed to train on a specific subject, as one parameter was wrongly tuned. The deepnet library results were not so easy to achieve with the darch and theano libraries. A possible explanation was that one or more parameters that are static with deepnet were wrongly set in those libraries for this specific task. More detailed investigations are recommended; however, it is out of the scope of the present study.

From Hidden to Visible Units. By fixing a synthetic state of hidden units, it was possible to calculate the corresponding visible unit states and to reconstruct an artificial patch, as shown for synthetic patches for all the possible binary states of a small ANN of 4 hidden units (Figure 3).

When no unit was activated, the ANN reconstructed the baseline without the lowest signal. When one unit was activated, the ANN reconstructed a specific zone in the chromatogram. For example, some hidden units were specialized in the detection of the zone start, while other hidden units were specialized in the zone end. This result was expected, but it is important to note that the training was unsupervised. The ANN system was told at no time what feature to learn. Different combinations of hidden units generated different patterns. This demonstrated the collaborative behavior of the ANN, meaning by specializing different units for the detection of different patterns, the global reconstruction of the chromatogram was getting closer to the original input. Note that the weight parameter initialization at the beginning was random, so the result from one training to another could be slightly different. Nevertheless, the ANN specialized its units to recognize the same kind of pattern for a given data set with a given set of hyperparameters. This underlined the robustness of its processing.

From Visible to Hidden Units. When training the ANN, hidden units were trained to specialize in the detection of different patterns. Thus, it was possible to use such an ANN as a feature extraction tool. By only crossing up the ANN (Figure 1, step 3), the hidden unit states of each pixel of the image were available and an array in a new format was reconstructed (Figure 1, step 4b). By carefully selecting the hyperparameters for a given data set, it was possible to obtain a system that learned to represent specific information inside the hidden layer that was not available before.

In the bioautogram, the three RGB channels were highly correlated after the denoising, and the ANN focused on Cartesian differences. Figure 3 could not be reconstructed with a 16-units ANN, as it would lead to 2^{16} possibilities. Nonetheless, it was possible to look at the characteristics of the 16 units by looking at their activation on a specific band

DOI: 10.1021/acs.analchem.8b01298
Anal. Chem. 2018, 90, 6984−6991

(Figure 4). A few characteristics were outstanding. Hidden units 11, 12, and 16 worked oppositely to other units. An explanation was that those were inhibitory units responsible for correction of the intensity. Units 5 and 14 were activated in a completely different way, possibly responsible for other patterns on the plate. Hidden units 9 and 10 were, respectively, activated for the start and end of a zone or peak. The last two zoomed bands R and M (Figure 4) showed the RGB chromatograms after reconstruction and median filtering. In both cases, the chromatograms and videodensitograms were smoothened. In contrast, the videodensitograms for the hidden units were sharper, making them better suitable for interpretation rather than the reconstruction. Such new features opened new avenues and might improve quantitative results as discussed later.

For the HPTLC chromatogram at UV 366 nm, the pixel patches population was broader than the one for the previously evaluated bioautogram. Thus, the ANN needed more hidden units to capture and reconstruct the intrinsic features of the data set. For feature extraction, the ANN used the compounds fluorescent at UV 366 nm (Figure 5A) and showed that the first hidden unit was specialized in the discrimination of the orange colored bands (Figure 5B), while the second one was responsible for the discrimination of blue hues (Figure 5C). As for the bioautogram at white light illumination in Figure 4, the same inhibitory or promoting characteristics were evident for this chromatogram at UV 366 nm, though in a more selective way. The specialization on orange worked by deactivating the hidden unit states, while the blue one worked by activating it. For both units, there was a return to the baseline, which was impossible to achieve using the RGB channel or the grayscale image. The resolution between two blue fluorescent adjacent zones (Figure 5B, framed) was 0.95 after baseline correction via the blue channel versus 1.18 without baseline correction in this specific ANN unit (Figure 5E versus F). The same resolution comparison was also performed on two orange fluorescent bands for several tracks (Figure 5C, framed). An up to 2-fold improvement in the resolution (from 0.77 to 1.57) was observed with this ANN unit (Figure 5E versus F).

The last hidden unit (Figure 5D) showed a pattern difficult to interpret. An explanation was that it was responsible for discrimination between peak start and peak end, letting the rest of the ANN focus on the color itself. Those three ANN units were the most remarkable, but a variety of other patterns can be observed in the other ANN units (Figure S-1).

Those newly created channels, among which a selective deconvolution phenomenon was observed, could be used to draw more profound conclusions. This was illustrated with principal component analysis (Figure 6). The best result with the original data set was obtained by the blue channel after standard normal variate preprocessing. Only the artichoke samples were clearly separated from the other samples (Figure 6A). By using the three extracted features presented in Figure 5 instead (also blue channel), it was possible to clearly distinguish the melissa samples from the other plants and thus improve the analysis (Figure 6B).

Quantitative Evaluation. As shown before, depending on the hyperparameters chosen, the data could be deformed during the process. This was not a problem as long as the signal of interest was still present. However, for quantitative evaluation, it was important to verify that the ANN was able to truly reconstruct the input data set. To compare the power of the ANN evaluation versus other techniques, a UV 254 nm chromatogram was used showing UV absorbing compounds in tea extracts (Figure 7). For each track out of the 18 tracks in the data set, the absorbance of EGCG was measured by conventional densitometry versus videodensitometry using three different processings, i.e., median filtering (Figure 7A), ANN denoising (Figure 7B), or ANN feature extraction (Figure 7C). All chromatograms were processed with baseline correction; the respective peaks were integrated, and the EGCG standard levels were assigned. The conventional densitometric EGCG response (y-axes) was plotted against each of the three different videodensitometric responses (x-axes) and depicted as a scatterplot. The determination coefficients R^2 of the linear and quadratic correlation models were determined.

For this data set, the reconstruction (green channel) after denoising by ANN (Figure 7B) showed worse results than after conventional median filtering (Figure 7A). The ANN deformed the signal during the process and the data correlated less. Instead of using the green channel of the reconstruction, better results were obtained when using one of the extracted features (Figure 7C) of the ANN trained with 8 hidden units for 50 epochs. The extracted feature result was equivalent with the median filtering (Figure 7A) for the quadratic model and even much better for the linear model. To conclude, ANN showed the power to improve the quantitative results. Some hidden units were specialized on quantitative aspects, while others were specialized on more abstract features. In this example, three hidden units out of eight in the ANN had linear quantitative properties with $R^2 > 0.9900$ and the best with an R^2 of 0.9958 (Figure 7C).

CONCLUSIONS

Inspired by biological processes, the way the brain processes HPTLC chromatograms was simulated. The new potential of ANNs was clearly demonstrated. The objective and challenge for this new ANN algorithm to cope with noisy images successfully resulted in improved videodensitograms (noise reduction, improved resolution, and quantitative performance). The developed ANN systems were also suited for extraction of specialized features, baseline correction, or deconvolution. ANNs encode data in an abstract way and offer new avenues for evaluations and advantageous options. In most applications, the inside of an ANN is difficult to interpret. By training such an ANN on HPTLC chromatograms, it was proven that the encoding correlated with the concentrations and results in the real world (compared to conventional densitometry), which verified the trueness of the new ANN. Such a rewarding processing by this new ANN algorithm is easily integrated in an analytical pipeline using the actual chromatogram as a training set or a previously trained ANN instead. The code was released as an open-source user interface and as an R package, so developers may explore more deeply its potential. This newly developed powerful open-source tool will boost image evaluation, which is an emerging key feature for fast profiling and fingerprinting of a high number of samples in a wide field of applications due to its many benefits.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.analchem.8b01298.

DOI: 10.1021/acs.analchem.8b01298
Anal. Chem. 2018, 90, 6984−6991

Table S-1: List of chemometric studies on TLC/HPTLC data (not comprehensive); Table S-2: Settings of the flexible parameters for different objectives and data sets; Figure S-1: The 128 extracted features from the model trained with the HPTLC chromatogram at UV 366 nm (PDF)

AUTHOR INFORMATION

Corresponding Author

*Tel.: +49-641-99-39141. Fax: +49-641-99-39149. E-mail: Gertrud.Morlock@uni-giessen.de (G.E.M.).

ORCID

Gertrud E. Morlock: 0000-0001-9406-0351

Notes

The authors declare no competing financial interest.

REFERENCES

(1) LeCun, Y.; Bengio, Y.; Hinton, G. *Nature* **2015**, *521*, 436−444.
(2) Ranzato, M.; Poultney, C.; Chopra, S.; Cun, Y. L. Efficient Learning of Sparse Representations with an Energy-Based Model. In *Advances in neural information processing systems 19*. Schölkopf, B., Platt, J. C., Hoffman, T., Eds.; MIT Press: Cambridge, MA, 2007; pp 1137−1144.
(3) Krizhevsky, A.; Sutskever, I.; Hinton, G. E. ImageNet Classification with Deep Convolutional Neural Networks. In *Advances in neural information processing systems 25*; Pereira, F., Burges, C. J. C., Bottou, L., Weinberger, K. Q., Eds.; Curran Associates, Inc.: Red Hook, NY, 2012; pp 1097−1105.
(4) Bengio, Y.; Lamblin, P.; Popovici, D.; Larochelle, H. Greedy Layer-Wise Training of Deep Networks. In *Advances in neural information processing systems 19*; Schölkopf, B., Platt, J. C., Hoffman, T., Eds.; MIT Press: Cambridge, MA, 2007; pp 153−160.
(5) Hinton, G. E.; Osindero, S.; Teh, Y.-W. *Neur. Comput.* **2006**, *18*, 1527−1554.
(6) Freund, Y.; Haussler, D. Unsupervised Learning of Distributions on Binary Vectors Using Two Layer Networks. In *Advances in neural information processing systems 4*; Moody, J. E., Hanson, S. J., Lippmann, R. P., Eds.; Morgan-Kaufmann: Burlington, MA, 1992; pp 912−919.
(7) Hinton, G. E. *Neur. Comput.* **2002**, *14*, 1771−1800.
(8) Hinton, G. E.; Salakhutdinov, R. R. *Science* **2006**, *313*, 504−507.
(9) Vincent, P.; Larochelle, H.; Lajoie, I.; Bengio, Y.; Manzagol, P.-A. *J. Mach. Learn. Res.* **2010**, *11*, 3371−3408.
(10) Burger, H. C.; Schuler, C. J.; Harmeling, S. *IEEE Conf. on Comp. Vis. and Pat. Recog.* **2012**, 2392−2399.
(11) Tang, Y.; Salakhutdinov, R.; Hinton, G. *IEEE Conf. on Comp. Vis. and Pat. Recog.* **2012**, 2264−2271.
(12) Roth, S.; Black, M. J. *Int. J. Comput. Vis.* **2009**, *82*, 205−229.
(13) Morlock, G.; Schwack, W. *J. Planar Chromatogr.–Mod. TLC* **2007**, *20*, 399−406.
(14) Ristivojević, P. M.; Morlock, G. E. *J. Planar Chromatogr.–Mod. TLC* **2016**, *29*, 310−317.
(15) Milojković-Opsenica, D.; Ristivojević, P.; Andrić, F.; Trifković, J. *Chromatographia* **2013**, *76*, 1239−1247.
(16) Komsta, Ł. *Anal. Chim. Acta* **2009**, *641*, 52−58.
(17) Kaczmarek, K.; Walczak, B.; Jong, S. D. *Acta Chromatogr.* **2005**, *15*, 82−96.
(18) Komsta, Ł. *J. Chromatogr. A* **2009**, *1216*, 2548−2553.
(19) Choma, I. M.; Choma, A.; Komaniecka, I.; Pilorz, K.; Staszczuk, K. *J. Liq. Chromatogr. Relat. Technol.* **2004**, *27*, 2071−2085.
(20) Choma, I. M.; Grzelak, E. M. *J. Chromatogr. A* **2011**, *1218*, 2684−2691.
(21) Jamshidi-Aidji, M.; Morlock, G. *J. Chromatogr. A* **2015**, *1420*, 110−118.
(22) Klingelhöfer, I.; Morlock, G. *Anal. Chem.* **2015**, *87*, 11098−11104.
(23) Jamshidi-Aidji, M.; Morlock, G. E. *Anal. Chem.* **2016**, *88*, 10979−10986.
(24) Morlock, G.; Ristivojević, P.; Chernetsova, E. S. *J. Chromatogr. A* **2014**, *1328*, 104−112.
(25) Tian, R.; Xie, P.; Liu, H. *J. Chromatogr. A* **2009**, *1216*, 2150−2155.
(26) Wong, K. H.; Razmovski-Naumovski, V.; Li, K. M.; Li, G. Q.; Chan, K. *J. Pharm. Biomed. Anal.* **2014**, *95*, 11−19.
(27) Agatonovic-Kustrin, S.; Loescher, C. M. *Anal. Chim. Acta* **2013**, *798*, 103−108.
(28) Ristivojević, P.; Morlock, G. *Food Chem.* **2018**, *260*, 344.
(29) *Ph. Eur.* 9.0, 1254 (01/2017).
(30) Schwegler, R. Master thesis, Justus Liebig University Giessen, 2016.
(31) R Core Team. *R, a language and environment for statistical computing*; R Foundation: Vienna, 2016.
(32) Rong, X. *deepnet: deep learning toolkit in R*; 2014, https://CRAN.R-project.org/package=deepnet.
(33) Komsta, Ł.; Cieśla, Ł.; Bogucka-Kocka, A.; Józefczyk, A.; Kryszeń, J.; Waksmundzka-Hajnos, M. *J. Chromatogr. A* **2011**, *1218*, 2820−2825.
(34) Hovde Liland, K.; Mevik, B.-H.; Canteri, R. *Baseline: Baseline Correction of Spectra*, Version 1.2-1; 2015, https://CRAN.R-project.org/package=baseline.
(35) Drees, M.; Rueckert, J.; Hinton, G.; Salakhutdinov, R.; Rasmussen, C. E. *darch: Package for Deep Architectures and Restricted Boltzmann Machines*; 2015, https://CRAN.R-project.org/package=darch.
(36) Bergstra, J.; Breuleux, O.; Bastien, F.; Lamblin, P.; Pascanu, R.; Desjardins, G.; Turian, J.; Warde-Farley, D.; Bengio, Y. Theano: A CPU and GPU Math Compiler in Python. In *Proceedings of the 9th Python in Science Conference*; June 30−July 3, Austin, TX, 2010; pp 3−10.

DOI: 10.1021/acs.analchem.8b01298
Anal. Chem. 2018, 90, 6984−6991

Publication 3 - Supporting information

Supporting Information

Powerful Artificial Neural Network for Planar Chromatographic Image Evaluation, Shown for Denoising and Feature Extraction

Dimitri Fichou and Gertrud E. Morlock*

Chair of Food Sciences, Institute of Nutritional Science, and Interdisciplinary Research Center (IFZ), Justus Liebig University Giessen, Heinrich-Buff-Ring 26-32, 35392 Giessen, Germany

*Corresponding author. Tel.: +49-641-99-39141; fax: +49-641-99-39149; E-mail address: Gertrud.Morlock@uni-giessen.de (G. E. Morlock).

Table of contents

Table S-1. List of chemometric studies on TLC/HPTLC data (not comprehensive)

Year	Sample	Unsupervised	Supervised	Reference
2009	herbs	PCA	K-nn, ANN	1
2009	herbs	PCA	-	2
2011	herbs	PCA	-	3
2011	propolis	PCA, HCA, hierarchical fuzzy clustering	-	4
2011	pharmaceutical formulations	PCA, HCA	-	5
2012	herbs	PCA	PLS-DA, OPLS-DA	6
2013	herbs	-	ANN	7
2014	microalgea	PCA	-	8
2014	propolis	PCA, HCA	LDA	9
2014	herbs	-	K-nn, CART, SPA-LDA, PCA-DA, SVM, PLS-DA	10
2014	propolis	PCA, HCA	PLS-DA	11
2014	herbs	PCA	-	12
2014	propolis	Similarity analysis, HCA, K-means clustering	ANN, SVM	13
2014	herbs	PCA	-	14
2014	herbs	PCA	-	15
2016	microalgae	PCA	-	16
2016	propolis	PCA	-	17
2016	herbs	PCA	-	18
2016	propolis	PCA	-	19
2016	propolis	PCA	-	20
2017	herbs	-	PLS	21
2017	herbs	PCA, HCA	-	22
2017	herbs	PCA, HCA	-	23
2017	biopolymers	PCA, HCA	-	24
2017	propolis	PCA	-	25
2017	herbs	PCA, HCA, heatmap	-	26
2018	propolis	PCA	-	27
2018	herbs	PCA	-	28
2018	herbs	PCA	OPLS	29

Table S-2. Settings of the flexible parameters for different objectives and datasets

Objectives	**Datasets**	**Hidden units**	**Epochs***
Denoising	Bioautogram	4	50
Qualitative feature extraction	Bioautogram	16	50
	Fluorescence chromatogram (UV 366 nm)	128	50
Quantitative feature extraction	Absorbance chromatogram (UV 254 nm)	8	50

*By chance the number of epochs turned out to be constant.

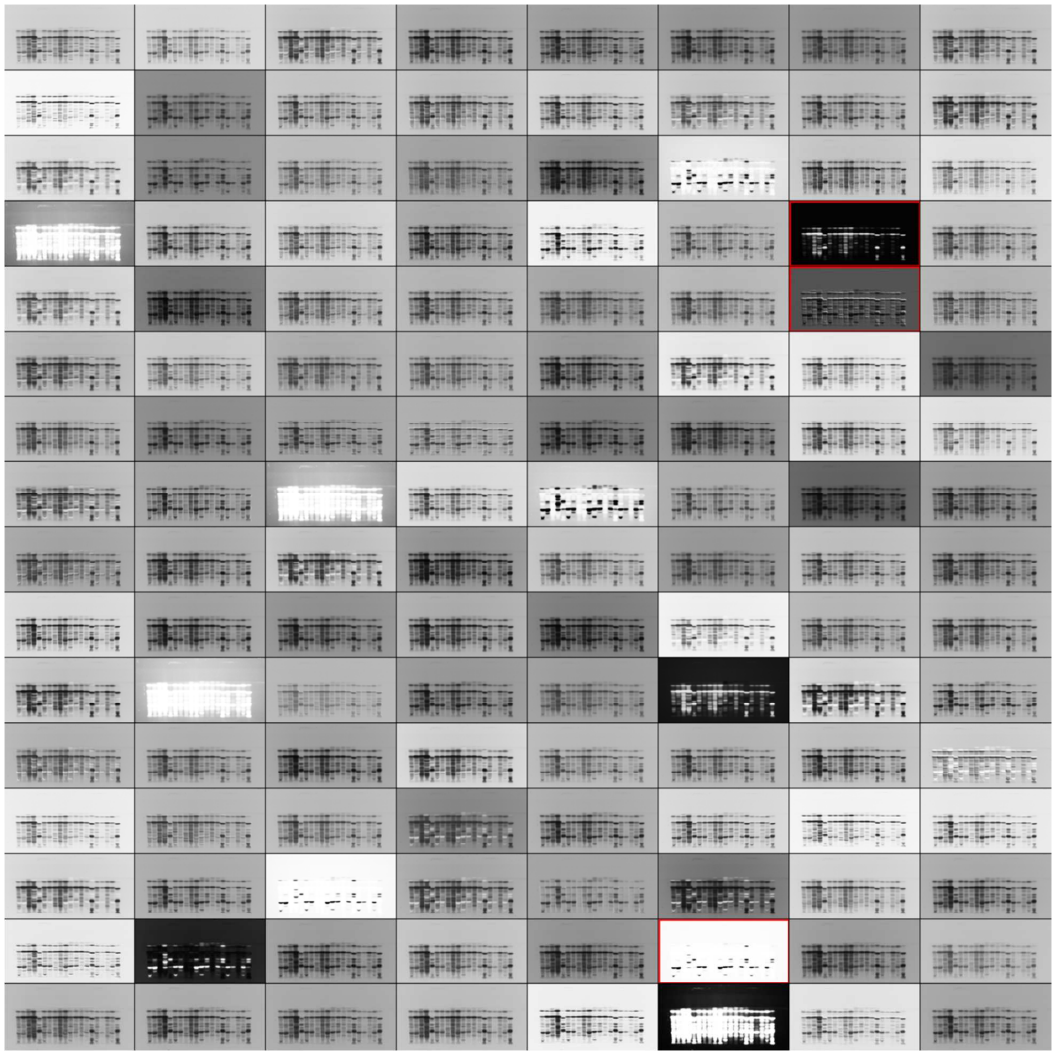


Figure S-1. The 128 extracted features from the model trained with the chromatogram at UV 366 nm; ANN units presented in the manuscript highlighted in red.

References

(1) Tian, R.; Xie, P.; Liu, H. *J. Chromatogr. A* **2009**, 1216, 2150–2155.

(2) Wang, X. H.; Xie, P. S.; Lam, C. W. K.; Yan, Y. Z.; Yu, Q. X. J. *Pharma. Biomed. Anal.* **2009**, 49, 1221–1225.

(3) Komsta, Ł.; Cieślab, Ł.; Bogucka-Kockac, A.; Józefczykd, A.; Kryszeńe, J.; Waksmundzka-Hajnos, M. *J. Chromatogr.* A **2011**, 1218, 2820–2825.

(4) Zarzycki, P. K.; Zarzycka, M. B.; Clifton, V. L.; Adamski, J.; Glod, B. K. *J. Chromatogr. A* **2011**, 1218, 5694–5704.

(5) Sârbu, C.; Moţ, A. C. *Talanta* **2011**, 85, 1112–1117.

(6) Ogegbo, O. L.; Eyob, S.; Parmar, S.; Wanga, Z.-T.; Bligh, S. W. A. *Anal. Methods* **2012**, 4, 2522–2527.

(7) Agatonovic-Kustrin, S.; Loescher, C. M. *Anal. Chim. Acta* **2013**, 798, 103–108.

(8) Audoin, C.; Holderith, S.; Romari, K.; Thomas, O. P.; Genta-Jouve, G. *J. Planar Chromatogr. - Mod. TLC* **2014**, 27, 328–332.

(9) Morlock, G.; Ristivojević, P.; Chernetsova, E. S. *J. Chromatogr. A* **2014**, 1328, 104–112.

(10) Wong, K. H.; Razmovski-Naumovski, V.; Li, K. M.; Li, G. Q.; Chan, K. *J. Pharma. Biomed. Anal.* **2014**, 95, 11–19.

(11) Ristivojević, P.; Andrić, F.; Trifković, J.; Vovk, I.; Stanisavljević, L. Ž.; Tešić, Ž. L.; Dušanka M. Milojković-Opsenica. *J. Chemom.* **2014**, 28, 301–310.

(12) Sagi, S.; Avula, B.; Wang, Y. H.; Zhao, J.; Khan, I. A. *J. Sep. Sci.* **2014**, 37, 2797–2804.

(13) Tang, T. X.; Guo, W. Y.; Xu, Y.; Zhang, S. M.; Xu, X. J.; Wang, D. M.; Zhao, Z. M.; Zhu, L. P.; Yang, D. P. *Phytochem. Anal.* **2014**, 25, 266–272.

(14) Pei-Shan, X.; Shuai, S.; Shunjun, X.; Longgang, G. *J. Chromatogr. Sep. Tech.* **2014**, 5, 249, doi:10.4172/2157-7064.1000249

(15) Mavimbela, T.; Viljoen, A.; Vermaak, I. *J. Appl. Res. Med. Aromat. Plants* **2014**, 1, 8–14.

(16) Agatonovic-Kustrin, S.; Morton, D. W.; Ristivojević, P. *J. Chromatogr. A* **2016**, 1468, 228–235.

(17) Chasset, T.; Häbe, T. T.; Ristivojevic, P.; Morlock, G. E. *J. Chromatogr. A* **2016**, 1465, 197–204.

(18) Maldini, M.; Montoro, P.; Addis, R.; Toniolo, C.; Petretto, G. L.; Foddai, M.; Nicoletti, M.; Pintore, G. *Ind. Crops Prod.* **2016**, 94, 665–672.

(19) Milojković Opsenica, D.; Ristivojević, P.; Trifković, J.; Vovk, I.; Lušić, D.; Tešić, Ž. *J. Chromatogr. Sci.* **2016**, 4, 1–7.

(20) Ristivojević, P. M.; Morlock, G. E. *J. Planar Chromatogr. - Mod. TLC* **2016**, 29, 310–317.

(21) Boka, V. I.; Stathopoulou, K.; Benaki, D.; Gikas, E.; Aligiannis, N.; Mikros, E.; Skaltsounis, A. L. *Phytochem. Lett.* **2017**, 20, 379–385.

(22) Guzelmeric, E.; Ristivojević, P.; Vovk, I.; Milojković-Opsenica, D.; Yesilada, E. *J. Pharma. Biomed. Anal.* **2017**, 132, 35–45.

(23) Hage, S.; Morlock, G. E. *J. Chromatogr. A* **2017**, 1490, 201–211.

(24) Ristivojević, P.; Morlock, G. E. *Food Hydrocolloids* **2017**, 64, 78–84.

(25) Ristivojević, P.; Trifković, J.; Vovk, I.; Milojković-Opsenica, D. *Talanta* **2017**, 162, 72–79.

(26) Shawky, E.; Selim, D. A. *J. Chromatogr. B* **2017**, 1061-1062, 134–138.

(27) Guzelmeric, E.; Ristivojević, P.; Trifković, J.; Dastan, T.; Yilmaz, O.; Cengiz, O.; Yesilada, E. *LWT–Food Sci. Technol.* **2018**, 87, 23–32.

(28) Ibrahim, R. S.; Fathy, H. *J. Chromatogr. B* **2018**, 1080, 59–63.

(29) Liu, X.; Ahlgren, S.; Korthout, H. A.; Salomé-Abarca, L. F.; Bayona, L. M.; Verpoorte, R.; Choi, Y. H. *J. Chromatogr. A* **2018**, 1532, 198–207.

Publication 4

quanTLC, an online open-source solution for video-densitometric quantification

Dimitri Fichou[1] and Gertrud E. Morlock[1]

1. Chair of Food Science, Institute of Nutritional Science, and Interdisciplinary Research Center (IFZ), Justus Liebig University Giessen, Heinrich-Buff-Ring 26-32, 35392 Giessen, Germany

Published in:

Journal of chromatography A (2018) 1560:78-81

Submitted: March 30, 2018, Accepted: May 13, 2018, Available online: May 19, 2018

Journal of Chromatography A, 1560 (2018) 78–81

Contents lists available at ScienceDirect

Journal of Chromatography A

journal homepage: www.elsevier.com/locate/chroma

quanTLC, an online open-source solution for videodensitometric quantification

Dimitri Fichou, Gertrud E. Morlock*

Justus Liebig University of Giessen, Institute of Nutritional Science, Chair of Food Science, and Interdisciplinary Research Center, Heinrich-Buff-Ring 26-32, 35392, Giessen, Germany

ARTICLE INFO

Article history:
Received 30 March 2018
Received in revised form 7 May 2018
Accepted 13 May 2018
Available online 19 May 2018

Keywords:
Online software
Quantitative image analysis
High-performance thin-layer chromatography
Videodensitometry
Open-source software

ABSTRACT

The image is the key feature of planar chromatography. Videodensitometry by digital image conversion is the fastest way of its evaluation. Instead of scanning single sample tracks one after the other, only few clicks are needed to convert all tracks at one go. A minimalistic software was newly developed, termed quanTLC, that allowed the quantitative evaluation of samples in few minutes. quanTLC includes important assets such as open-source, online, free of charge, intuitive to use and tailored to planar chromatography, as none of the nine existent software for image evaluation covered these aspects altogether. quanTLC supports common image file formats for chromatogram upload. All necessary steps were included, *i.e.,* videodensitogram extraction, preprocessing, automatic peak integration, calibration, statistical data analysis, reporting and data export. The default options for each step are suitable for most analyses while still being tunable, if needed. A one-minute video was recorded to serve as user manual. The software capabilities are shown on the example of a lipophilic dye mixture separation. The quantitative results were verified by comparison with those obtained by commercial videodensitometry software and opto-mechanical slit-scanning densitometry. The data can be exported at each step to be processed in further software, if required. The code was released open-source to be exploited even further. The software itself is online useable without installation and directly accessible at http://shinyapps.ernaehrung.uni-giessen.de/quanTLC.

1. Introduction

Quantitative evaluation is an important part of the chromatographic analysis. In planar chromatography, this evaluation is performed using data obtained by either opto-mechanical slit-scanning densitometry of sample tracks or videodensitometry of digitally converted chromatographic images. The latter were recorded by a flat-bed scanner [1–4], camera [5] or light emission diodes [6]. The performance of such approaches has recently been compared [6,7], showing that videodensitometry is a viable option for quantification. Instead of scanning single sample tracks one after the other, as known in conventional densitometry, only few clicks are needed to convert all tracks at one go in videodensitometry, leading to performance data comparable to slit-scanning densitometry [8].

Per se, the videodensitometric analysis is fast and comparatively low-cost. It is the method of choice for laboratories limited in the budget, due to the ubiquitous access to a smart phone, digital camera or flat-bed scanner for image recording and documentation. So far, nine different software have been reported for videodensitometric analyses, *i.e.,* ImageJ [7], VideoScan [8], Sorbfil TLC Videodensitometer [9], Macherey Nagel TLC scanner [10], JustTLC [11], TLC Analyzer [12], qtlc [13], TLSee [14] and Matlab's Imaging Processing Toolbox [15]. However, advantageous assets such as open-source, online, free of charge, intuitive to use and tailored to planar chromatography were only partially covered by some of the existent software for image evaluation (Table 1). For example, TLC Analyzer was written in Matlab and is available free of charge, but a quantification of the data was not integrated in the software. Or the open-source R package qtlc is very rudimental in its functionality and does not have a graphical user interface, which means that it is difficult to operate for a non-programmer. Or the software is not free of charge (VideoScan, Sorbfil TLC Videodensitometer, Macherey Nagel TLC scanner, JustTLC, TLSee and Matlab's Imaging Processing Toolbox). Or other software is not easy to operate (ImageJ, qtlc, and Matlab's Imaging Processing Toolbox) or less tailored to the chromatographic evaluation (ImageJ and Matlab's Imaging Processing Toolbox). So far, only two software (ImageJ

* Corresponding author.
E-mail address: Gertrud.Morlock@uni-giessen.de (G.E. Morlock).

https://doi.org/10.1016/j.chroma.2018.05.027
0021-9673/

Table 1
Comparison of nine existent software for image evaluation with the newly developed quanTLC software regarding important software assets for the HPTLC user.

Software name	Manufacturer or Developer	Important software asset for the HPTLC user				
		Open-source	Online	Free of charge	Tailored	Quantitative
ImageJ	U.S. National Institute of Health, Bethesda, MD, USA	x		x		x
VideoScan	CAMAG, Muttenz, Switzerland				x	x
Sorbfil TLC Video- densitometer	Jsc Sorbpolymer, Krasnodar, Russia				x	x
Macherey Nagel TLC scanner	Macherey Nagel, Düren, Germany				x	peak integration only
JustTLC	Sweday, Sodra Sandby, Sweden				x	peak integration only
TLC Analyzer	A. Victoria, I. Hess [12]			x	x	
qtlc	I. D. Pavicevic [13]	x		x	x	via R
TLSee	AlfaTech, Genova, Italy				x	peak integration only
Matlab's Imaging Processing Toolbox	MathWorks, Natick, MA, USA					via Matlab
quanTLC	D. Fichou, G. E. Morlock	x	x	x	x	x

and qtlc) have been written as open-source software with an open-source code.

This clearly underlines that sophisticated software comprising important assets (open-source, free of charge, online, *i.e.* easily accessible, intuitive, minimalistic, yet tailored to quantitative image evaluation) would be beneficial in the field of videodensitometry. Recently, rTLC [16] was introduced as online open-source software for multivariate data analysis in planar chromatography, based on the R programming language [17]. This software is tailored and focused on the chemometric evaluation of planar chromatograms. Other open-source software explored the potential and possibilities of unsupervised artificial neural network for denoising of chromatograms and feature extraction [18]. Exploring further the new options that offer open-source technologies, newly developed software, termed quanTLC, is demonstrated in this study. It is a minimalistic, yet sophisticated solution for videodensitogram evaluation and its capabilities are discussed and shown on the example of a lipophilic dye mixture separation.

2. Material and methods

2.1. Chemicals and materials

Dye mixture III composed of 6 lipophilic dyes, i.e., Oracet Violet 2R, Solvent Blue 22, Sudan Red G, Solvent Blue 35, Oracet Red G, and Dimethyl Yellow (in ascending hR_F order) [19], was purchased from CAMAG, Muttenz, Switzerland. HPTLC plates silica gel 60 F_{254} (10 cm × 10 cm format, 0.2 mm layer thickness) and toluene (gradient grade) were obtained from Merck, Darmstadt, Germany.

2.2. Planar chromatography

The lipophilic dye mixture III was diluted 1:10 in toluene and applied on an HPTLC plate silica gel 60 F_{254} as 8-mm bands using the Automated TLC Sampler 4 (CAMAG). The applied sample volume range was 0.5–2.5 μL/band. The track distance was 10 mm. The distance from the lower edge was 7 mm and from the side edge 25 mm. Development was performed with 4 mL of toluene up to a migration distance of 60 mm in the Twin-Trough Chamber, 10 cm × 10 cm (CAMAG). The chromatogram was documented under white light illumination using the DigiStore 2 Documentation System (CAMAG).

2.3. Newly developed quanTLC software

Videodensitometric quantification was performed with the newly developed quanTLC software. It was written in the R programming language [17]. The code was released open-source to be exploited even further. The shiny package [20] was used to create a user interface in HTML, directly accessible from an internet browser. Thus, the software itself is online useable without installation and directly accessible at http://shinyapps.ernaehrung.uni-giessen.de/quanTLC. In the form of an R package, instructions for installation and operation are available on GitHub at http://github.com/dimitrif/quanTLC

2.4. Verification of results by software comparison

Densitometry was performed via absorbance measurement with the multi-wavelength scan at 450, 500, 530, and 620 nm of the TLC Scanner 4 and winCATS software, version 1.4.6 (CAMAG) [19]. The scanning speed was 20 mm/s with a measuring slit dimension of 3.0 mm × 0.3 mm. Videodensitometry was performed via the VideoScan software (CAMAG). Parameters used for integration in each of the software are summarized (Table 2). Polynomial regression was selected in all software and applied to all dyes.

3. Results and discussion

3.1. Description of the workflow in quanTLC

In the newly developed software quanTLC, the videodensitometric analysis was structured in five steps, *i.e.* data input, preprocessing, integration, statistical data analysis and data export. A one-minute video serves as illustration (Video S1). As data input, the user had to upload a chromatogram as file. JPEG, PNG, BMP and TIFF are accepted as graphic file formats. As further data input, the plate dimension as well as distances of sample application and chromatographic development were required to extract the individual videodensitograms automatically. The pixel values across a line were averaged to achieve a mean pixel value for each hR_F value. The software is also able to deal with an application from both sides, used for example in high-throughput horizontal development.

Preprocessing improved the compatibility of the videodensitograms for quantitative analysis [21,22]. Several preprocessing options were integrated in quanTLC to be selectable, *i.e.*, negative peak inversion, smoothing with the Savitzky-Golay algorithm [23], baseline correction via the R package baseline [24] and warping for peak alignment with the parametric time warping (ptw) and dynamic time warping algorithms (dtw) via the R packages ptw and dtw [25–27]. During selection of the preprocessing, visual inspection of the resulting videodensitograms in real time helped to choose the best options.

Automatic peak integration was necessary to increase reproducibility and minimize user action. Peak integration was done via the R package pracma [28]. Peak starts and peak ends were detected based on the succession of increasing and decreasing steps before a peak. A minimum peak height was set to remove all peaks of a smaller intensity.

For calibration, quantification and statistical data analysis, the user had to select the standard tracks and input their values (amount per band or target concentration). By clicking on a peak

Table 2
Settings used as integration parameters by the three software compared.

Dye	hR_F	Videodensitometry via quanTLC or VideoScan		Densitometry via winCATS	
		Image or channel	Signal: Peak	Wavelength (nm)	Signal: Peak
Oracet Violet 2R	2	grayscale	area	530	height
Solvent Blue 22	6	red	area	620	area
Sudan Red G	15	blue	area	500	height
Solvent Blue 35	21	red	area	620	area
Oracet Red G	35	blue	height	530	height
Dimethyl Yellow	51	blue	area	450	area

depicted in the RGB channel or grayscale image, signal intensities were automatically extracted via peak area or peak height. Then, regression models were presented that automatically depicted the best regression on those data values as well as the calibration function. The software used the R package EnvStats [29] to build the regression model. A *p*-value of 0.05 was used for the stepwise regression to determine which polynomial model to use, *i.e.,* linear or quadratic. A summary on the regression showed the achieved metrics, *e.g.,* determination coefficient (multiple r-squared and adjusted r-squared) and standard error. Any zones of sample tracks (the ones that were not assigned as standard tracks in the chromatogram) were calculated based on the obtained calibration function. For linear regression models, limit of detection (LOD) and limit of quantification (LOQ) were automatically calculated based on the standard deviation of the response divided by the slope (calibration curve method, Eq. (1)) [30].

$$\mathrm{LOD} = \frac{3.3 \cdot \text{Standard deviation of the response}}{\text{Slope of the calibration curve}} \quad (1)$$

With regard to data export, a final report can be downloaded in PDF, HTML or MS Word format. The data at each step can be exported either as CSV format to be analyzed by another software or as Rdata format to be uploaded in another quanTLC session.

3.2. Verification of quanTLC results

For illustration of the software capabilities and for verification of the results generated, a mixture of lipophilic dyes was separated (Fig. 1a). For calibration, different volumes were applied. Exemplarily, the extracted videodensitogram of track 3 is illustrated that shows the reconstructed chromatogram from the pixel values (Fig. 1b). After preprocessing of the red channel, *e.g.*, negative peak inversion, smoothing and baseline correction, the data values were ready for peak detection, selection and integration (Fig. 1c). A five-point quadratic regression model fitted for Solvent Blue 35 (Fig. 1d).

The generated results were compared with those obtained by existent software, using the same chromatogram of the dye mixture separation. Quantitative results were calculated for each dye using a polynomial regression. The first five tracks were used for the five-point calibration and the last three tracks for calculation of the trueness and repeatability (relative standard deviation) of the quantification of the five dyes. As the determination coefficient of the calibration curve was not available with the VideoScan and winCATS software, it was recalculated with the R package EnvStats [29] and the same trueness and repeatability were found. Comparable results were obtained, which verified the accuracy of the new software (Table 3). The Oracet Violet 2R was close to the application zone, which worsened the trueness and precision. For all other dyes, the results were acceptable with determination coefficients superior to 0.9972, truenesses between 95% and 103% and relative standard deviations below 1.8%. The respective means over all 6 dyes were as good as or even better as those obtained by existent software (Table 3). The chromatogram analyzed with quanTLC, including all regression models and options selected during the analysis, can be loaded as demonstration dataset in the software.

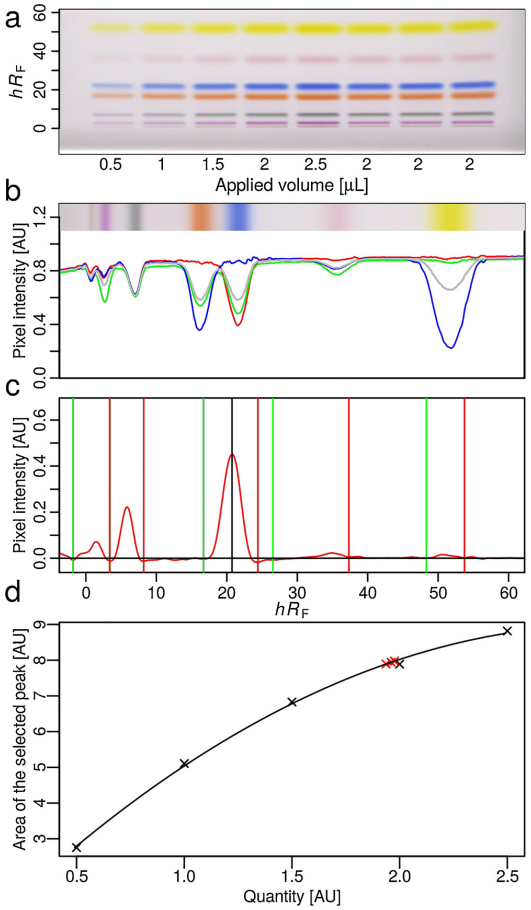


Fig. 1. HPTLC chromatogram of lipophilic dyes (a), overlayed RGB channel and grayscale videodensitograms of track 3 (b), red channel after preprocessing by negative peak inversion, smoothing and baseline correction and peak detection (c, peak start as green bar and peak end as red bar, selected peak as black bar), and exemplarily, calibration curve for Solvent Blue 35 with a determination coefficient of 0.9987 (d). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

4. Conclusions

The newly developed quanTLC software was proven to be tailored to the needs of quantitative evaluation of planar chro-

Table 3
For the same dye mixture chromatogram, comparison of the calibration quality and quantitative results (mean trueness and repeatability) obtained by three different software (quanTLC *versus* VideoScan *versus* winCATS).

	Videodensitometry				Densitometry	
	quanTLC		VideoScan		winCATS	
	R^2	Mean trueness (%) ± %*RSD* (*n* = 3)	R^{2a}	Mean trueness (%) ± %*RSD* (*n* = 3)	R^{2a}	Mean trueness (%) ± %*RSD* (*n* = 3)
Oracet Violet 2R	0.9999	92.9 ± 2.1	0.9947	86.5 ± 3.1	0.9998	93.6 ± 1.0
Solvent Blue 22	0.9995	103.0 ± 0.8	0.9983	101.5 ± 1.7	0.9989	104.7 ± 3.0
Sudan Red G	0.9990	97.3 ± 0.4	0.9983	91.3 ± 2.2	0.9966	98.3 ± 0.8
Solvent Blue 35	0.9987	97.9 ± 1.0	0.9965	99.0 ± 1.5	0.9980	102.6 ± 3.8
Oracet Red G	0.9996	102.2 ± 1.8	0.9972	106.5 ± 2.5	0.9976	100.9 ± 2.1
Dimethyl Yellow	0.9972	94.8 ± 1.6	0.9993	97.3 ± 1.9	0.9972	101.7 ± 3.7
Mean (*n* = 6)	0.9990	98.0 ± 1.3	0.9974	97.0 ± 2.2	0.9980	100.3 ± 2.4

[a]Recalculated with the R package EnvStats [29] to obtain the determination coefficient

matography images. Advantageous assets such as open-source, online, free of charge and tailored to planar chromatography were covered altogether by quanTLC, in contrast to existent software for image evaluation. As features, quanTLC supports the upload of common file formats, videodensitogram extraction, preprocessing options, peak integration, calibration and further statistical data analysis, reporting and data export. The user-interface of quanTLC was designed to be minimalistic and intuitive to use, while fully satisfying relevant needs. Fast chromatogram evaluations were performed in few minutes. Its results were comparable to commercially available solutions for both, videodensitometry and opto-mechanical slit-scanning densitometry. quanTLC is directly useable without installation and online accessible free of charge. A one-minute video serves as user manual. The code is released open-source and can be modified to expand the current possibilities, create new features and verify their functionalities.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.chroma.2018.05.027.

References

[1] J. Stroka, T. Peschel, G. Tittelbach, G. Weldner, R. van Otterdjik, E. Anklam, Modification of an office scanner for the determination of aflatoxins after TLC separation, J. Planar Chromatogr. – Mod. TLC 14 (2001) 109–112.
[2] M.E. Johnson, Rapid, simple quantitation in thin-layer chromatography using a flatbed scanner, J. Chem. Educ. 77 (2000) 368–372, http://dx.doi.org/10.1021/ed077p368.
[3] A. Abbaspour, E. Mirahmadi, A. Khajehzadeh, Disposable sensor for quantitative determination of hydrazine in water and biological sample, Anal. Methods 2 (2010) 349–353, http://dx.doi.org/10.1039/b9ay00291j.
[4] O. Kaynar, M. Ileriturk, A. Hayirli, Evaluation of computational modifications in HPTLC with gel analysis software and flatbed scanner for lipid separation, J. Planar Chromatogr. – Mod. TLC 26 (2013) 202–208, http://dx.doi.org/10.1556/JPC.26.2013.3.1.
[5] S. Babić, D. Mutavdžić, D. Ašperger, A.J.M. Horvat, M. Kaštelan-Macan, Determination of veterinary pharmaceuticals in production wastewater by HPTLC-videodensitometry, Chromatographia 65 (2006) 105–110, http://dx.doi.org/10.1365/s10337-006-0109-2.
[6] D. Fichou, G. Morlock, Office Chromatography: Miniaturized All-in-One Open-Source System for Ultrathin-Layer Chromatography, in submission.
[7] N. Popovic, J. Sherma, Comparative study of the quantification of thin-layer chromatograms of a model dye using three types of commercial densitometers and image analysis with ImageJ, Trends Chromatogr. 9 (2014) 21–28.
[8] G. Morlock, S. Kopacz, Fast and precise SBSE-HPTLC/FLD method for quantification of six polycyclic aromatic hydrocarbons frequently found in water, J. Liq. Chromatogr. Rel. Technol. 31 (2008) 1925–1942, http://dx.doi.org/10.1080/10826070802194856.
[9] I. Kowalska, Ł. Cieśla, T. Oniszczuk, M. Waksmundzka-Hajnos, W. Oleszek, A. Stochmal, Comparison of two TLC-DPPH•-image processing procedures for studying free radical scavenging activity of compounds from selected varieties of *Medicago sativa*, J. Liq. Chromatogr. Rel. Technol. 36 (2013) 2387–2394, http://dx.doi.org/10.1080/10826076.2013.790776.
[10] F. Soponar, A.C. Moț, C. Sârbu, Quantitative determination of some food dyes using digital processing of images obtained by thin-layer chromatography, J. Chromatogr. A 1188 (2008) 295–300, http://dx.doi.org/10.1016/j.chroma.2008.02.077.
[11] R. Johnsson, G. Träff, M. Sundén, U. Ellervik, Evaluation of quantitative thin layer chromatography using staining reagents, J. Chromatogr. A 1164 (2007) 298–305, http://dx.doi.org/10.1016/j.chroma.2007.07.029.
[12] A. Victoria, I. Hess, Digitally enhanced thin-layer chromatography: an inexpensive, new technique for qualitative and quantitative analysis, J. Chem. Educ. 84 (2007) 842–847, http://dx.doi.org/10.1021/ed084p842.
[13] I.D. Pavicevic, Package 'qtlc', 2016 https://cran.r-project.org/web/packages/qtlc/.
[14] C. Witold, D. Katarzyna, W. Robert, H. Adam, Analysis of sulfide ions by densitometric thin-layer chromatography and use of TLSee software, J. Planar Chromatogr. – Mod. TLC 23 (2010) 343–347.
[15] I. Rezic, T. Rolich, Quantitative evaluation of chromatograms from scanned thin-layer chromatography plates, Tekstil-Zagreb 60 (2011) 74–80.
[16] D. Fichou, P. Ristivojević, G.E. Morlock, Proof-of-principle of rTLC, an open-source software developed for image evaluation and multivariate analysis of planar chromatograms, Anal. Chem. 88 (2016) 12494–12501, http://dx.doi.org/10.1021/acs.analchem.6b04017.
[17] R.C. Team, R a language and environment for statisticalcomputing, 2016 https://www.r-project.org/.
[18] D. Fichou, G. Morlock, Powerful artificial neural network for planar chromatographic image evaluation, shown for denoising and feature extraction, Anal. Chem. (2018), http://dx.doi.org/10.1021/acs.analchem.8b01298.
[19] G.E. Morlock, N. Brett, Correct assignment of lipophilic dye mixtures? A case study for high-performance thin-layer chromatography-mass spectrometry and performance data for the TLC-MS Interface, J. Chromatogr. A 1390 (2015) 103–111, http://dx.doi.org/10.1016/j.chroma.2015.02.011.
[20] W. Chang, J. Cheng, J. Allaire, Y. Xie, J. McPherson, Shiny, Web Application Framework Forr, 2016 https://cran.r-project.org/package=shiny.
[21] Ł. Komsta, Suppressing the charged coupled device noise in univariate thin-layer videoscans: a comparison of several algorithms, J. Chromatogr. A 1216 (2009) 2548–2553, http://dx.doi.org/10.1016/j.chroma.2009.01.042.
[22] P.M. Ristivojević, G.E. Morlock, The influence of preprocessing methods on multivariate image analysis in high-performance thin-layer chromatography fingerprinting, J. Planar Chromatogr. – Mod. TLC 29 (2016) 310–317, http://dx.doi.org/10.1556/1006.2016.29.4.10.
[23] A. Stevens, L. Ramirez Lopez, Package 'prospectr', 2014, pp. 1–22 https://cran.r-project.org/package=prospectr.
[24] K. Hovde Liland, B.-H. Mevik, R. Canteri, Package 'baseline', 2015 https://cran.r-project.org/package=baseline.
[25] J. Gerretzen, P. Eilers, H. Wouters, T. Bloemberg, R. Wehrens, Package 'ptw', 2017 https://cran.r-project.org/package=ptw.
[26] P.H.C. Eilers, Parametric time warping, Anal. Chem. 76 (2004) 404–411, http://dx.doi.org/10.1021/ac034800e.
[27] T. Giorgino, Computing and visualizing dynamic time warping alignments in R: the dtw package, J. Stat. Software. 31 (2009) 1–24 http://www.jstatsoft.org/v31/i07.
[28] H.W. Borchers, Package 'pracma', 2018 https://cran.r-project.org/package=pracma.
[29] S.P. Millard, Package 'EnvStats', 2017 https://cran.r-project.org/package=EnvStats.
[30] Int. Conf. Harmonization ICH Topic Q2 (R1) Validation of Analytical Procedures: Text and Methodology 1994 17 (2005).

Publication 5

In silico and in situ signal highlighting in planar chromatography high-resolution mass spectrometry to ease structure elucidation

Dimitri Fichou[1#], Imanuel Yüce[1#] and Gertrud E. Morlock[1]

1. Chair of Food Science, Institute of Nutritional Science, and Interdisciplinary Research Center (IFZ), Justus Liebig University Giessen, Heinrich-Buff-Ring 26-32, 35392 Giessen, Germany

#. The authors D. F. and I. Y. contributed equally to this work.

Submitted: May 1, 2018

In silico and *in situ* signal highlighting in planar chromatography high-resolution mass spectrometry to ease structure elucidation

Dimitri Fichou[a#], Imanuel Yüce[a#] and Gertrud E. Morlock[a*]

[a]Chair of Food Science, Institute of Nutritional Science, and Interdisciplinary Research Center (IFZ), Justus Liebig University Giessen, Heinrich-Buff-Ring 26-32, 35392 Giessen, Germany

*Corresponding author: Gertrud.Morlock@ernaehrung.uni-giessen.de

ABSTRACT: Analyzing unknown compounds present at trace amounts, optimization of the ionization parameters does not guarantee a sufficient response and successful signal assignment. The interpretation is difficult as the mass signals of impurities are often less intense than background signals. Important information may be overlooked. Such critical and time consuming data analysis was overcome by developing a new strategy and open source software called eicCluster offering unsupervised machine learning algorithms and powerful interactive visualization tools that made data processing fast and intuitive. Using eicCluster for high-performance thin-layer chromatography coupled with high-resolution mass spectrometry, mass signals of impurities (degradation products) were highlighted in a stressed formulation, which were hardly found until linked to the new software. Owed to (1) preprocessing leading to intensity-agnostic signals and (2) the tSNE algorithm, clustering mass signals based on their similarity, even compound ions present at low intensities were separated in subclusters from background signals (*in silico* highlighting). The resulting 2D maps allowed a new view on the dataset to target molecules (impurities) in complex mixtures. In addition to this new source of information, targeted on-surface synthesis of degradation products (*in situ* highlighting) was shown to support a fast structure elucidation when standards are not commercially available. It allowed a better understanding of the proposed degradation reactions in the formulation. As proof of principle, an original example formulation, its stressed samples as well as the proposed degradation products of on-surface synthesis were compared. *In silico* and *in situ* signal highlighting substantially eased data processing in structure elucidation.

Hyphenating chromatography with high-resolution mass spectrometry (HRMS), data interpretation is fronted as a main problem. The tremendous amount of data produced during the recording as well as the presence of chemical and electronic interferences challenge the assignment of mass signals. Algorithms that distinguish background from analyte signals are considered as helpful tool being environmentally friendly (no chemical waste) and time-saving (taking only few minutes). Though the use of algorithms (*e.g.*, sequential paired covariance[1], component detection algorithm (CODA)[2], windowed mass selection method[3], matched filtration with experimental noise determination[4] and centWave[5]) was introduced for data interpretation in high-performance liquid chromatography (HPLC)-MS and gas chromatography (GC)-MS[6], it could not be transferred to high-performance thin-layer chromatography (HPTLC)-MS due to its peculiarities. In 2017, the Themis algorithm preprocessed petroleum products measured by ultra HRMS. This software tool was able to work with big data sets and separate peaks from noise signals within minutes[7]. Another metabolic sample was analyzed via HR widescan LC hyphenated to HRMS and the data processed with the R package AssayR[8]. Several open source (*e.g.*, enviMass[9] and XCMS[10]) and vendor software (*e.g.*, TraceFinder™) focused on an automatic processing. However, such existing software dealing with HRMS data was tailored either to GC or HPLC or metabolomic analyses and hardly suited to manage other challenges.

For HPTLC-MS, such immense potential has not been exploited so far[11]. Especially, for the analysis of samples incompatible with HPLC and GC[12,13], HPTLC would be a good solution, especially as the elution head-based TLC-MS interface[14–16] was fully automated recently[17,18], allowing high-throughput analysis and hands-free operation during the elution sequence. Challenges of HPTLC-MS analysis are the chemical background that may cause ion suppression and thus less intense analyte signals[19]. Some of these background signals in MS can be removed by plate prewashing[20]. Other peculiarities of HPTLC-MS were (1) the chronogram, which depicts the zone elution order, *i.e.* eluted analyte and background (blank) zones in the same data file, (2) the time-resolved elution (additional separation dimension) by the different desorption characteristics of molecules during elution of a zone, leading to non-Gaussian elution peaks in the total ion current chronogram, and (3) the coelution of analytes in a zone transferred to HRMS. The latter could be overcome by insertion of an orthogonal column in the transfer line between the elution-based interface and the MS ionization source[21-23].

In this study, the potential of *in situ* and *in silico* signal highlighting in HPTLC-HRMS was investigated on the example of the analysis of degradation products (impurities) in a formulation, as HPTLC is a treasured complementary method to HPLC or GC for impurity analysis [12,24,25]. The structure elucidation of such analytes present at low amounts was hardly possible due to the predominance of other signals. On one hand, the potential of open source unsupervised

machine learning algorithms and interactive visualization tools was investigated to allow or ease the interpretation of HPTLC-HRMS data. On the other hand, on-surface synthesis followed by HPTLC-HRMS was exploited to gather further structural information about potential degradation reactions in complex mixtures for confirmation.

EXPERIMENTAL SECTION

Chemicals and Materials. Methanol was obtained from Carl Roth, Karlsruhe, Germany and ammonia solution (25%) was purchased from Bernd Kraft, Duisburg, Germany. HPTLC plates silica gel 60 F_{254} (20 cm × 10 cm, 0.2 mm layer thickness) and L-ornithine monohydrochloride were provided by Merck, Darmstadt, Germany. D-fructose was received from Sigma-Aldrich, Steinheim, Germany. Double-distilled water was made with Heraeus Destamat Bi-18E (Thermo Fisher Scientific, Schwerte, Germany). The ingredients of a formulation contained *i.a.* sugars (*e.g.* fructose), amino acids (*e.g.* L-ornithine), organic acids and flavoring agents. Methanol (≥99.9%, Optima LC/MS) was bought from Thermo Fisher Scientific.

Preparation of Original and Stressed Formulations. The original formulation (100 mg) was dissolved in 1 mL double-distilled water and transferred into a vial. For the temperature-stress study, the solid original formulation was stored for 24 months at 40 °C and 75-% air humidity. The color of the formulation changed from orange to brown. A 100-mg/mL aqueous solution was prepared (temperature-stressed formulation).

On-Surface Temperature-Stressed Formulation. The original formulation solution (6 µL) was applied on a HPTLC plate and heated on a hotplate at 120 °C for 10 min. The color of the formulation changed from orange to brown.

On-Surface Synthesis of the Degradation Products. For degradation products a and b, aqueous solutions of L-ornithine monohydrochloride (42 mg/mL) and D-fructose (54 mg/mL) were prepared. Both solutions were applied as overspotted bands, *i.e.* 2.5 µL each of L-ornithine monohydrochloride, (105 µg, 0.623 µmol) and D-fructose (135 µg, 0.749µmol, 1.2 eq). The plate was heated at 120 °C on a hotplate for 10 min, and the zone color changed to brown. For degradation products c and d, aqueous solutions of 48 mg/mL citric acid and 54 mg/mL D-fructose were prepared. Citric acid (5 µL, 240 µg, 1.25 µmol) and D-fructose (5 µL, 270 µg, 150 µmol, 1.2 eq) were applied as overspotted bands. The plate was heated at 120 °C on a hotplate for 10 min and the zone color changed to brown.

Traditional Synthesis of Degradation Products a and b. A mixture of 20 mL pyridine, 20 mL glacial acetic acid, 1.37 g L-ornithine hydrochloride (8.1 mmol) and 531 mg zinc powder (8.1 mmol) was stirred for 40 min at room temperature in a round bottom flask. D-fructose (1.73 g, 9.6 mmol, 1.2eq) was subsequently added and the mixture was stirred at room temperature for 48 h. The mixture was filtered and 125 mL acetone were added. The resulting solid was washed with 150 mL acetone and 75 mL diethyl ether. The crude product was dissolved in 25 mL methanol and filtered. After evaporation of methanol, 1.42 g of a yellow-brownish solid were obtained.

HPTLC-UV/vis/FLD. HPTLC plates were dried at 125 C in an oven for 15 min. Samples were applied as 8-mm bands with a dosage speed of 150 nL/s by the Automatic TLC Sampler ATS 4 (CAMAG, Muttenz, Switzerland). The distance between tracks was 17 mm and the distance from the lower edge was set to 10 mm. Application zones were dried in a warm stream of air for 30 s. Development was performed in the Twin Trough Chamber 20 cm × 10 cm (CAMAG). After chamber saturation for 5 min with mobile phase solvent, development was performed with a mixture of methanol - water - ammonia (25%), 15:4:4.5, *V/V/V*, up to 7 cm. The TLC Visualizer (CAMAG) was used to capture the chromatogram at white light illumination (vis) in transmission and reflection modes, at ultra violet (UV) light at 254 nm and at 366 nm. Instruments were controlled and data were evaluated by the winCATS software version 1.4.7.2018 (CAMAG).

HPTLC-HRMS. The Plate Express (Advion, Ithaca, NY, USA), equipped with an oval elution head (cutting edge of 4×2 mm) was used to elute the zones of interest with methanol - water, 6:4, *V/V*, at a flow rate of 0.09 mL/min from the HPTLC plate into the mass spectrometer via an inline filter containing a 0.5-mm stainless steel frit (Upchurch Scientific A-356 and PEEK-Frit Blue UPA-703, Techlab, Erkerode, Germany)[19] The Q Exactive Plus Hybrid Quadrupole-Orbitrap mass spectrometer equipped with a heated electrospray ionization (HESI-II) source (Thermo Fisher Scientific) was operated and data processed by Xcalibur 3.0.63 software (Thermo Fisher Scientific).

Data Analysis. The .RAW files obtained from the HRMS analysis were converted into the .MZXML format with ProteoWizard[26]. A user interface was then designed in R[27]and shiny[28]. The .MZXML file was uploaded and read with the readMzXmlData package[29]. The signal ions below a threshold intensity of $<1\ e^{+6}$ cps were removed. This noise-reduced data set was then transformed from an unstructured (unordonned) list into a matrix compatible with the clusterisation algorithm.

The newly developed eicCluster software used the bucketting method[30,31], *i.e.*, at each time step, all *m/z* values were merged together into a given range (0.01 Dalton). This led to a matrix with one extracted ion chronogram (EIC) per *m/z*. Each EIC signal was normalized by preprocessing with the Standard Normal Variate (SNV) algorithm, which divided each EIC by its standard deviation and subtracted the means[32].

The t-distributed stochastic neighbour embedding (tSNE) algorithm was used on the resulting matrix for dimension reduction[33-35]. The Rtsne package with the Barnes-Hut implementation was used with an initial dimension reduction of 30, perplexity of 3 and maximal iteration of 1000. Instruction for software installation is available on GitHub (https://github.com/DimitriF/eicCluster).

RESULTS AND DISCUSSION

***In Silico* Signal Highlighting by Newly Developed Software**. A strategy and open source software was developed to deal with the high complexity of HRMS data and to ease the interpretation of HPTLC-HRMS data of unknown compound zones present at the trace level. Exemplarily shown for a pharmaceutical formulation, if impurities should be discovered, regular signals of the formulation and its ingredients as well as background signals from the system, plate and solvents used had to be excluded. Hence, new software called eicCluster was developed to help users with data interpretation based on unsupervised machine learning algorithms and interactive visualization tools. Exemplarily, this software was used to point to degradation products in a stressed formulation. The signals of the degradation products

were hardly found until HRMS data were linked to the eicCluster software treatment. Preprocessing and clustering of mass signals was investigated for highlighting such target molecule signals in complex mixtures. The workflow was structured in the following steps, *i.e.*, data input, preprocessing and clustering as well as interactive visualization and output of the cluster (Figure 1).

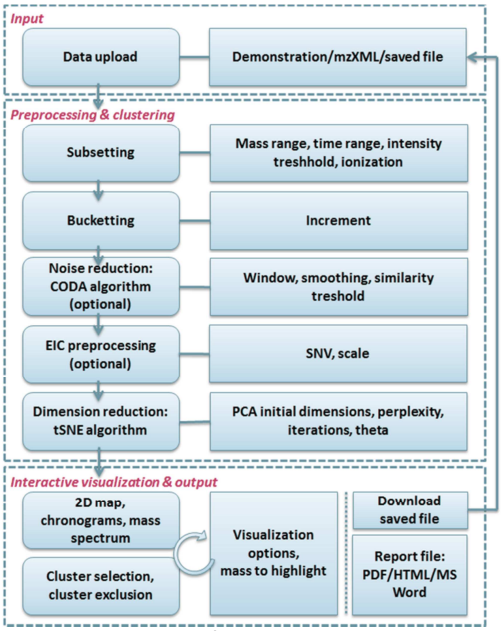


Figure 1. Block-diagram of the strategic operations in the newly developed eicCluster software.

The data input was facilitated by previously developed open source tools[26,29]. For preprocessing, it was important to first reduce the number of considered signals by imposing an intensity threshold. Signals below this threshold were discarded. This speeded up the subsequent processing by bucketting and clustering. The software also allowed to subset the dataset, meaning working with part of the mass signals by selecting the measurement time range and mass range. For example, using only the first 10 min of the experiment or working with signal ions between *m/z* 100 and 500 only. Applying bucketting to HRMS data was simple but came along with a trade-off: If the number of digits after the decimal point, called increment, was too small, a signal could split into several *m/z* values, whereas if it was too big, noise could be added to the signal of interest or two different ions with the same sum formula could join and be not distinguishable. Keeping those aspects in mind during data analysis was important to avoid errors in interpretation. Before the clustering, the SNV algorithm standardized the EICs to ensure that EICs of the same molecule look similar by being independent on signal intensity.

EICs were clustered with the tSNE algorithm, which reduced the high-dimensional matrix to two dimensions. Optimization of the perplexity parameter was most important. It represented the optimal number of neighbors. The number of iterations was important, but its optimization was not crucial as the map did not change much after typically 1000 iterations. The effects of these two parameters on the clustering were exemplarily shown (Figure S-1). Principal component analysis was tested as alternative. The resulting 2D map was more difficult to interpret (Figure S-2), if compared to that of the tSNE algorithm.

Once the clustering was finished, the user was able to interactively explore a map were each point corresponded to an *m/z* EIC. By brushing, clicking and double clicking on this map, it was possible to zoom, de-zoom and select a cluster of interest to plot the corresponding EICs and mass spectra and compare them to the ones obtained by Xcalibur. Finally, once a cluster corresponding to a signal of interest was found, it was possible to save it for later reporting in a PDF, HTML or MS Word format.

***In Situ* Signal Highlighting.** The verification and proof of concept of *in silico* signal highlighting were made by *in situ* signal highlighting. The original example formulation (Figure 2A-D, track 1) was compared to three stressed samples. One formulation was stored stressed by temperature and humidity (at 40 °C and 75% air humidity for 24 months, Figure 2A-D, track 2) and another original example formulation was heated on the HPTLC layer (stressed on-surface) forcing degradation products (Figure 2A-D, track 3). As third sample for comparison, degradation products were synthesized on the HPTLC layer (micromole on-surface synthesis) from the proposed ingredients L-ornithine and D-fructose, which were oversprayed on the same start zone and heated (Figure 2A-D, track 4). Practically, tracks 3 and 4 were applied first and heated on a plate heater at 120 °C for 10 min. The colors of the zones changed from colorless to brownish under white light illumination, so did the fluorescence color at UV 366 nm (Figure 2B). Then, the original sample was applied on track 1 and the sample stored stressed on track 2 (Figure 2C). The HPTLC plate was developed under standardized conditions and the compound zone at hR_F 34 was online eluted into the HRMS (Figure 2D).

Proof of Concept. The HRMS data were analyzed and the main output was the 2D map containing clustered masses (Figure 2E). The TIC (Figure 2F) showed the online eluted samples as well as the rinsing cycles performed in between. The processing parameters were previously optimized to get well separated and defined cluster structures (Figure S-2). The formulation samples with known ingredients and unknown degradation products challenged the algorithm. The expectation was to cluster together such masses, which represented the same substance in one logical cluster. Several clusters were presented in the map, but only four are discussed in the following. These mass clusters represent: (1) background signals from the system, (2) and (3) signals from ingredients of the original sample and (4) the degradation product **a**. The masses and attribution in each of those clusters were summarized in Table 1. For example, the EIC of the selected background cluster (Figure 2G) showed a signal that did not increase during the sample zone elutions. The selected mass spectrum showed a base peak at *m/z* 100.08 assigned to be *N*-methyl-2-pyrrolidone, a floor stripper (Figure 2H). There were several other masses in the same cluster, but they were not assigned to substances.

The first cluster coming from the formulation was assigned to the ingredient L-ornithine. The EIC profile was highly similar to the one of the TIC and counted for 85% of the

intensity (Figure 2G). In the reconstructed mass spectrum, the molecular ion, the isotopes, the fragment after ammonia loss and its dimer together with the ^{13}C isotope were found (Fig 2J, Table 1).

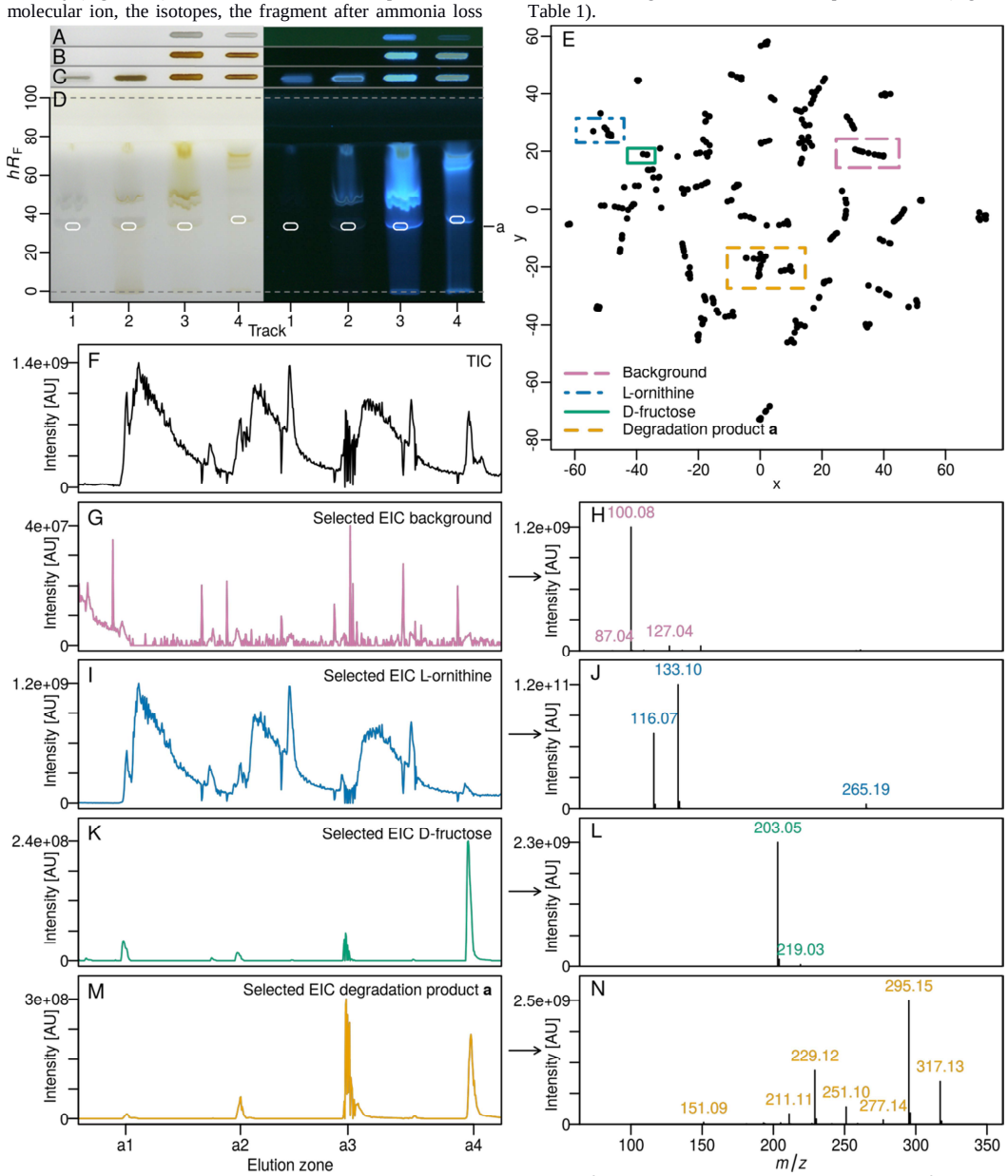


Figure 2. HPTLC plate images illuminated with white and UV light (366 nm) after sample application (A), heating at 120 °C for 10 min (B), further sample application after heating (C), development (D), loaded with original formulation (track 1), formulation stored stressed (track 2), on-surface stressed formulation (track 3), on-surface synthesized degradation product (track 4); 2D map after tSNE clusterization of mass signals (E); chronograms of TIC (F), selected EIC cluster (G, I, K, M) and the corresponding mass spectra of the selected cluster (H, J, L, N).

The second cluster coming from the formulation was assigned to the ingredient D-fructose. The EIC profile was more specific with signals only appearing during sample zone elutions (Figure 2K). In the mass spectra, the sodium adduct as base peak and its ^{13}C isotope were present (Figure 2L, Table 1).

Table 1. Assignment of the calculated masses (m/z_{calc}) to the substances detected by HRMS and observed (m/z_{obs}) after EIC clustering.

Sum formula	Signal assignment	m/z_{calc}	m/z_{obs}
Ingredient: L-ornithine			
$C_5H_{12}N_2O_2$	$[M+H]^+$	133.09715	133.10
$C_5H_{12}N^{15}NO_2$	$[M+H]^+$	134.09419	134.09
$C_4{}^{13}CH_{12}N_2O_2$	$[M+H]^+$	134.10051	134.10
$C_5H_{12}N_2O^{18}O$	$[M+H]^+$	135.10140	135.10
$C_5H_9NO_2$	$[M+H\text{-}NH_3]^+$	116.07060	116.07
$C_4{}^{13}CH_9NO_2$	$[M+H\text{-}NH_3]^+$	117.07396	117.07
$C_{10}H_{25}N_4O_4$	$[2M+H]^+$	265.18703	265.19
$C_9{}^{13}CH_{25}N_4O_4$	$[2M+H]^+$	266.19039	266.19
Ingredient: D-fructose			
$C_6H_{12}O_6$	$[M+Na]^+$	203.05261	203.05
$C_5{}^{13}CH_{12}O_6$	$[M+Na]^+$	204.05596	204.06
Degradation product **a**			
$C_{11}H_{22}N_2O_7$	$[M+Na]^+$	317.13192	317.13
$C_{10}{}^{13}CH_{22}N_2O_7$	$[M+Na]^+$	318.13528	318.14
$C_{11}H_{22}N_2O_7$	$[M+H]^+$	295.14998	295.15
$C_{10}{}^{13}CH_{22}N_2O_7$	$[M+H]^+$	296.15333	296.15
$C_{10}H_{20}O_6N_2$	$[M+H\text{-}H_2O]^+$	277.13941	277.14
$C_{10}H_{16}N_2O_4$	$[M+Na\text{-}CH_6O_3]^+$	251.10023	251.10
$C_{10}H_{16}N_2O_4$	$[M+H\text{-}CH_6O_3]^+$	229.11828	229.12
$C_{10}H_{14}N_2O_3$	$[M+H\text{-}CH_8O_4]^+$	211.10772	211.11
Background: N-methyl-2-pyrrolidone			
$C_5H_{10}NO$	$[M+H]^+$	100.07569	100.08

The last cluster was assigned to the proposed degradation product **a** (Figure 2M). The EIC cluster showed low signal intensity in the original sample and a higher intensity in the stressed sample (Figure 2D, track 2). The signal increased further after stressing the sample on-surface (track 3) and on-surface synthesis (track 4). The EIC profile of the degradation product **a** (Figure 2M) was close to the one of the D-fructose (Figure 2K), but the algorithm managed to separate both in the 2D map (Figure 2E). In the mass spectra (Figure 2N), the molecular ion, sodium adduct of the degradation product **a,** and the fragments after one, two and three water losses were found, in addition to a loss of formaldehyde. Thus, the *m/z* 229.12 was assigned to two water losses and a loss of formaldehyde. This fragment was also observed as sodium adduct.

The confirmation of the proposed sum formula of degradation product **a** was done via the isotopic pattern, MS^2 experiments and on-surface synthesis of the Maillard reaction degradation product. By processing the mass spectra with Xcalibur, the degradation product **a** was not present in the original example formulation, and only a low intensity signal at *m/z* 295.15012 was present in the stressed sample (Figure S-3). The cleanest spectrum was received after µg-scaled on-surface synthesis with the starting materials L-ornithine and D-fructose. The predicted sum formula of the degradation product **a** coming from the stressed formulation was consistent to those after synthesis. On-surface synthesis highlighted the mass signal (*in situ* highlighting), which made it easy to interpret HRMS data. The signal ion at *m/z* 295.15 of the degraded sample stored stressed and of the on-surface synthesis was fragmented (Figure 3). The MS^2 spectra showed a similar fragmentation pattern to the cluster in the 2D map (Figure 2E) with especially the water losses (Figure 2N, Table 1). The absence of the mass 229.12 could be explained as being an in source fragmentation not observed in the MS^2 experiment.

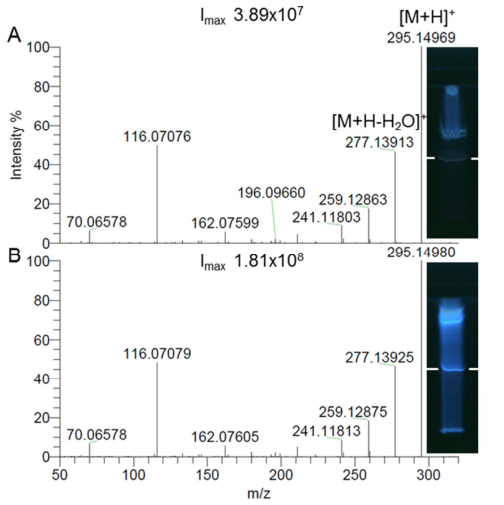


Figure 3. MS^2 with mass isolation at *m/z* 295.15 ± 0.4 and normalized collision energy (10 NCE) of the degraded formulation stored stressed (A) and after on-surface synthesis with L-ornithine and D-fructose (B).

This concept was used for the further degradation products **b**, **c** and **d** discovered in the formulation. The Maillard reaction product **b** was found in the stressed sample, on-surface synthesis sample and also in a traditionally synthesized sample (using L-ornithine and D-fructose, Figure S-4). The protonated molecule $[M+H]^+$ of **b** was measured to be 457.20285±0.1 ppm (Figure S-5). The degradation products **a** and **b** (proposed structures[36] in Figure S-5) were separated in two different clusters in the 2D map (Figure S-4). The degradation products **c** and **d** were found in the stressed sample (Figure 4). However, both molecules were not clearly separated as shown in the previous maps. This can be explained by the low amount of degradation product **d**, for which this software was at its limit. The eicCluster separated both in the 2D map subcluster, and they were confirmed by on-surface synthesis by reaction of citric acid with D-fructose (Figure S-6), most likely according to an esterification mechanism. Degradation product **c** was found at *m/z* 353.07278±0.7 ppm $[M-H]^-$ and product **d** at *m/z* 515.12608±1.4 ppm $[M-H]^-$ (proposed structures in Figure S-6).

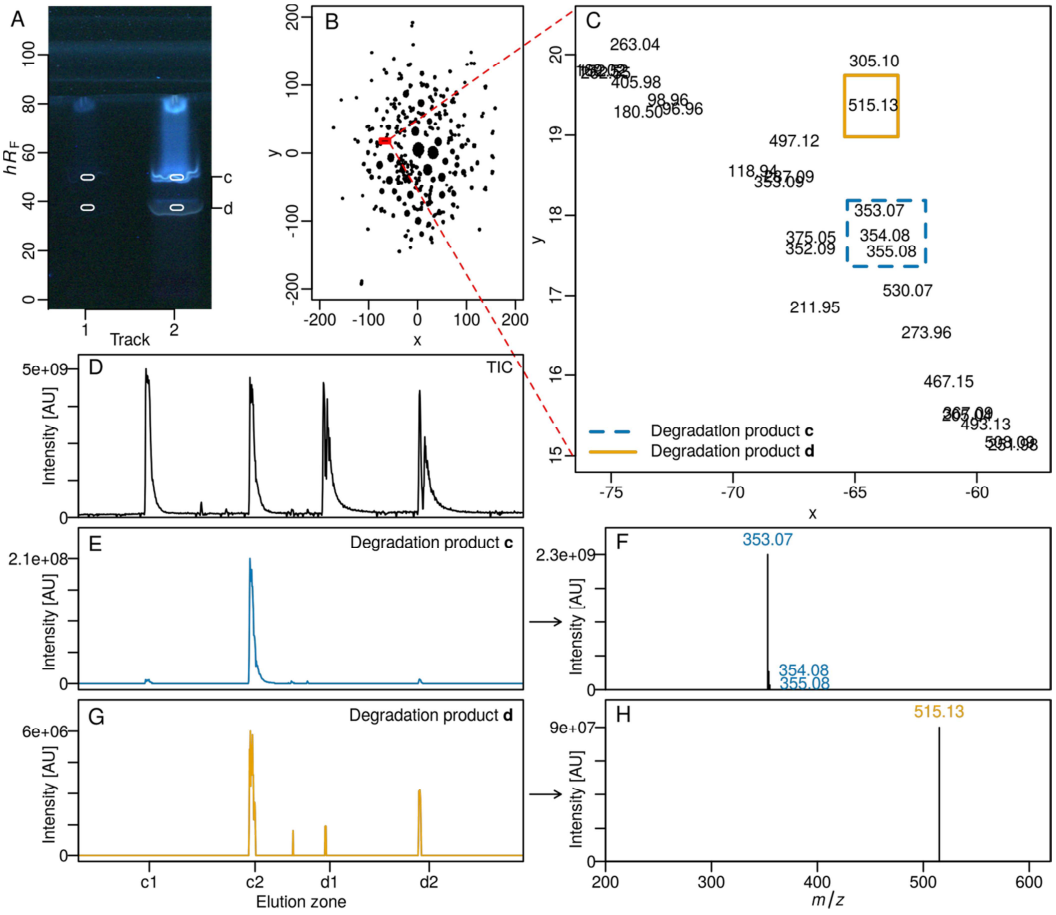


Figure 4. Signal highlighting via t-SNE clusterisation for degradation products **c** and **d**: HPTLC chromatogram at UV 366 nm (A) showing the degradation products **c** and **d** in the original formulation (track 1, 6 µL, 100 mg/mL) *versus* on-surface stressed formulation (track 2, 5 µL, 100 mg/mL). After elution of the respective zones into HRMS (D, TIC chronogram) and t-SNE clusterisation of the mass signals obtained, the 2D map of all mass signals (B and C zoomed) as well as EIC chronograms of responsible clusters for both degradation products (E and G) and the corresponding mass spectra (F and H).

CONCLUSIONS

In silico and *in situ* signal highlighting in HPTLC-HRMS successfully pointed to molecules present at low amounts in two complex mixtures compared. Though HRMS in combination with on-surface synthesis is a powerful tool for structure elucidation, the big data created during an experiment can be overwhelming and render the mass signal interpretation difficult, especially when the signal of interest has a low intensity. We showed that signal highlighting in HRMS can be conducted after preprocessing by using an algorithm as an electronic filter and for clustering of mass signals. This way, the molecular ion, adduct ions, dimer ions *etc.* of known or unknown compounds as well as background signals were found rapidly in one HRMS run, providing a fast identification of the impurities in the formulation.

Most of the existing software for HRMS data management are not dedicated to HPTLC. Hence, a new software called eicCluster was developed to help users with data interpretation based on unsupervised machine learning algorithms and interactive visualization tools. This software was proven to properly assign degradation products in a stressed formulation. Note that the signals of the degradation products were hardly found until HRMS data were linked to the eicCluster software treatment. To further confirm the origin of the new degradation products found and explain its formation reaction, on-surface synthesis was hyphenated to HRMS and also submitted to the data analysis software. The proposed Maillard reaction between D-fructose and L-ornithine was performed within 10 min on the HPTLC surface. A comparison of the corresponding mass spectra at equal hR_F values was consistent. Surface synthesis was successfully used to highlight the mass signals of the targeted degradation product. The new concept led to the discovery of three further new degradation products and was verified by this.

MS^2 experiments of the degradation product zones of the sample stored stressed and the respective on-surface synthesis sample revealed the same fragmentation patterns. The tSNE algorithm assigned some fragments in the full scan spectra that were confirmed by MS^2 fragmentation. This open source eicCluster software was created to support research in the field of HPTLC-HRMS, when the amount of data is high and the processing is too difficult and time-consuming for a scientist.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website.
Figure S-1. Compilation of 32 different t-SNE clusters using different values for iteration and perplexity, and thereof, highlighted parameters used; Figure S-2. Principal component analysis of the EIC mass signal matrix and mass signals of interest zoomed; Figure S-3. Comparison of three HPTLC-HRMS mass spectra of the degradation product a in original sample, temperature-stressed stored sample and sample of on-surface synthesis; Figure S-4. Signal highlighting via t-SNE clusterisation for degradation products a and b; Figure S-5. HPTLC-HRMS spectra and respective chemical structures of the traditionally synthesized degradation products a and b; Figure S-6 HPTLC-HRMS mass spectra and respective chemical structures of the degradation product c in original sample and temperature-stressed stored sample as well as degradation product d in original sample, temperature-stressed stored sample and sample of on-surface synthesis. (DOCX)

AUTHOR INFORMATION

Corresponding Author

*Corresponding author. Tel.: +49-641-99-39141; fax: +49-641-99-39149; E-mail address: Gertrud.Morlock@uni-giessen.de (G. E. Morlock).

Author Contributions

#The authors I. Y. and D. F. contributed equally to this work.

Notes

The authors declare no competing financial interest.

REFERENCES

(1) Muddiman, D. C.; Rockwood, A. L.; Gao, Q.; Severs, J. C.; Udseth, H. R.; Smith, R. D.; Proctor, A. *Anal. Chem.* **1995**, *67*, 4371–4375.

(2) Windig, W.; Phalp, J. M.; Payne, A. W. *Anal. Chem.* **1996**, *68*, 3602–3606.

(3) Fleming, C. M.; Kowalski, B. R.; Apffel, A.; Hancock, W. S. *J. Chromatogr. A* **1999**, *849*, 71–85.

(4) Andreev, V. P.; Rejtar, T.; Chen, H. S.; Moskovets, E. V.; Ivanov, A. R.; Karger, B. L. *Anal. Chem.* **2003**, *75*, 6314–6326.

(5) Tautenhahn, R.; Bottcher, C.; Neumann, S. *BMC Bioinf.* **2008**, *9*, 1–16.

(6) Hilario, M.; Kalousis, A.; Pellegrini, C.; Müller, M. *Mass Spectrom. Rev.* **2006**, *25*, 409–449.

(7) Gavard, R.; Rossell, D.; Spencer, S. E.; Barrow, M. P. *Anal. Chem.* **2017**, *89*, 11383–11390.

(8) Wills, J.; Edwards-Hicks, J.; Finch, A. J. *Anal. Chem.* **2017**, *89*, 9616–9619.

(9) Martin Loos. blosloos/enviMass: Utilities to Process Mass Spectrometry (LC-MS) Data for Environmental Trend Analysis, 2016.

(10) Smith, C. a; Want, E. J.; O'Maille, G.; Abagyan, R.; Siuzdak, G. *Anal. Chem.* **2006**, *78*, 779–787.

(11) Morlock, G. In *Reference module in chemistry, molecular sciences and chemical engineering*; J. Reedijk, Ed.; Elsevier: Waltham, USA, 2014; pp 101–121.

(12) Yüce, I.; Morlock, G. *J. Chromatogr. A* **2016**, *1469*, 120–127.

(13) Rejšek, J.; Vrkoslav, V.; Vaikkinen, A.; Haapala, M.; Kauppila, T. J.; Kostiainen, R.; Cvačka, J. *Anal. Chem.* **2016**, *88*, 12279–12286.

(14) Luftmann, H. *Anal. Bioanal. Chem.* **2004**, *378*, 964–968.

(15) Alpmann, A.; Morlock, G. *Anal. Bioanal. Chem.* **2006**, *386*, 1543–1551.

(16) Morlock, G.; Jautz, U. *J. Planar Chromatogr. - Mod. TLC* **2008**, *21*, 367–371.

(17) Luftmann, H.; Aranda, M.; Morlock, G. *Rapid Commun. Mass Spectrom.* **2007**, *21*, 3772–3776.

(18) Häbe, T. T., G. Morlock, Automated hyphenation of HPTLC to DART-MS and ESI-MS, Poster P-66, International Symposium for HPTLC, Berlin, 2017.

(19) Morlock, G. *J. Liq. Chromatogr. Rel. Technol.* **2014**, *37*, 2892–2914.

(20) Glavnik, V.; Vovk, I.; Albreht, A. *J. Chromatogr. A* **2017**, *1482*, 97–108.

(21) Morlock, G.; Schwack, W. *J. Chromatogr. A* **2010**, *1217*, 6600–6609.

(22) Mroczek, T. *J. Pharma. Biomed. Anal.* **2016**, *129*, 155–162.

(23) Schwack, W.; Pellissier, E.; Morlock, G. *Anal. Bioanal. Chem.* **2018**, *DOI: https://doi.org/10.1007/s00216-018-0945-6*.

(24) Yüce, I.; Morlock, G. E. All on one HPTLC plate: solvent-free on-surface synthesis, workup and online HRMS analysis for structure elucidation of impurities, *in submission*.

(25) Lomax, S. Q.; Learner, T. *J. Am. Inst. Conserv.* **2006**, *45*, 107–125.

(26) Chambers, M. C.; Maclean, B.; Burke, R.; Amodei, D.; Ruderman, D. L.; Neumann, S.; Gatto, L.; Fischer, B.; Pratt, B.; Egertson, J.; Hoff, K.; Kessner, D.; Tasman, N.; Shulman, N.; Frewen, B.; Baker, T. a; Brusniak, M.-Y.; Paulse, C.; Creasy, D.; Flashner, L.; Kani, K.; Moulding, C.; Seymour, S. L.; Nuwaysir, L. M.; Lefebvre, B.; Kuhlmann, F.; Roark, J.; Rainer, P.; Detlev, S.; Hemenway, T.; Huhmer, A.; Langridge, J.; Connolly, B.; Chadick, T.; Holly, K.; Eckels, J.; Deutsch, E. W.; Moritz, R. L.; Katz, J. E.; Agus, D. B.; MacCoss, M.; Tabb, D. L.; Mallick, P. *Nat. Biotechnol.* **2012**, *30*, 918–920.

(27) R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2018.

(28) Chang, W.; Cheng, J.; Allaire, J.; Xie, Y.; McPherson, J. Shiny, web application framework for r, 2016.

(29) Gibb, S. readMzXmlData: Reads Mass Spectrometry Data in mzXML Format, 2015.

(30) Stolt, R.; Torgrip, R. J. O.; Lindberg, J.; Csenki, L.; Kolmert, J.; Schuppe-Koistinen, I.; Jacobsson, S. P. *Anal. Chem.* **2006**, *78*, 975–983.

(31) Cappadona, S.; Levander, F.; Jansson, M.; James, P.; Cerutti, S.; Pattini, L. *Anal. Chem.* **2008**, *80*, 4960–4968.

(32) Barnes, R. J.; Dhanoa, M. S.; Lister, S. J. *Appl. Spectrosc.* **1989**, *43*, 772–777.

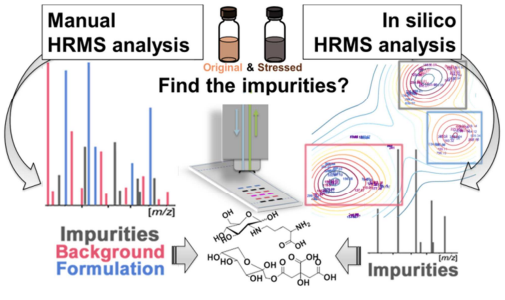
(33) Maaten, L. V. D. *JMLR Proc. vol. 5 (AISTATS)* **2009**, 384–391.

(34) Van Der Maaten, L. J. P.; Hinton, G. E. *J. Mach. Learn. Res.* **2008**, *9*, 2579–2605.

(35) Krijthe, J.; Maaten, L. van der. Package 'Rtsne', 2017.

(36) Srinivas, S. M.; Harohally, N. V. *J. Agric. Food Chem.* **2012**, *60*, 1522–1527.

For TOC only


Manual
HRMS analysis
In silico
HRMS analysis
Original & Stressed
Find the impurities?
[m/z]
Impurities
Background
Formulation
[m/z]
Impurities

Publication 5 - Supporting information

Supporting information

In silico and *in situ* signal highlighting in planar chromatography

high-resolution mass spectrometry to ease structure elucidation

Dimitri Fichou[a#], Imanuel Yüce[a#] and Gertrud E. Morlock[a*]

[a]Institute of Nutritional Science and Interdisciplinary Research Center (IFZ), Justus Liebig University Giessen, Heinrich-Buff-Ring 26-32, 35392 Giessen, Germany

[#]The authors I. Y. and D. F. contributed equally to this work.

*Corresponding author. Phone: +49-641-99-39141; Fax: +49-641-99-39149;

E-mail: Gertrud.Morlock@uni-giessen.de.

Table of contents

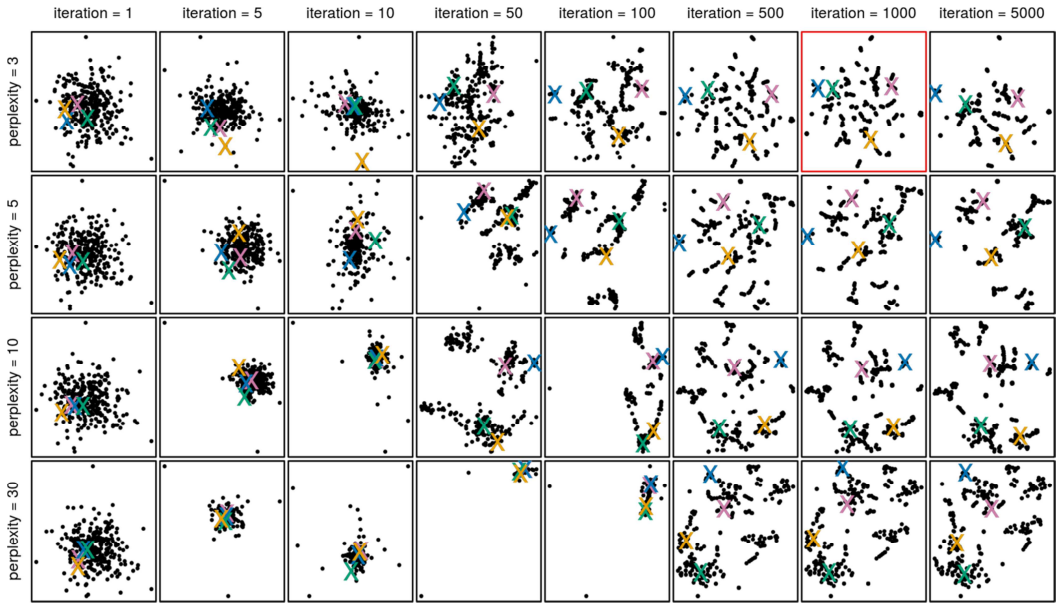


Figure S-1. Compilation of 32 different t-SNE clusters using different values for iteration (x-axis) and perplexity (y-axis), and thereof, highlighted parameters used (framed red; initial dimension 30, max. iteration 1000, perplexity 3) with mass signals of interest (crossed) derived from background (pink), L-ornithine (blue), D-fructose (green) and degradation product **a** (yellow).

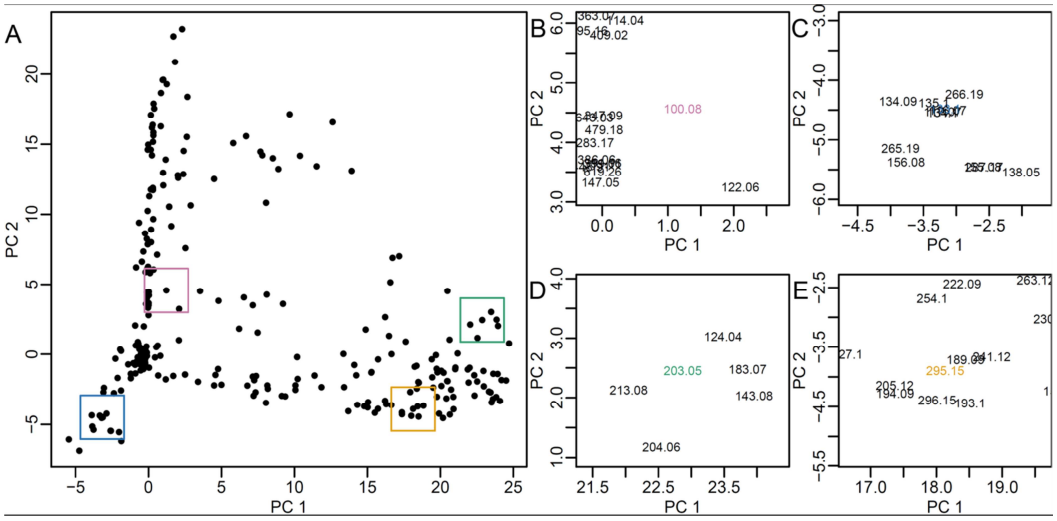


Figure S-2. Principal component analysis of the EIC mass signal matrix (A) and zoom on mass signals of interest (B-E) derived from background (pink), L-ornithine (blue), D-fructose (green) and degradation product **a** (yellow).

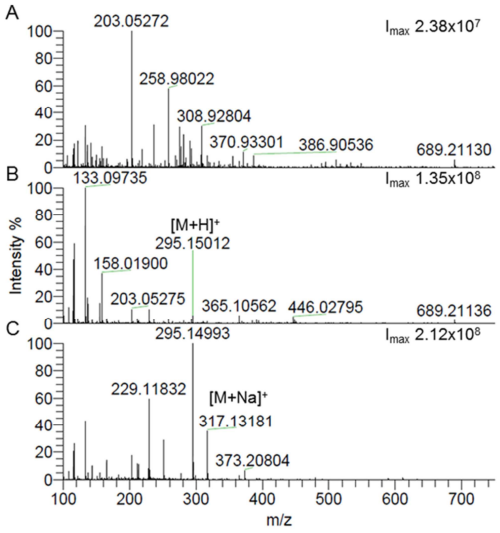


Figure S-3. Comparison of three HPTLC-HRMS mass spectra of the degradation product **a** in original sample (A), temperature-stressed stored sample (B) and sample of on-surface synthesis with L-ornithine and D-fructose (C).

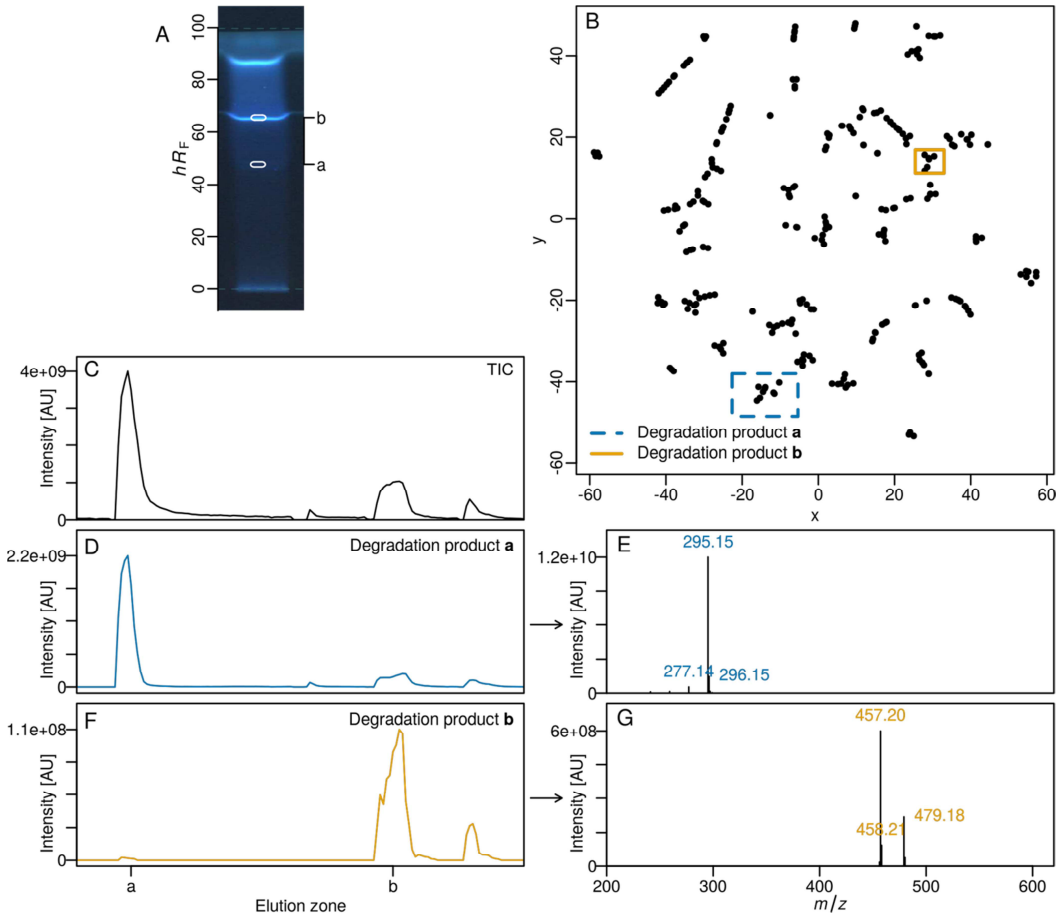


Figure S-4. Signal highlighting via t-SNE clusterisation for degradation products **a** and **b**: HPTLC chromatogram at UV 366 nm (A) showing the traditionally synthesized degradation products **a** and **b** (25 µL, 10 mg/mL). After their elutions into HRMS (C, TIC chronogram) and t-SNE clusterisation of both mass spectra, the 2D map of all mass signals (B) as well as EIC chronograms of responsible clusters for both degradation products (D and F) and the corresponding mass spectra (E and G).

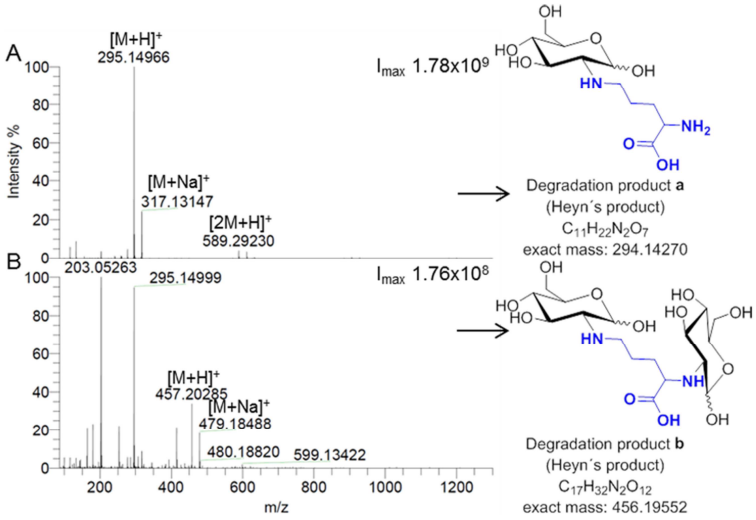


Figure S-5. HPTLC-HRMS spectra and respective chemical structures of the traditionally synthesized degradation products **a** (A) and **b** (B).

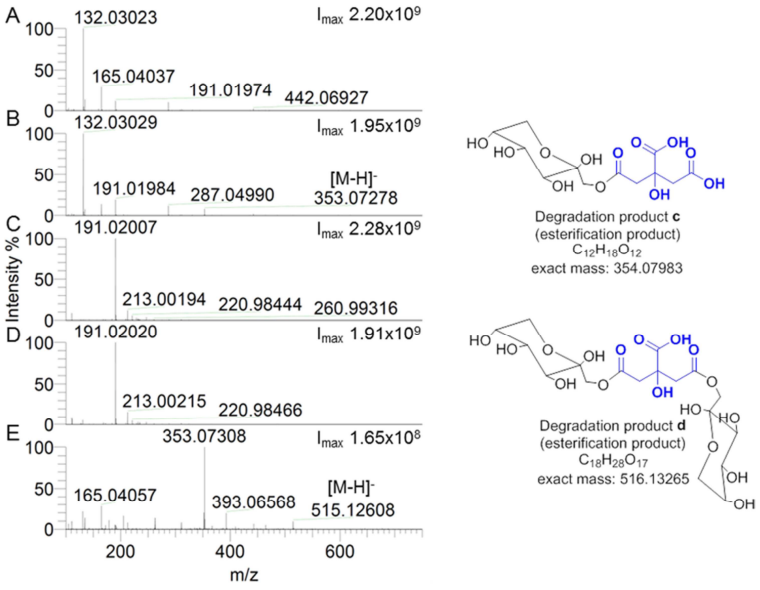


Figure S-6. HPTLC-HRMS mass spectra and respective chemical structures of the degradation product **c** in original sample (A) and temperature-stressed stored sample (B) as well as degradation product **d** in original sample (C), temperature-stressed stored sample (D) and sample of on-surface synthesis with L-ornithine and citric acid (E).

Publication 6

Office Chromatography: Miniaturized All-in-One Open-Source System for Ultrathin-Layer Chromatography

Dimitri Fichou[1] and Gertrud E. Morlock[1]

1. Chair of Food Science, Institute of Nutritional Science, and Interdisciplinary Research Center (IFZ), Justus Liebig University Giessen, Heinrich-Buff-Ring 26-32, 35392 Giessen, Germany

Submitted: June 25, 2018

Office Chromatography: Miniaturized All-in-One Open-Source System for Planar Chromatography

Dimitri Fichou[a] and Gertrud E. Morlock[a,*]

[a]Chair of Food Science, Institute of Nutritional Science, and Interdisciplinary Research Center (IFZ), Justus Liebig University Giessen, Heinrich-Buff-Ring 26-32, 35392 Giessen, Germany

*Corresponding Author: fax +49-641-99-39149; Gertrud.Morlock@uni-giessen.de

ABSTRACT: Current high-performance thin-layer chromatography instrumentation is offline and stepwise automated. However, a moderate miniaturization offers many advantages, and together with the transfer of modern print and media technologies to the field of chromatography (Office Chromatography) it opens up new avenues. This is demonstrated in an all-in-one open-source system developed for planar chromatography, and especially, ultrathin-layer chromatography. Using the InkShield board to control a thermal inkjet cartridge, picoliter drops were sharply printed at a resolution of 96 dpi on the adsorbent layer. Using Marlin, a popular firmware in 3D printing, a Cartesian movement of the print head was made possible for full control on the printing process. Open-source software was developed to control the device in each operation step. Sample solutions and mobile phase were inkjet-printed, exemplarily shown for the analysis of dye or paraben mixture solutions. Light-emitting diodes (LEDs) were investigated for documentation. For example, deep UV LEDs gave access to 254-nm light and RGB LEDs to the visible light range. Calibration functions with correlation coefficients superior to 0.999 were obtained by videodensitometry. The developed modular open-source hardware was compact (26x31x26 cm^3), light (<3 kg) and affordable (810 Euro). For the given analyses, the footprint of needed current instrumentation was miniaturized by a factor of 9. The highly reduced material design complies with Green Chemistry and Lean Laboratory. The design and instruction to reproduce the all-in-one open-source system were made freely available at https://github.com/OfficeChromatography. It is intended to boost progress and understanding by the nature of *do it yourself*.

The concept of Office Chromatography (OC) was introduced in 2010[1,2] to address and merge miniaturization in planar chromatography, both concerning the apparatus and the adsorbent layer. Thin-layer chromatography (TLC) dates back to 1938,[3] and high-performance thin-layer chromatography (HPTLC) to 1975.[4] HPTLC triggered the use of adsorbents of a smaller particle size and a narrower particle distribution along with state-of-the-art instrumentation.[5] Unlike to gas chromatography or high-performance liquid chromatography, automatization was not covering all steps of an analysis in a single system, but implied the use of several automated devices to perform the analysis. The latest idea of a fully automatized planar chromatography system was patented in 2001,[6] nonetheless a prototype was never shown. This one and other all-in-one approaches of the past decades[5] missed the notion of plate miniaturization and the integration of innovative technologies,[2] not used in the field of chromatography so far. Hence, traditional perspectives for automatization stranded.

In 2001, Merck introduced ultrathin-layer chromatography (UTLC) via binder free and monolithic silica gel layers.[7,8] In the following decade, further approaches for UTLC layers were proposed.[9–13] Most UTLC layers were more fragile and the spray-on sample application was abrading the layer material. Also current instrumentation for development and densitometry were oversized for the UTLC plate sizes. Hence, innovative techniques had to be explored, especially in a disruptive, interdisciplinary thinking to open up new avenues, other than the traditional perspectives, and fill the current instrumental gap for UTLC.

Inkjet printing allows a contact-less dosage of liquids via drop-on-demand application.[14] In planar chromatography, inkjet printing was first demonstrated in 2007.[15] For derivatization, the nínhydrin reagent was homogeneously inkjet-printed on the hR_F window of taurine in energy drinks separated by HPTLC. The results obtained were as precise as for traditional derivatization by immersion of the plate in the reagent. Next, inkjet-printing was successfully demonstrated for the sharp and accurate application of sample solutions on UTLC layers.[1,16] Inkjet-print of solutions was a novel approach in the field of chromatography, which was realized by demounting and modification of commercially available low-cost printers. However, the full operational control on the print head of mainstream printers was limited and considered as an obstacle to progress. Thus, focus was laid on open-source technologies that allow modifications.

Additive manufacturing, commonly called 3D printing is gaining interest in the analytical chemistry laboratory.[17–19] In particular, open-source RepRap 3D printers offer an affordable solution (https://reprap.org).[20] In addition to its price, one of the main advantages of those printers is their modularity due to their open-source aspect. The electronics and software are easily modified and can be re-purposed. For the hardware, the philosophy behind the RepRap apparatus is that they could

print their own parts, making them a great inspiration for those willing to develop new automated devices. Those 3D printers are controlled with GCODE, a numerical control programming language which instruct the machine to move, turn on temperature and to use its tools installed. The open-source aspect of 3D printing was important for its progress. The technique went from patent-protected and company-owned to an affordable mainstream technology. The modification of a RepRap system for laboratory usage was demonstrated several times. Liquid chromatography and fluorescence detection were hyphenated with the help of a Prusa i3 3D printer for assisted laboratory automation.[21] Or, a Printrbot 3D printer was modified for automated sample preparation.[22] In our previous study, a slurry doser was designed to be mounted on a Prusa i3 3D printer for production of TLC/HPTLC layers in planar chromatography[13]. Based on this experience, the same 3D printer was further modified for first sample application and development approaches with an inkjet print head installed.[23] Still in the field of planar chromatography, but outside of the RepRap concept, layers were also produced with a polyjet printer[24] and a smartphone-based visualization chamber was 3D printed.[25]

For detection, flatbed scanners were demonstrated as an alternative to conventional densitometers in planar chromatography.[16,26–29] In contrast to typically used light bulbs or tubes, light-emitting diodes (LEDs) present several advantages over conventional sources of illumination, *e.g.* a longer lifetime, narrower emission spectra and low price. Dated back to 1973,[30] its use in analytical chemistry was recently reviewed.[31,32] For example, the use of deep UV LEDs at 235 nm was investigated for liquid chromatography.[33] Or, blue LEDs were used in combination with a Raspberry Pi minicomputer and its camera for fluorescence detection in capillary electrophoresis.[34]

In this study, a newly built miniaturized all-in-one open-source system for planar chromatography with focus on UTLC is presented, evaluated and compared to current state-of-the-art instrumentation. It is the first example of an Office Chromatography system that is able to perform the main steps of the planar chromatography workflow in a single system, *i.e.* application, development and detection. It is exemplarily shown for separations of preservatives and food dyes on three different adsorbents, *i.e.* normal phase silica gel (NP; only application), middle polar cyanopropyl phase (CN) and reversed phase (RP).

Experimental Section

Chemicals and Materials. Sodium sulfate as well as preservatives propyl *p*-hydroxybenzoate (propyl paraben, PP, purity ≥99.0%), ethyl *p*-hydroxybenzoate (ethyl paraben, EP, purity ≈99.0%), methyl *p*-hydroxybenzoate (methyl paraben, MP, purity ≥99.0%) were purchased from Sigma-Aldrich, Buchs, Switzerland. Brilliant black BN (E151) was purchased from Riedel-de Haen, Seelze, Germany. Fast yellow AB (E105, purity ≈95%), Acid blue (E131, purity ≈50–51% pure, also Patent blue V), Sunset yellow FCF (E110), Indigo carmine (E132, also Indigotine), and Ponceau 4R (E124, purity ≈80%, also Cochenille red A) were purchased from Ringe and Kuhlmann, Hamburg, Germany. Azorubine (E122) was obtained from Schuhmann and Son, Karlsruhe, Germany. Bidestilled water was produced by a Heraeus Destamat Bi-18E, Thermo Fisher Scientific, Schwerte, Germany. HPTLC plates silica gel 60 F_{254} (NP), HPTLC plates silica gel 60 CN F_{254} (CN), HPTLC plates silica gel 60 RP 18 W (RP), prototype UTLC plates (NP-UTLC, 4 µm particles, 75 µm layer thickness), benzoic acid (BA, purity >99.9%) and all solvents (HPLC quality) were purchased from Merck, Darmstadt, Germany.

Standard Solutions. The preservatives were dissolved together in methanol (2 µg/µL) and diluted with water (100 ng/µL). The water-soluble food dyes were dissolved together in water as follows: E110, E124, E132 and E151 (482, 414, 1648 and 312 ng/µL, respectively) for RP system and E122, E131 and E105 (541, 214 and 1363 ng/µL, respectively) for NP system.

Hardware of the OC System. The newly built apparatus was designed in OpenSCAD (http://openscad.org). The bill of material was compiled (Table S-1). For self-mounting and reproduction, an assembly instruction (Instruction S-1) together with an assembly video (Video S-1) are available at https://github.com/OfficeChromatography/OCLab. Briefly, the frame was made of aluminum profiles. 3D printing with polylactic acid (PLA; ColorFabb, Belfeld, The Netherlands) on a Prusa i3 MK2 3D printer (Prusa Research, Praha, Czech Republic) was used to produce the functional parts. The InkShield board was purchased from NerdCreationLab (Mount Vernon, WA, USA).[35] Second hand HP C6602 ink cartridges were purchased on ebay (San Jose, CA, USA). The cartridges were sawn to remove the ink, the sponge and to freely access the reservoir. They were rinsed with deionized water prior to use. Deep UV LEDs at 254 nm were purchased from Crystal IS (Green Island, NY, USA) and RGB LEDs from Conrad (Hirschau, Germany).

Firmware of the OC System. The apparatus was controlled by an Arduino Mega 2560 microcontroller board (https://arduino.cc) associated with a RAMPS 1.4 shield (http://reprap.org/wiki/RAMPS_1.4). The cartridges were controlled with the InkShield board. The Marlin firmware (http://marlinfw.org) was uploaded on the Arduino. Modifications had to be made to control the InkShield board by the Arduino as follows: A new GCODE command M700 for control of the InkShield was introduced before[36] and new arguments were added for finer control, *i.e.* argument I for drop per dot and argument L for control of the pulse length. The latter was kept constant at 5 µs.

Software of the OC System. To create GCODE files and send these to the apparatus, a dedicated open-source software called OC_manager was developed in R,[37] and it used the shiny package to create a user interface.[38] The software used the package reticulate[39] to call the python software Printrun (http://www.pronterface.com) and communicate with the Arduino. The software was hosted on a Raspberry Pi 3 single-board computer (https://raspberrypi.org) connected to the local network. As software manual, a standard operating procedure was written (Instruction S-2). The software is freely available at https://github.com/OfficeChromatography/OC_manager.

Analysis of Dye and Preservative Mixtures. Unless stated otherwise, the solutions as well as respective mobile phases were printed with the OC system controlled by the OC_manager software. For the sample application, microliter volumes of the respective sample solution (ca. 30 µL) were filled in the sawn cartridge. The dosage speeds were adjusted by setting the speed of the cartridge movement to 10 mm/s and 2 to 10 drops per dot were used (I value). No delay between paths was used. The volume was then adjusted via the number

of printing paths. The resulting dosage speed roughly averaged by the software was between 10 and 20 nL/s.

For the development by inkjet-printing, ca. 1.7 mL-volumes of mobile phase were filled in the sawn cartridge. A counter glass plate (11 x 5 cm, 1 mm thick) was installed in the plate holder to simulate a closed environment. This dosage speed was controlled indirectly by settings multiple parameters. The speed of the cartridge movement was 20 mm/s, the number of drop per dot was 40 (I value) and the length of the path was 98 mm. The mobile phase was printed on each way (instead of single way). No delay between paths was used, as well as no increasing factor of the delay at each path (that would reduce the dosage speed during the analysis). The resulting dosage speed roughly averaged by the software was 1.5 µL/s.

The preservative mixture solution was used for application and development studies. For the first study on the sample application quality, 0.3-1.5 µL/band were inkjet-printed on a 5x10 cm^2 CN plate. The band length was 6 mm, track distance 8 mm, distance to lower edge 8 mm and first application position 7 mm. As only the application quality was intended to be compared, the plate was conventionally developed with ethanol – water - acetic acid, 14:40:0.1, *V/V/V*[40] up to a migration distance of 50 mm (ca. 12 min), as described in the subsequent comparison section. For the second study on the development, 0.1-0.6 µL/band were inkjet-printed on a 10x5 cm^2 CN plate. The band length was 2 mm, track distance 3.5 mm, distance to middle of the plate 3 mm and first application position 5 mm. The previously mentioned mobile phase was inkjet-printed for development up to a migration distance of 25 mm (ca. 10 min).

The food dye mixture solution was inkjet-printed on a 10x5 cm^2 RP plate (0.1-0.6 µL/band of E110, E124, E132 and E151). The band length was 4 mm, track distance 6 mm, distance to middle of the plate 3 mm and first application position 8 mm. The mobile phase consisting of methanol – water containing 5% sodium sulfate, 3:4 *V/V*,[41] was inkjet-printed for development up to a migration distance of 25 mm (ca. 10 min).

Another food dye mixture solution (0.2-1.0 µL/band of E122, E131 and E105) was inkjet-printed on a 10x5 cm^2 NP-UTLC plate. The band length was 6 mm, track distance 8 mm, distance to lower edge 8 mm and first application position 13 mm. As described in the subsequent comparison section, the plate was conventionally developed with ethyl acetate – methanol – water - acetic acid, 65:23:11:1 *V/V/V/V*,[42] up to a migration distance of 25 mm (ca. 4 min).

Comparison to Status Quo. The common spray-on application via the Automated TLC Sampler 4 (ATS4) was used as status-of-the-art reference. For comparison, the same application parameters were selected as used for inkjet printing. Vertical development was performed in the Twin-Trough Chamber (10x10 cm^2 or 10x5 cm^2). Chromatograms were documented at UV 254 nm or white light illumination in the automatic reflectance mode using the TLC Visualizer. Respective densitograms (absorption measurement) were recorded at the maximum wavelength of preservatives at 254 nm using the TLC Scanner 4. The measuring slit dimension was 4.0 mm × 0.3 mm. The scanning speed was 20 mm/s. Data recording and evaluation was performed with winCATS software, version 1.4.7. All instrumentation used for the comparison was from CAMAG, Muttenz, Switzerland.

Results and Discussion

The self-replicating aspect of the 3D printers was a great source of inspiration. The modification of a RepRap 3D printer for instrumental use in analytical chemistry was demonstrated for 3D printing of TLC/HPTLC layers.[13] This gain in knowledge was applied and expanded to design and built a miniaturized, all-in-one open-source system for planar chromatography with special focus on UTLC (Figure 1).

Hardware of the OC System. The newly built apparatus was designed in OpenSCAD. The selection of aluminum profiles as frame, also used for several RepRap 3D printers (https://reprap.org), made the system highly modular and ensured the required rigidity. Preferably, parts that are easily available were selected to facilitate the reproduction of the OC system and to keep the price low. Not available, tailored functional parts (most black parts in Figure 1) were printed with PLA on a Prusa i3 MK2 3D printer. The relevant OpenSCAD files for self-printing are freely available at Github (https://github.com/OfficeChromatography/OCLab). All required materials for self-mounting including suppliers and prices were compiled in the bill of material (Table S-1). Focus was also laid on the assembly time needed, which could be done in one day. In order to motivate for the self-mounting and assist in the assembly, a 30-min video illustrates the assembly of the OC system (Video S-1).

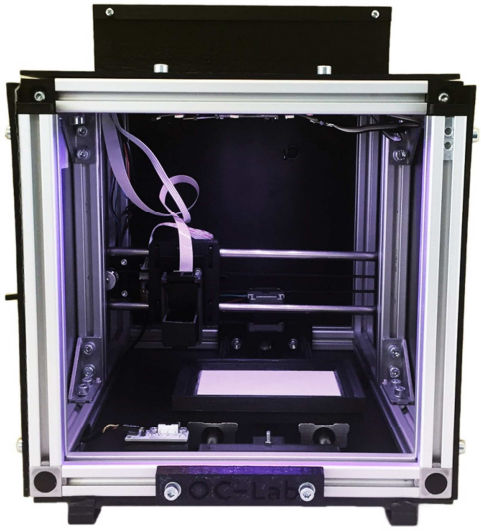


Figure 1. Development of an open-source OC system for planar chromatography that is compact (26 x 31 x 26 cm^3), light (< 3 kg) and affordable (810 Euro).

Second hand HP C6602 ink cartridges were swan, emptied, rinsed and filled with the solutions for dosage. Mostly 30 µL of the sample solutions were filled in the reservoir for sample application or 1.7 mL of the mobile phase mixture for development. A zoom on the swan HP C6602 cartridge shows the reservoir, the nozzles and the electronic connections (Figure 2A). An important plastic part was the plate holder system which required numerous iterations until a versatile version was found that was compatible with several modes of application and development. This plate holder supported the use of

10x5 cm^2 and 10x10 cm^2 plates and had sufficient room for a counter glass plate to cover the mobile phase migration during the development, preventing its evaporation. The 10x5 cm^2 plates were used for parallel development from the plate middle in two opposite directions, which doubled the sample throughput. Along the plate holder, a reservoir to be filled with liquids was integrated for an optional control of the gas phase during development or increase of the humidity for printing bioassays on the plate. The plate holder was hold by a magnet for a fast loading and was made compatible with the compartment for the bioassay plate of the Bioluminizer (CAMAG), as a more efficient (expensive) optical system for luminescence recording has not been integrated in the current OC system so far.

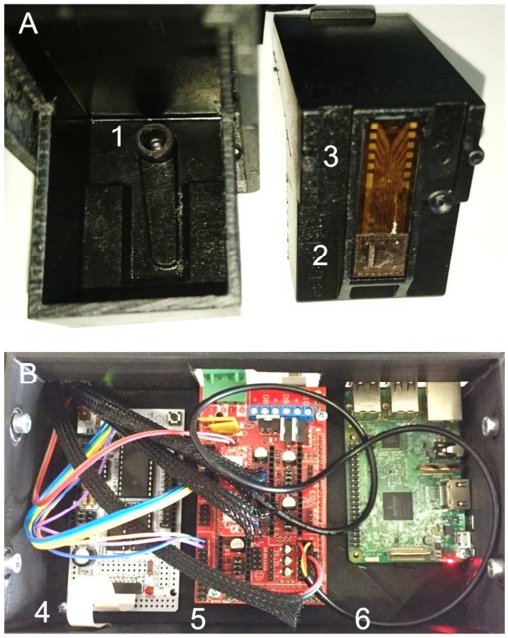


Figure 2. Zoom on a swan HP C6602 (A) showing the reservoir (1), the nozzles (2) and the electronic connections (3) as well as electronic part of the open-source OC system (B) showing the InkShield board (4), the Ramps 1.4 with the Arduino Mega 2560 microprocessor board (5) and the Raspberry Pi single-board computer (6).

For the electronic part, the RepRap environment was once again highly valuable. Three different boards were used (Figure 2B). The InkShield board was able to control each of the 12 nozzles of the HP C6602 cartridge by applying selectively a 20 V voltage pulse for a given time, typically 5 µs. The diameter of each nozzle was about 60 µm. During this pulse, a resistance in the nozzle was heated forming a gas bubble, which collapsed at the end of the pulse, releasing the drop.[14] The cartridge itself was driven above and along the plate surface for the different analytical steps. This was done with the Ramps 1.4 shield associated with the Arduino Mega 2560 microprocessor board loaded with the Marlin firmware. The board also controlled the InkShield drop release and the LEDs. The full range of given functionalities of the board has not been exploited so far, *e.g.*, heater, fan or servo motors were not controlled and only two stepper motors out of five were used. This means that more is possible with this set-up. Finally, the Raspberry Pi board hosted the user interface to control the device by sending GCODE files to the Arduino and capturing the chromatograms with the Raspberry Pi camera.

OC_manager Software. Inspired by the RepRap project and especially the octoprint software (https://octoprint.org/), a new software named OC_manager was developed in the R programming language.[37] The shiny package was used to create a user interface that is intuitively to use.[38] This control software was hosted on a Raspberry Pi single-board computer. Connected to the local network, it was accessed via a mounted screen on the Raspberry Pi, or via the network on another computer. The choice of the Raspberry Pi facilitated the software installation and deployment, as just the image of the operative system had to be loaded on the secure digital (SD) card of the Raspberry Pi. The use of this minicomputer was also important for detection and documentation of the chromatograms, as a mini camera was already available for it.

Several features expected from software for analytical chemistry were integrated, *e.g.*, user login, save method file order, and log files. The software was created to be user-friendly with all the steps available in a simple layout (Figure S-1). Methods were created for each step of the analytical workflow owed to the use of GCODE files as standard to communicate with the system. As input by the user, important parameters were made selectable in a table (Figure S-1). A visual representation of the actual printing was automatically plotted and important information like print volume and dimensions were displayed as feedback to the user (Figure S-1). The software did not aim to comply with regulations like good manufacturing practice, but to give the user utmost freedom on instrumental control. For example, the chromatograms were directly available as JPEG format, and the function to create the GCODE file can be investigated and modified. This way, new steps could easily be implemented for new user cases as well as further GCODE controlled devices.

Inkjet-Print of Sample Solutions. First, the inkjet technology installed in this OC system was used for sample application. For evaluation of its suitability, it was exemplarily applied for the analysis of preservatives and compared to the state-of-the-art spray-on application of liquids as aerosol. This comparison was important for evaluation of the drop size and the accuracy obtainable by the inkjet technology. The sharpness of the application zones was found comparable between spray-on application (Figure 3A) and inkjet print (Figure 3B). The nominal resolution achievable by the inkjet print was 96 dpi through the nozzles (Figure 2A), *i.e.* 0.26 mm per dot. However, the applied bands were larger depending on the diffusion, applied volume and dosage speed. The drop per dot I value was the most important to obtain sharp bands and a small value was recommended to avoid diffusion. The dosage speed estimated by the software was 10 to 20 nL/s, which was by a factor of 10 slower than spray-on application of typically 100 to 200 nL/s. The inkjet technology was found to be a comparatively soft application technique, if compared to spray-on application with its aerosol beam formation and intensive gas-liquid contact. The drop-by-drop printed solution was applied comparatively more on the surface of the layer. This may improve the capability of detection, especially suited for UTLC layers with comparatively lower sample loadings.

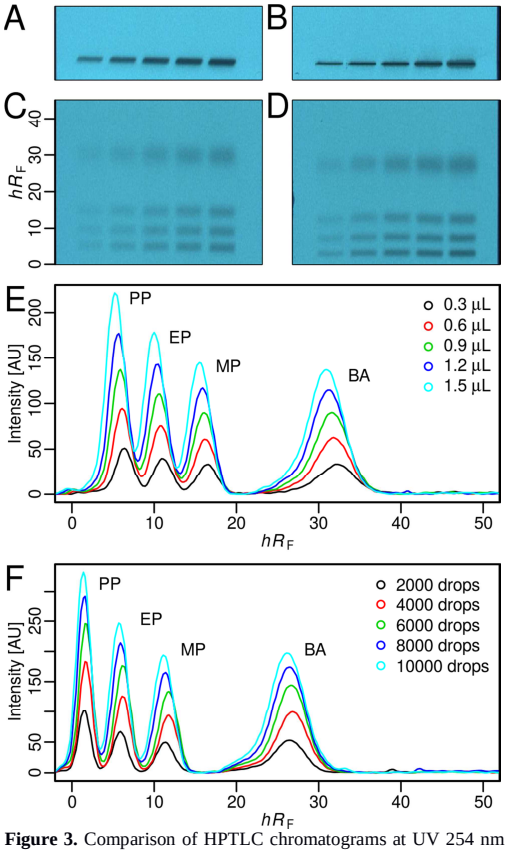


Figure 3. Comparison of HPTLC chromatograms at UV 254 nm of preservatives applied on CN plate by spray-on dosage (A; 0.3-1.5 µL/band using the ATS4) versus inkjet-print (B; 2000-10000 drops/band equal to *ca.* 0.3–1.5 µL/band) as well as after their respective developments (C versus D) with densitograms at 254 nm (E versus F).

After separation, the obtained UV 254 nm chromatograms (Figure 3C *versus* D) and densitograms of the two application techniques (Figure 3E *versus* F) were found to be equivalent. The linearity of different inkjet-printed concentration levels achieved coefficients of determination superior to 0.999. Comparing the slopes of the calibration curves allowed to determine the drop volume. In order to verify the reproducibility of the drop volume between the individual nozzles of ca. 60 µm diameter, a further experiment was conducted without development. Peak integration was made via videodensitometry. Four out of the 12 nozzles were individually activated for printing and compared (Figure S-2). The mean drop volume was 139 pL ± 17 pL (n = 4).

Inkjet-Print of Mobile Phase. This was the most challenging step to implement in the OC system (Figure 4A). Exemplarily, the mobile phase was inkjet-printed for the separation of preservatives (Figure 4B) and water-soluble food dyes (Figure 4C). In order to demonstrate the high sample throughput that is achievable on a small 10x5 cm^2 plate, the preservative mixture solution was inkjet-printed 50 times arranged on two lines in the plate middle (each line 3 mm from plate middle, spaced 6 mm). In between both application lines, the mobile phase was inkjet-printed for 10 min, and by this generated flow, the development and separation of the mixture started (Figure 4B). Mobile phases mainly containing water, methanol and ethanol were preferred to be compatible with the currently integrated HP C6602 ink cartridge that was limited in its solvent resistance. Such solvents were suited best as mobile phase for separations on middle polar plates like CN plates or RP plates (Figure 4B and C).

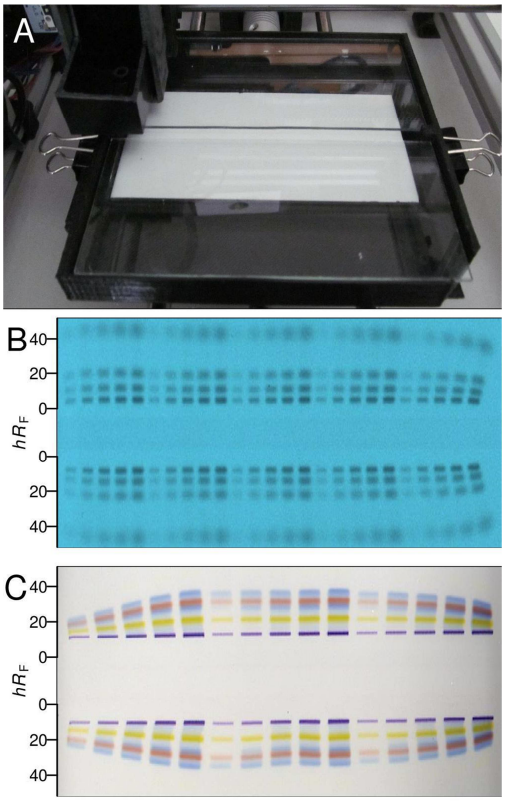


Figure 4. Inkjet-print of different mobile phases: Print in one direction with the installed plate holder capable for 5x10 cm^2 or 10x10 cm^2 plates and counter glass plate to cover the mobile phase migration (A) as well as anti-parallel development from the middle to opposite directions of preservatives on CN plate (B; in ascending hR_F order: PP, EP, MP and BA) and of food dyes on RP plate (C; in ascending hR_F order: E151, E110, E132a, E124, E132b), for both 800-4000 drops/band, ca. 0.1-0.6 µL/band.

The dosage speed of the mobile phase had to be adjusted to the capacity of the layer (layer thickness) to avoid a flooding of the stationary phase with the mobile phase. The 800 µs delay necessary for the printing nozzle to rest between drop firings was capping the dosage speed to 2 µL/s without considering the cartridge movement. This maximal dosage speed was still above the capacity of TLC/HPTLC layers and could even be decreased for printing of the mobile phase. A good flow of the mobile phase was obtained by printing it in the middle of a 10x5 cm^2 plate for parallel development from the plate middle

in the two opposite directions. However, a significant border effect was observed (Figure 4B and C). Altogether, the chamber tightness/fitting to better control the gas phase and the mobile phase print parameters were crucial for the reproducibility of the separation, which is solved and described in detail in a separate study.

Detection and Documentation. The Raspberry Pi single-board computer was already the main controller for this device. Its camera could easily be controlled from the software and unlike other cameras, several options were available directly from the command line, *e.g.*, the International Organization for Standardization (ISO) speed and exposure time. LED sources were preferred for the illumination, as conventional light tubes were not compatible with the desired miniaturization and cost-efficiency of the system. RGB LEDs were available, and more recently (still more expensive), also deep UV LEDs at 254 nm.

The homogeneity of the illumination of the plate and the lens effect of the camera were the main concerns during the implementation of the documentation step. For the illumination, a wide angle deep UV LED was selected and an algorithmic correction was necessary. In order to evaluate the homogeneity, 25 bands (each 500 ng/band) of the preservative mixture solution were printed on a 10x10 cm^2 NP plate and evaluated in a first prototype for the illumination study (Figure S-3) to be integrated into the OC system later. The resulting images and pixel values were compared for different ISO speeds and exposure times, with and without algorithmic correction (Figure 5). For comparison, the plate image was also captured with a state-of-the-art commercial instrument.

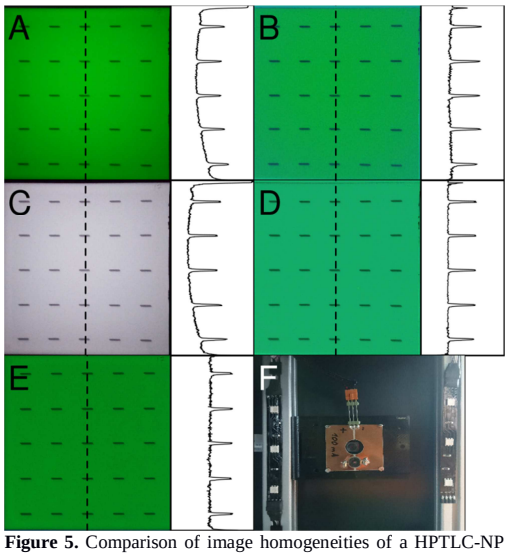


Figure 5. Comparison of image homogeneities of a HPTLC-NP plate containing 25 applied preservative bands (*%RSD*, as precision within each image*;* videodensitogram of dashed line) for different exposure times and ISO speeds, without and with algorithmic correction: 800 ms and ISO 200, uncorrected: 6.4 *%RSD* (A) and corrected: 6 *%RSD* (B), 200 ms and ISO 800 uncorrected: 4.6 *%RSD* (C) and corrected: 3.4 *%RSD* (D); as reference, TLC Visualizer with 50 ms: 3.7 *%RSD* (E) and detection part with LEDs and mini camera (F).

The image inhomogeneity was evident (Figure 5A and C). With algorithmic correction (Figure 5B and D), the image inhomogeneity was comparable to that obtained by the commercial instrument (Figure 5E). The algorithm, written in R, subtracted the pixel values of a blank plate and renormalized those values afterwards. It must be stressed that the commercial instrument also performed such a correction.[43] Using the OC system, the best precision (*%RSD)* was achieved with an exposure time of 800 ms and an ISO speed of 200. Interestingly, the green color of the fluorescence indicator plate background disappeared for longer exposure times. The small detection part (Figure 5F) substantially supported the desired miniaturization. This outcome verified the use of LEDs and such a mini camera for detection and documentation.

Water-soluble food dyes were separated on an UTLC-NP plate[44] to illustrate the advantages of RGB LEDs. Under white light illumination, the chromatograms and videodensitograms were comparable to the ones obtained from commercial instrumentation in automatic reflectance mode (Figure 6A versus B). Looking at one of the channels only, such selective evaluation is one advantage of the extracted videodensitograms of RGB images. However, this selectivity is not optimal due to the illumination over the full visible spectrum and the way the information is encoded in the RGB color space. For example, the dye E131 or E122 was still present, even though the blue or red channel was evaluated (Figure 6D and F). Those signals were not present, when the plate was illuminated with a single LED light (Figure 6C and E). The distinct narrow wavelength range of LEDs was clearly an advantage and can open new possibilities in terms of selectivity.

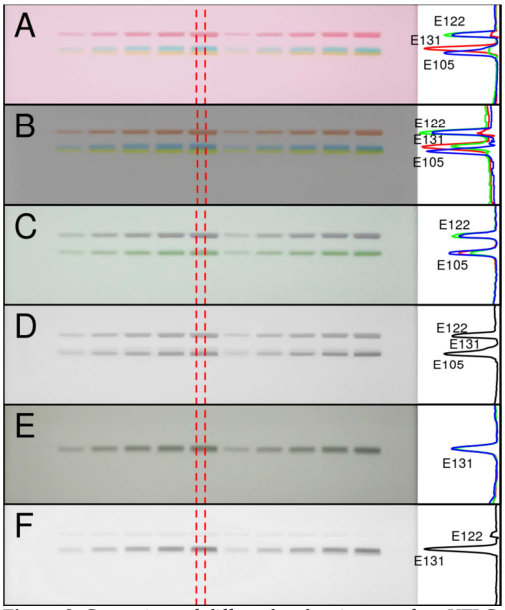


Figure 6. Comparison of differently taken images of an UTLC-NP chromatogram with the food dyes E105, E131 and E122 and respective videodensitograms (dashed lines): white illumination via LEDs in the OC system (A) TLC Visualizer as reference (B), blue LED illumination (C), blue channel of the white LED illumi-

nation (D), red LED illumination (E) and red channel of the white LED illumination (F).

The documentation was straightforwardly added to the OC system using the given electronics. This showed several advantages from the analytical perspective. The documentation between each step can be done automatically, reducing operator handling and leading to fast analyses. Reproducibility can be improved, especially for time-sensitive analyses. New options are possible, like videorecording during the development or derivatization.

Outlook. Several aspects of this device still need further improvement to finally reach all objectives of the OC concept. The solvent compatibility of the inkjet cartridge needs to be expanded. Currently, only aqueous and hydroalcoholic solutions can be printed with the given OC system, which makes it presently suited for separations on middle polar phases or RP phases only. In addition, nozzle failure of the second hand cartridge can occur, which is currently solved by replacing it. Further options for sample dosage need to be investigated, which is focus of a separate study. Instead of the current thermal inkjet technology, the piezoelectric inkjet technology could be an alternative[45] among other options. With such an inkjet system, new plate options will be possible, like the production of monolayer silica gel plates. The recently reported layer printing[13] is already implemented as a method step in the OC_manager software. Starting from the reported system, minimal hardware modifications were needed to make it compatible with the given OC system (Figure S-4). Still, the slurry doser is too big to be integrated, but mounting only the needle on the x-carriage could be the solution. An autosampler is also necessary for a final one-click system, similar to the ones in column chromatography. Further approaches need to be investigated for the development. The use of a moving pipette was already proposed.[46] The derivatization step is also supported by the current OC_manager software. It was not addressed in this study, as its proof-of-principle was already given.[15] Similarly, the possibility to heat the plate is supported. The current limitation is coming from the PLA, which will start to deform around 100 °C. This can be solved by a plate holder made of metal, polyether ether ketone or polytetrafluoroethylene, if needed, manufactured by another technology than 3D printing. Another option for local zone heating could be the integration of a laser, as demonstrated in the selective laser sintering 3D printing technique[17]. The print of bioassays for effect-directed analysis was already shown[2] and may be an alternative to existing approaches of dipping and spraying. For both, bioassays and derivatizations, inkjet printing can be an advantageous option due to the small volume consumption and the possibility to apply it freely selectable on distinct areas of the plate. Finally, the control and integration of further LED wavelengths, even as an LED array covering the full spectral range, will be advantageous to improve the selectivity range. Concerning data analysis, uncorrected images are available in the simple OC_manager software environment. Enhanced data processing as well as an easy link to other data analysis software like quanTLC[47] and rTLC[48] will be valuable.

Conclusions

This OC system demonstrated for the first time that an all-in-one planar chromatography system is possible. Its implementation was facilitated by the use open-source technologies, and the maker community was a great source of inspiration during its development. The proof-of-principle was given for the inkjet print of solutions and mobile phases on different adsorbents as well as for LEDs together with a mini camera for detection and documentation. The performance of the OC system was shown to be comparable to commercially available, mature instrumentation. The small footprint of the OC system should facilitate its integration in the lean and green laboratory. Further improvement of the current system is expected due to its high modularity and the open-source nature. Keeping the open-source aspect of the RepRap legacy, the OC system can give access to modern analytical technique on a broad basis, especially to the do it yourself maker community.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website.
Table S-1: Bill of material for the construction of the Office Chromatography system; Figure S-1: Screenshot with tabs of the OC_manager software; Figure S-2: Calculation of the drop volume; Figure S-3: 3D printed prototype for studying of the illumination at UV 254 nm; Figure S-4: Dedicated open-source system for layer printing; Instruction S-1: Instruction for the assembly of the Office Chromatography system; Instruction S-2: Instruction for the use of the OC_manager software. Separately, the Video S-1: Video of the assembly of the Office Chromatographic system, is available at https://github.com/OfficeChromatography/OCLab and also at https://youtu.be/E2v-KFa0v40.

AUTHOR INFORMATION

Corresponding Author

*Corresponding author. Tel.: +49-641-99-39141; fax: +49-641-99-39149; E-mail address: Gertrud.Morlock@uni-giessen.de (G. E. Morlock).

Notes

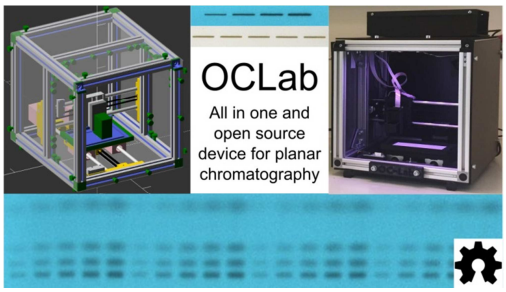
The authors declare no competing financial interest.

REFERENCES

(1) Morlock, G. E.; Oellig, C.; Bezuldenhout, L. W.; Brett, M. J.; Schwack, W. *Anal. Chem.* **2010**, *82*, 2940–2946.
(2) Morlock, G. E. *J.Chromatogr. A* **2015**, *1382*, 87–96.
(3) Izmailov, N.; Shraiber, M. *Farmatsiya* **1938**, *1*, 1–7.
(4) Kaiser, R.; Zlatkis, A. *High Performance Thin-layer Chromatography*; Elsevier Inc.: Amsterdam, 1977.
(5) Sherma, J.; Morlock, G. *J. Planar Chromatogr. - Mod. TLC* **2008**, *21*, 471–477.
(6) Nyiredy, S. Device for fully automatic TLC, Patent CH 69 008 A5, 2001.
(7) Hauck, H. E.; Schulz, M. *J. Chromatogr. Sci.* **2002**, *40*, 1–3.
(8) Hauck, H. E.; Schulz, M. *Chromatographia* **2003**, *57*, S/313–S/315.
(9) Frolova, A. M.; Chukhlieb, M. A.; Drobot, A. V.; Kryshtal, A. P.; Loginova, L. P.; Boichenko, A. P. *Open Surf. Sci. J.* **2009**, *1*, 40–45.
(10) Clark, J. E.; Olesik, S. V. *Anal. Chem.* **2009**, *81*, 4121–4129.
(11) Bezuidenhout, L. W.; Brett, M. J. *J. Chromatogr. A* **2008**, *1183*, 179–185.
(12) Song, J.; Jensen, D. S.; Hutchison, D. N.; Turner, B.; Wood, T.; Dadson, A.; Vail, M. A.; Linford, M. R.; Vanfleet, R. R.; Davis, R. C. *Adv. Funct. Mater.* **2011**, *21*, 1132–1139.
(13) Fichou, D.; Morlock, G. E. *Anal. Chem.* **2017**, *89*, 2116–2122.

(14) Li, J.; Rossignol, F.; Macdonald, J. *Lab Chip* **2015**, *15*, 2538–2558.
(15) Morlock, G.; Stiefel, C.; Schwack, W. *J. Liq. Chromatogr. Rel. Technol.* **2007**, *30*, 2171–2184.
(16) Häbe, T. T.; Morlock, G. E. *J. Chromatogr. A* **2015**, *1413*, 127–134.
(17) Gross, B. C.; Erkal, J. L.; Lockwood, S. Y.; Chen, C.; Spence, D. M. *Anal. Chem.* **2014**, *86*, 3240–3253.
(18) Gross, B.; Lockwood, S. Y.; Spence, D. M. *Anal. Chem.* **2017**, *89*, 57–70.
(19) Manzanares Palenzuela, C. L.; Pumera, M. *TrAC, Trends Anal. Chem.* **2018**, *103*, 110-118.
(20) Jones, R.; Haufe, P.; Sells, E.; Iravani, P.; Olliver, V.; Palmer, C.; Bowyer, A. *Robotica* **2011**, *29* (January 2011), 177–191.
(21) Siano, G. G.; Montemurro, M.; Alcaráz, M. R.; Goicoechea, H. C. *Anal. Chem.* **2017**, *89*, 10667-10672.
(22) Chan, K.; Coen, M.; Hardick, J.; Gaydos, C. A.; Wong, K. Y.; Smith, C.; Wilson, S. A.; Vayugundla, S. P.; Wong, S. *PLoS ONE* **2016**, *11*, 1–19.
(23) Fichou, D.; Morlock, G. E. Open-source developments for Office Chromatography, Poster P-3, International Symposium for HPTLC, Berlin, 2017.
(24) Macdonald, N. P.; Currivan, S. A.; Tedone, L.; Paull, B. *Anal. Chem.* **2017**, *89*, 2457–2463.
(25) Yu, H.; Le, H. M.; Kaale, E.; Long, K. D.; Layloff, T.; Lumetta, S. S.; Cunningham, B. T. *J. Pharma. Biomed. Anal.* **2016**, *125*, 85–93.
(26) Stroka, J.; Peschel, T.; Tittelbach, G.; Weldner, G.; Otterdjik; Anklam, E. *J. Planar Chromatogr. - Mod. TLC* **2001**, *14*, 109–112.
(27) Johnson, M. E. *J. Chem. Educ.* **2000**, *77*, 368–372.
(28) Abbaspour, A.; Mirahmadi, E.; Khajehzadeh, A. *Anal. Methods* **2010**, *2*, 349–353.
(29) Kaynar, O.; Ilerituerk, M.; Hayirli, A. *J. Planar Chromatogr. - Mod. TLC* **2013**, *26*, 202–208.
(30) Flaschka, H.; Mckeithan, C.; Barnes, R. *Anal. Lett.* **1973**, *6*, 585–594.
(31) Bui, D. A.; Hauser, P. C. *Anal. Chim. Acta* **2015**, *853*, 46–58.
(32) Macka, M.; Piasecki, T.; Dasgupta, P. K. *Annu. Rev.Anal. Chem.* **2014**, *7*, 183–207.
(33) Li, Y.; Nesterenko, P. N.; Paull, B.; Stanley, R.; Macka, M. *Anal. Chem.* **2016**, *88*, 12116−12121.
(34) Szarka, M.; Guttman, A. *Anal. Chem.* **2017**, *89*, 10673–10678.
(35) Inkshield. http://nicholasclewis.com/projects/inkshield (accessed June 20, 2018).
(36) Kelly, S. laboratory notebook. https://github.com/sjkelly/AMRI-Rice-201308-sjkelly (accessed June 20, 2018).
(37) R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2018.
(38) Chang, W.; Cheng, J.; Allaire, J.; Xie, Y.; McPherson, J. Shiny, web application framework for r, 2018, https://cran.r-project.org/web/packages/shiny.
(39) Allaire, J.; Ushey, K.; RStudio; Tang, Y.; Eddelbuettel, D.; Lewis, B.; Geelnard, M. Package 'reticulate', 2018, https://cran.r-project.org/web/packages/reticulate.
(40) Macherey Nagel, TLC Applications, Application no. 132.
(41) M. WERTHER. CAMAG Bibliogr. Service CBS **2002**, 88.
(42) Morlock, G. E.; Oellig, C. *J. AOAC Int.* **2009**, *92*, 745–756.
(43) Zhang, L.; Lin, X. *J. Chromatogr. A* **2006**, *1109*, 273–278.
(44) Kirchert, S.; Oberle, M.; Schulz, M.; Morlock, G.E. *J. Chrom. A* **2018**, in revision.
(45) De Maria, C.; Ferrari, L.; Montemurro, F.; Vozzi, F.; Guerrazzi, I.; Boland, T.; Vozzi, G. *Procedia Eng.* **2015**, *110*, 98–105.
(46) Hałka-Grysińska, A.; Skop, K.; Gorzkoska, M.; Klimek-Turek, A.; Dzido, T. New approach to development of planar chromatograms, Oral O-19, International Symposium for HPTLC, Berlin, 2017.
(47) Fichou, D.; Ristivojević, P.; Morlock, G.E. *Anal. Chem.* **2016**, *88*, 12494–12501.
(48) Fichou, D., Morlock, G.E. *J. Chrom. A* **2018**, *1560*, 78-81.

TOC



Publication 6 - Supporting information

Supporting information

Office Chromatography: Miniaturized All-in-One Open-Source System for Planar Chromatography

Dimitri Fichou[a] and Gertrud E. Morlock[a*]

[a]Institute of Nutritional Science and Interdisciplinary Research Center (IFZ), Justus Liebig University Giessen, Heinrich-Buff-Ring 26-32, 35392 Giessen, Germany

*Corresponding author. Phone: +49-641-99-39141; Fax: +49-641-99-39149;

E-mail: Gertrud.Morlock@uni-giessen.de.

Table of contents

Table S-1. Bill of material for the construction of the Office Chromatography system.

Part name	Number or comment	Distributor	Euro (VAT)
Plastic PLA	1 kg	Colorfabb	13
Zip ties	pack of 100	Local workshop	2
lm8uu bearing	8	Motedis	10
lm8uu housing	4	Motedis	12
623zz bearing	4	Amazon	2
gt2 belt	2 m	Motedis	3
gt2 pulley 20 teeths	2	Motedis	3
Nema 14	2	Emotiontech	30
8 mm stainless steel rod	4 x 255 mm	Motedis	15
20 x 20 profile 5 I-type	5 x 215 mm (X) + 4 x 255 mm (Y) + 6 x 210 mm (Z)	Motedis	10
Nut for 20 x 20 profile 5 I-type	pack of 100	Motedis	20
Cube connector for 20 x 20 profile	pack of 10 (6 needed)	Motedis	25
Angle connector for 20 x 20 profile	pack of 10 (10 needed)	Motedis	10
Aluminium frame for box	2 mm thick	Local workshop	30
Screw M5 x 10	pack of 100 hex type	Local workshop	5
Screw M5 x 16	pack of 100 hex type	Local workshop	5
Screw M3 x 30	pack of 100 phillips or slotted type	Local workshop	5
Screw M3 x 10	pack of 100 hex type	Local workshop	5
Screw M3 x 16	pack of 100 phillips or slotted type	Local workshop	5
Screw M4 x 20	pack of 100 phillips or slotted type	Local workshop	5
Nut M3	pack of 100	Local workshop	3
Nut M4	pack of 100	Local workshop	3
Nut M5	pack of 100	Local workshop	3
Washer M3	pack of 100	Local workshop	3
Washer M4	pack of 100	Local workshop	3
Washer M5	pack of 100	Local workshop	3
Magnet	8 cylinder, OD 8 mm, height 3 mm	Amazon	2
Glass plates	1 mm thick, recycled from commercial plates	Recycled	0
Raspberry Pi 3	also Raspberry Pi 2 is compatible	Conrad	35
5V 2.5 A power supply	for Raspberry Pi	Conrad	10
16 Giga bit SD card	1	Conrad	10
Ethernet cable	1	Local workshop	5
Raspberry Pi camera	1	Conrad	25
Raspberry Pi camera longer cable	200 mm is enough	Conrad	5
Arduino Mega 2560	1	Conrad	30
RAMPS 1.4	1	Amazon	30
A9688 motor driver	2	Amazon	20
Endstop	2	Amazon	10
Inkshield board	1	Nerdcreationlab	100
12 V 10 A power supply	5 A could be enough	Amazon	20
NPN transistor 2n3904h33	3	Conrad	5
LED strip	1 m	Conrad	10
Prototype soldering plate	60 x 40 mm	Amazon	5
Breadboard Jumper Wires Ribbon Cable		Amazon	5
255 nm LED	Optan 225P SMD 4-6 mW	Crystal IS	200
Power supply for UV LED	no information	Local workshop	50
		Total	**810**

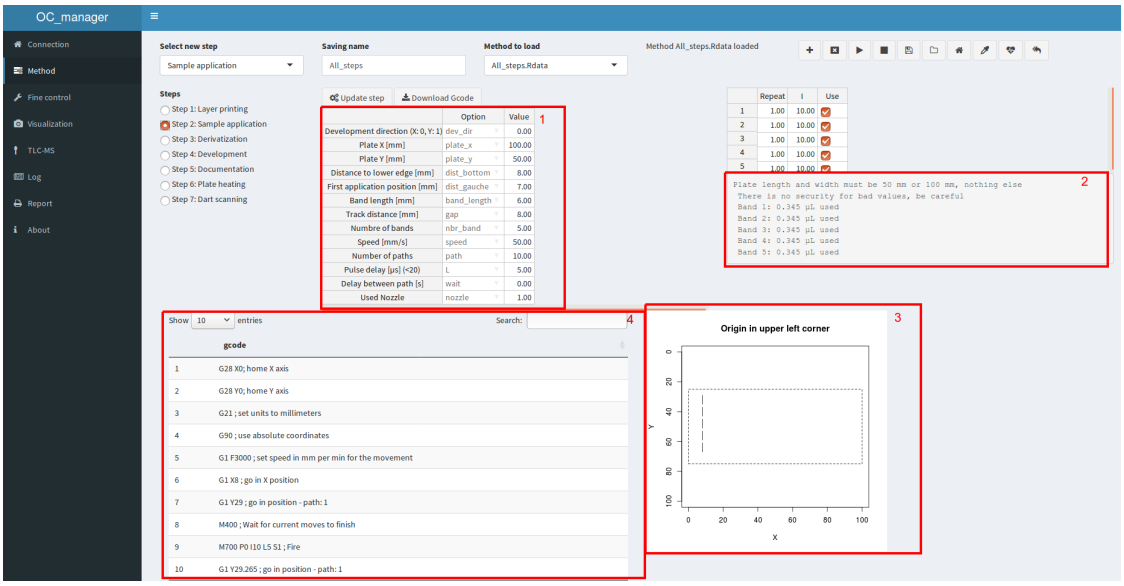


Figure S-1. Screenshot of the OC_manager software with option table (1), printed feedback (2), plotted feedback (3) and GCODE table (4).

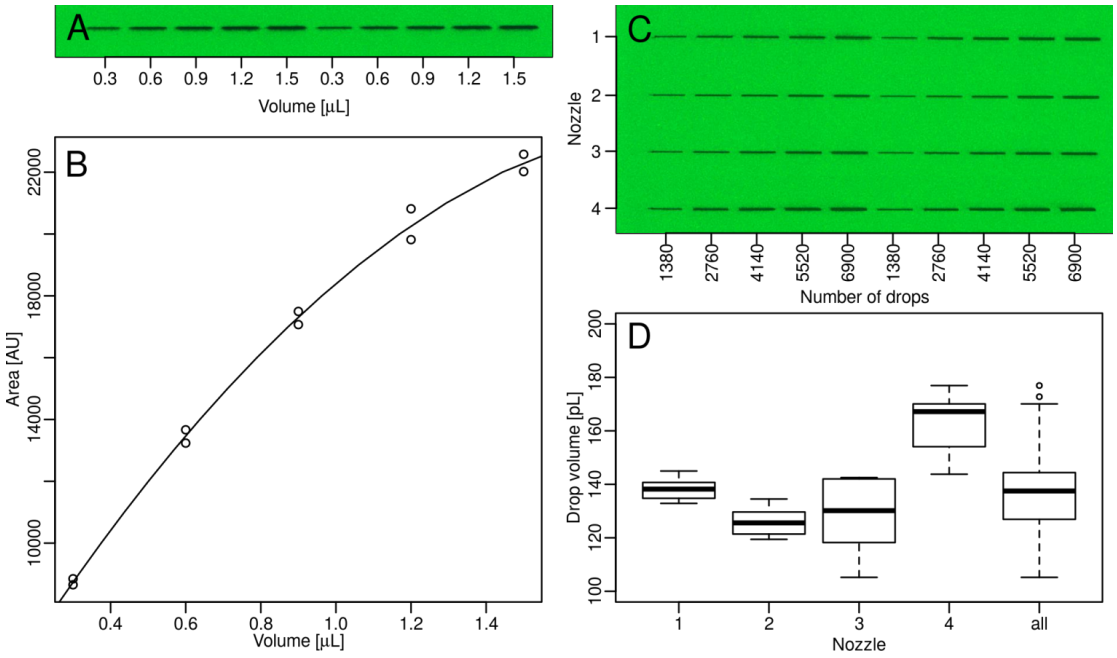


Figure S-2. Calculation of the drop volume with preservatives solution via scanning densitometry without development with spray-on application (A) used for calibration (B; R^2 = 0.9963), inkjet applied of 4 nozzles (C) and boxplot of the calculated drop volume for each nozzle and for all together (D; mean *RSD* = 6%, total *RSD* = 12%).

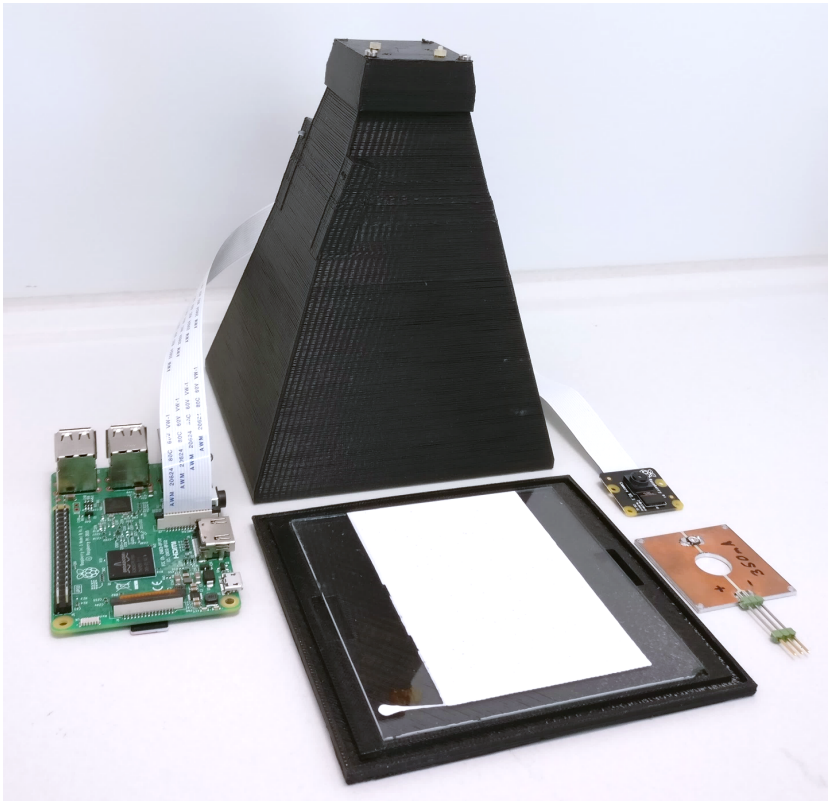


Figure S-3. 3D printed prototype for studying of the illumination at UV 254 nm, showing the baseplate carrying a self-printed layer on a 10 x 10 cm^2 glass plate and cover to be placed on baseplate.

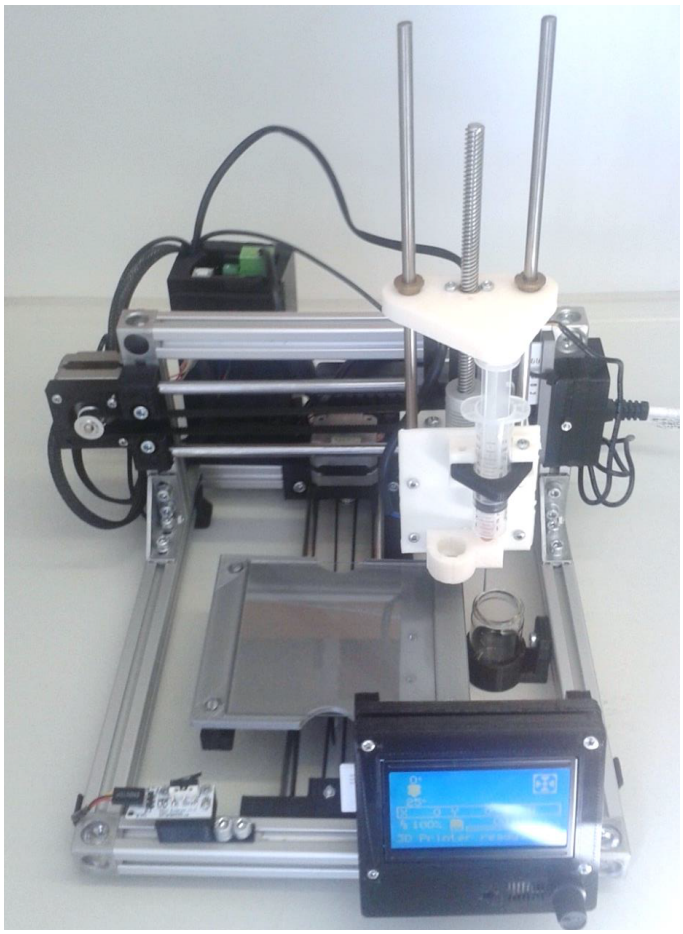


Figure S-4. Dedicated open-source system for layer printing.

Instruction S-1: Instruction for the assembly of the Office Chromatography system.

Useful resources

Some knowledge is mandatory to build this device. The following links should give an overview. Over this document, it is assumed a significant amount of time had been spend on those websites and that their content is known.

3D printing:

- Reprap project
- Gcode listing
- Course on 3D printing

Information on each electronic board used:

- Ramps 1.4 board (electronic of our machines)
- Inkshield board for inkjet, more info from people who used the board here, here, here. Place to buy the cartridge
- Raspberry Pi project
- Arduino project (see those courses for an introduction on the Internet of things)

Some software to install:

- Openscad software (for drawing, see this tutorial)
- Slic3r software (to create Gcode file for 3d printing, here is the main page but it's better to install the Prusa edition)
- Arduino IDE (to modify and upload the marlin firmware)

Others:

- Instructable for LED control from the Ramps 1.4

Print the parts

SCAD folder

SCAD file: **full_view("printed")**.

SCAD files can be found in the SCAD folder, normally, only the file OCLab.scad is of interest. Each different module can be observed separatly by changing the argument of the `full_view` module.

STL folder

The STL files in this folder contain several files, only the file `OCLab.stl` should be considered up to date. Parts are designed to be printed on a Prusa i3 MK2 with the *0.35mm fast* setting. Normally, no supports are needed but a brim could be usefull. Be sure to rotate the part so they print correctly.

Bill of material

Table 1: Bill of material - Estimated total price: 810 Euro (VAT)

Part name	Number or comment	Seller	Euro (VAT)
Plastic PLA	1 kg	Colorfabb	13
Zip ties	pack of 100	Local workshop	2
lm8uu bearing	8	Motedis	10
lm8uu housing	4	Motedis	12
623zz bearing	4	Amazon	2
gt2 belt	2 m	Motedis	3
gt2 pulley 20 teeth	2	Motedis	3
Nema 14	2	Emotiontech	30
8 mm stainless steel rod	4 x 255 mm	Motedis	15
20 x 20 profile 5 I-type	5 x 215 mm (X) + 4 x 255 mm (Y) + 6 x 210 mm (Z)	Motedis	10
Nut for 20 x 20 profile 5 I-type	pack of 100	Motedis	20
Cube connector for 20 x 20 profile	pack of 10 (6 needed)	Motedis	25
Angle connector for 20 x 20 profile	pack of 10 (10 needed)	Motedis	10
Aluminium frame for box	2 mm thick	Local workshop	30
Screw M5 x 10	pack of 100 hex type	Local workshop	5
Screw M5 x 16	pack of 100 hex type	Local workshop	5
Screw M3 x 30	pack of 100 phillips or slotted type	Local workshop	5
Screw M3 x 10	pack of 100 hex type	Local workshop	5
Screw M3 x 16	pack of 100 phillips or slotted type	Local workshop	5
Screw M4 x 20	pack of 100 phillips or slotted type	Local workshop	5
Nut M3	pack of 100	Local workshop	3
Nut M4	pack of 100	Local workshop	3
Nut M5	pack of 100	Local workshop	3
Washer M3	pack of 100	Local workshop	3
Washer M4	pack of 100	Local workshop	3
Washer M5	pack of 100	Local workshop	3
Magnet	8 cylinders, OD 8 mm, height 3 mm	Amazon	2
Glass plates	1 mm thick, recycled from commercial plates	Recycled	0
Raspberry Pi 3	also Raspberry Pi 2 compatible	Conrad	35
5 V 2.5 A power supply	for Raspberry Pi	Conrad	10
16 Giga bit SD card	1	Conrad	10
Ethernet cable	1	Local workshop	5
Raspberry Pi camera	1	Conrad	25
Raspberry Pi camera longer cable	200 mm is enough	Conrad	5
Arduino Mega 2560	1	Conrad	30
RAMPS 1.4	1	Amazon	30
A9688 motor driver	2	Amazon	20
Endstop	2	Amazon	10
Inkshield board	1	Nerdcreationlab	100
12 V 10 A power supply	5 A could be enough	Amazon	20
NPN transistor 2n3904h33	3	Conrad	5
LED strip	1 m	Conrad	10
Prototype soldering plate	60 x 40 mm	Amazon	5
Breadboard Jumper Wires Ribbon Cable	NA	Amazon	5
255 nm LED	Optan 225P SMD 4-6 mW	Crystal IS	200
Power supply for UV LED	no information	Local workshop	50

Frame

SCAD file: **full_view(“frame”)**.

Needed parts:

- 76x M5 sliding nuts
- 20x20 profiles type I
- 6x 20x20 cube connectors
- 10x 20x20 angle connectors
- M5 screw for cube connector (sold with it normally)
- M5*10 screw for angle connectors
- printed parts: `feets`; `green_top_front`
- closing plates

Before assembling the frame, set the nuts to have one nut for each M5 screw shown in Figure 1. There is 76 in total. The best is to use the openscad file and turn the model around.

The verticale profile in the middle as well as the horizontale one in the top-middle are situated 120 mm away from the ones in the back.

The outside frames can be found in the `outside frames` folder in pdf format. The plates to close the box can be added now or later but can make the assembly difficult, worth case scenario, they are put to early and need to be dismounted to grant access.

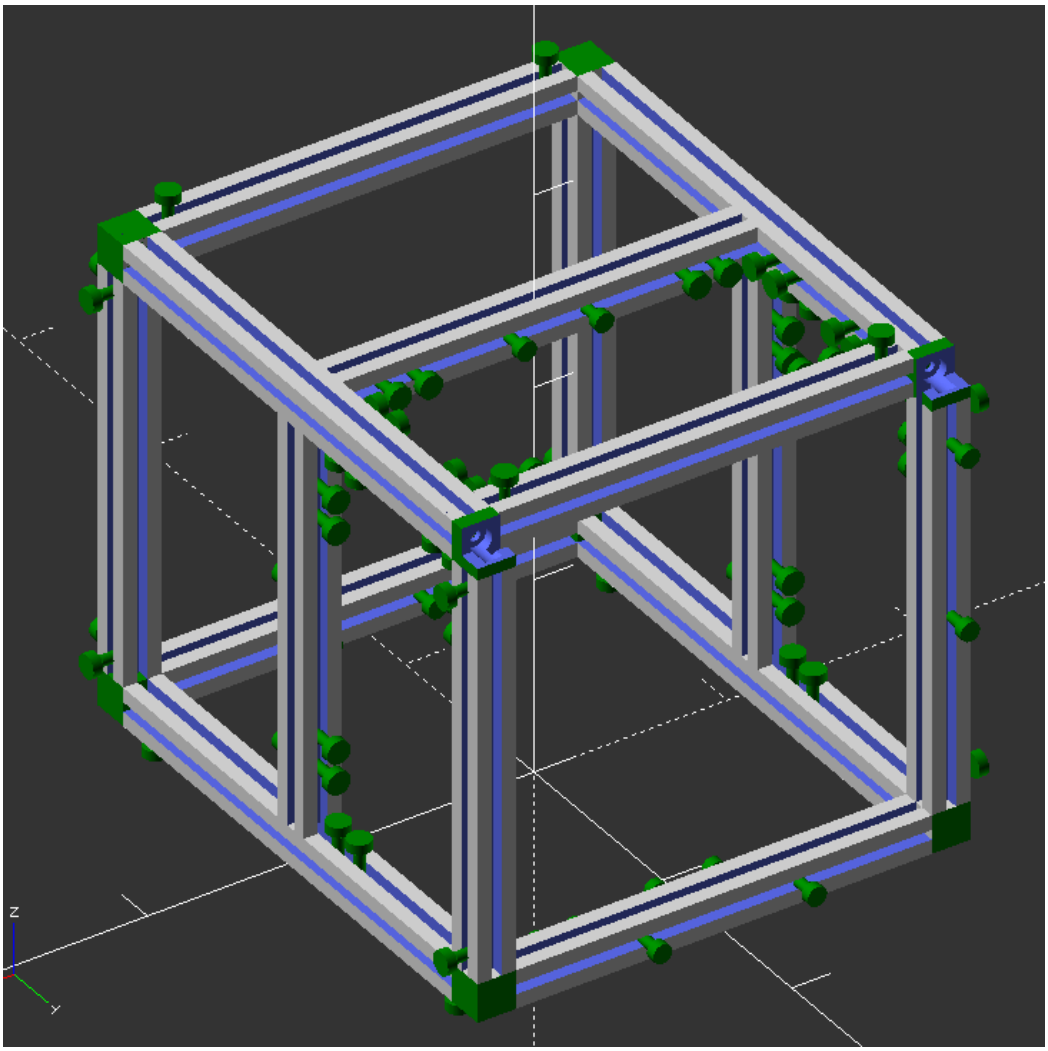


Figure 1: Frame assembly.

Y axis

SCAD file: **full_view("Y axis")**.

Needed parts:

- 2x 8mm rods
- 4x lm8uu bearing
- 4x lm8uu housing
- 2x 623zz bearing
- 1x endstop
- 1x nema14
- 1x gt2 pulley
- printed parts: `Y_endstop_holder`, `Y_moving`, `plate_holder`, `Y_belt_holder`, `Y_end` and `Y_motor`
- 8x magnets: cylinder OD = 8mm, height = 3mm
- 16x M4x10 (lm8uu housing)
- 3x M3x30 (belt holder + 623zz bearing)
- 4x M3x12 (motor)
- 2x M3x10 (endstop)
- 9x M3 washers
- 5x M3 nuts

Assemble the Y axis as in Figure 2. Start by the `Y_moving` part, screw the lm8uu housing (with the lm8uu in). For the belt holder, be carrefull with the direction, it can be changed when the belt will be tightned during the assembly to the frame if needed.

Pull the magnets in the `Y_moving` part and the `plate_holder` part, be careful with the direction.

Insert the rods in the lm8uu housing. The rods must then fit in the `Y_end` and `Y_motor` parts, if not, drill a 8mm hole.

The endstop must be screw to the `Y_endstop_holder` and can be put appart for the moment and will be set during the assembly.

For the plate holder, the one having its own STL file is the multipurpose one but multiple versions are available in the main STL file to fit different needs.

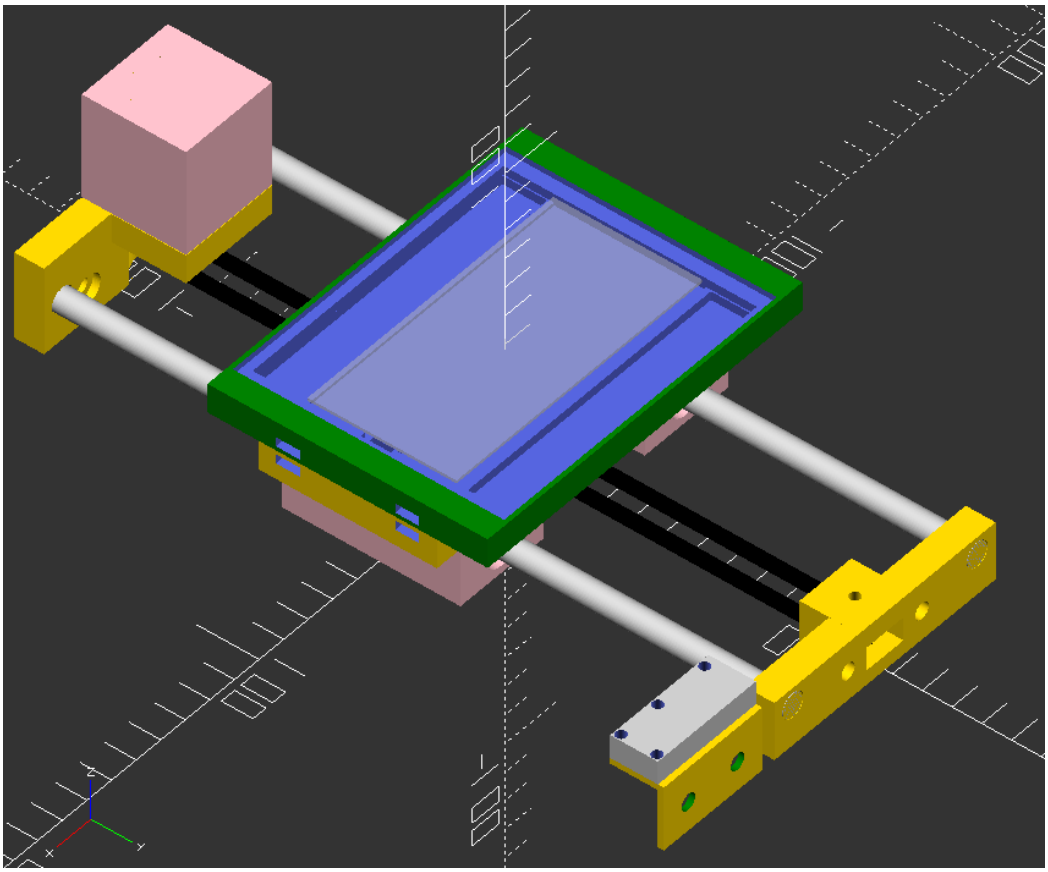


Figure 2: Y axis.

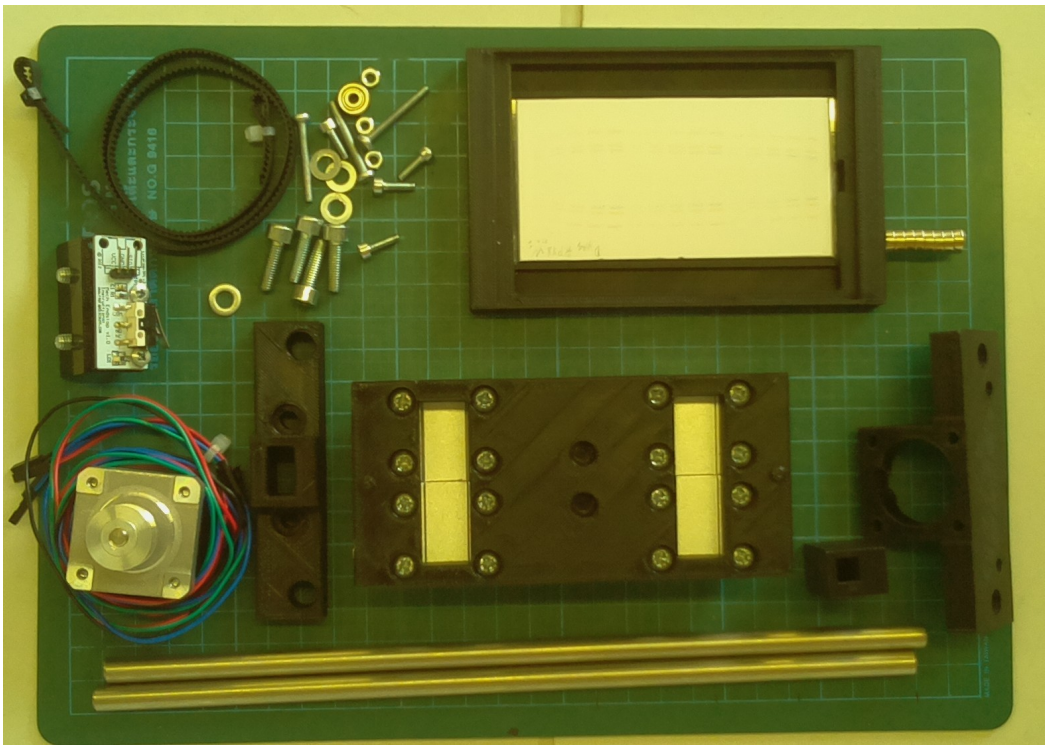


Figure 3: Y axis before.

X axis

SCAD file: **full_view("X axis")**.

Needed parts:

- 2x 8mm rods
- 4x lm8uu bearing
- 2x 623zz bearing
- 1x endstop
- 1x nema14
- 1x gt2 pulley
- printed parts: `X_motor`, `X_end`, `X_moving`, `HP_C6602_holder_holder`, `X_end_stop_holder`
- HP C6602 holder
- 2x M4x20 (HP_C6602_holder_holder)
- 2x M4 nut (HP_C6602_holder_holder)
- 1x M3x30 (623zz bearing)
- 4x M3x10 (motor)
- 2x M2*10 (HP_C6602_holder_holder)
- 1x M3 nut (623zz bearing)
- 2x M2 nut (HP_C6602_holder_holder)
- 4x M3 washer (motor)

Assemble the `X_motor`, `X_end` and `X_moving` parts as for the Y axis. Note that the `X_moving` and

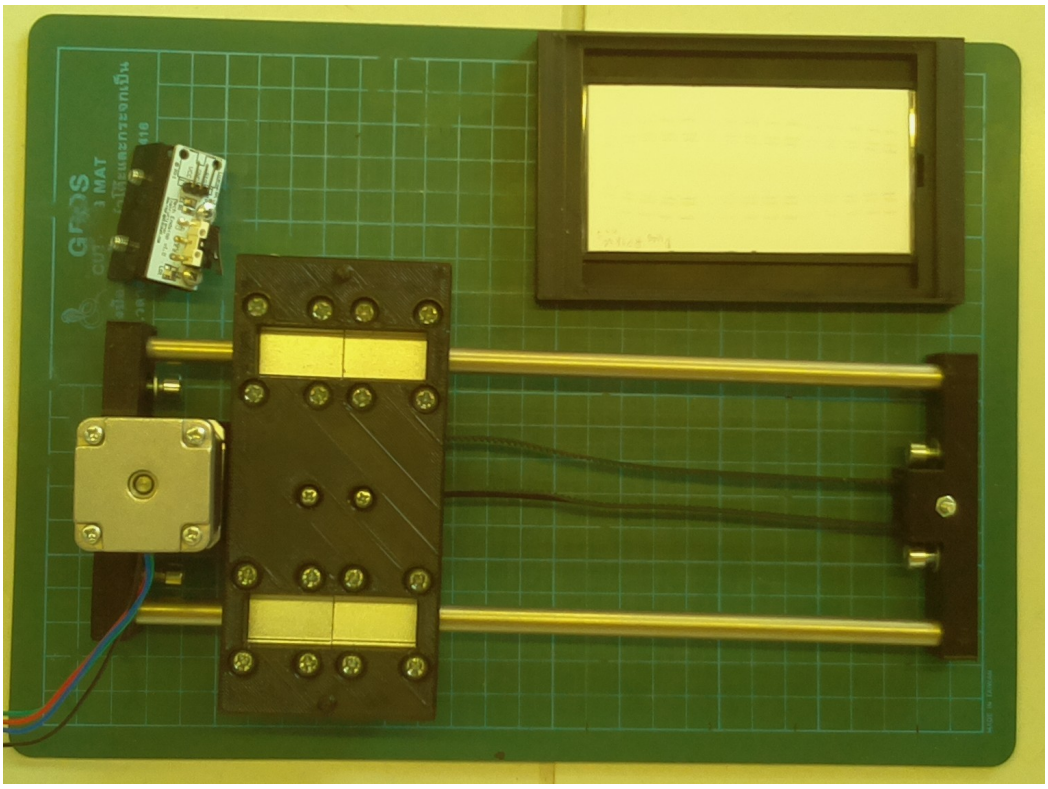


Figure 4: Y axis after.

`X_end_stop_holder` are taken from the other designs, the STL files are in the STL folder separatly but also in the main STL file, taken from prusa i3 rework.

The link between the `HP C6602 holder` and `X_moving` is made by the `HP_C6602_holder_holder` part, use M4 and M2 screws and nuts for this. M4 link to the `X_moving` and M2 to the `HP C6602 holder`.

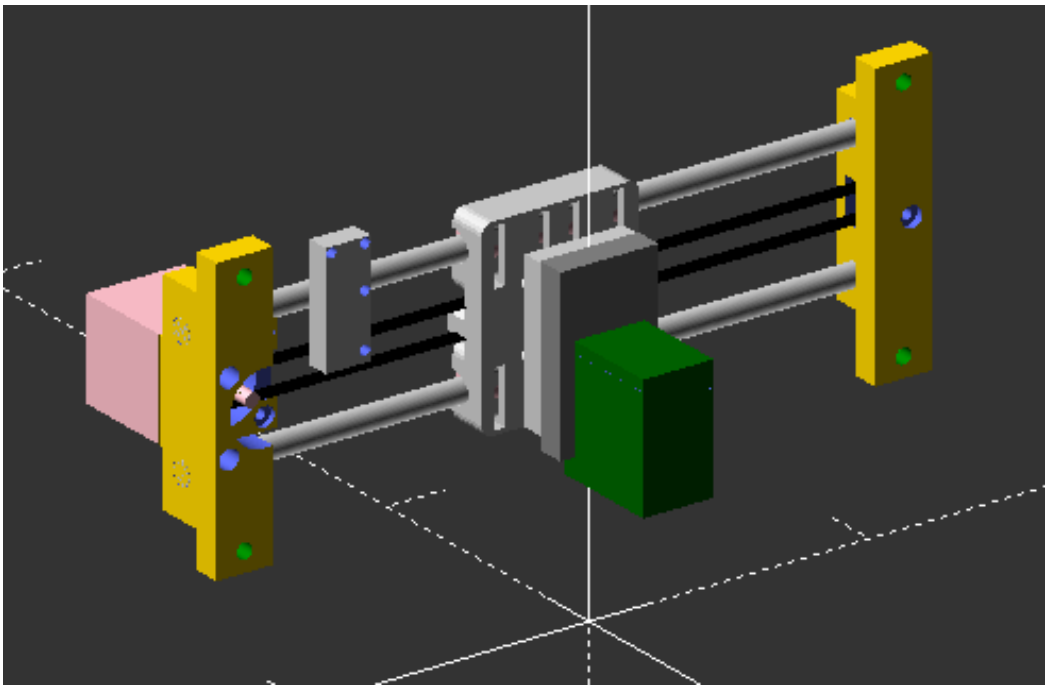


Figure 5: X axis SCAD.

Axis assembly

SCAD file: **full_view("assembled")**.

Needed parts:

- X axis
- Y axis
- frame
- 10x M5
- belts

Position the X axis and Y axis as in Figure 8. The belt must be tightened and blocked with zip-ties.

The X axis must be position as low as possible while avoiding collision with the plate holder. The Y axis will be positioned later when the cartridge will be able to fire.

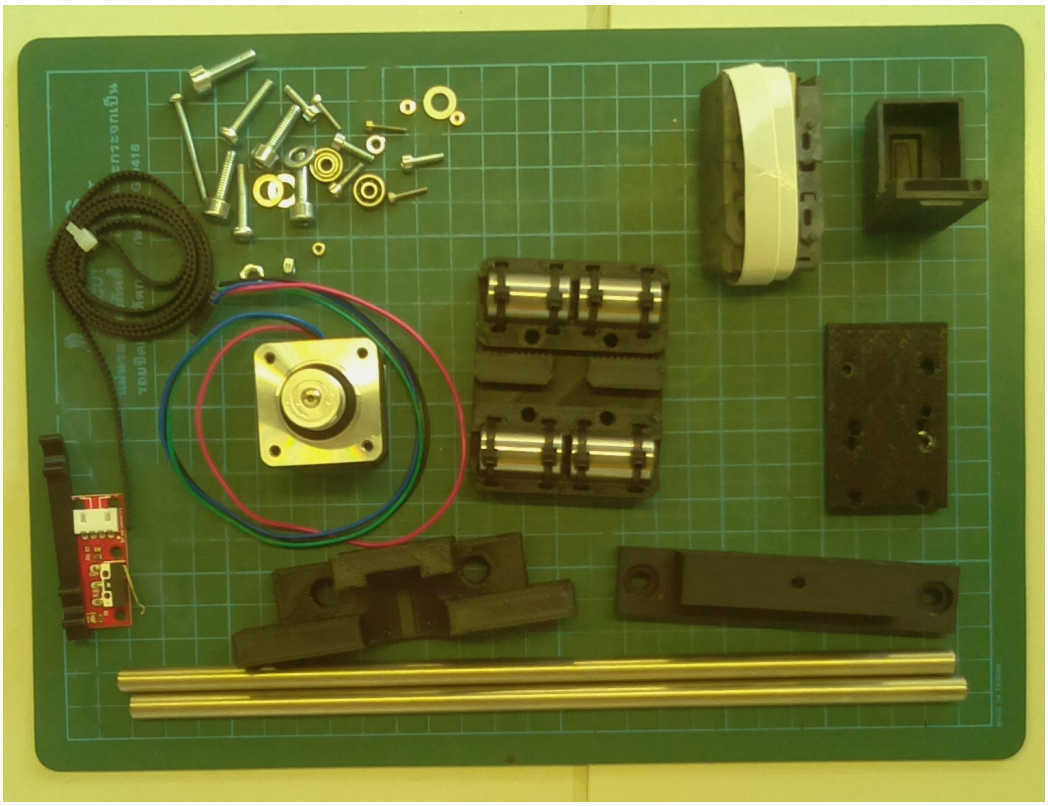


Figure 6: X axis before.

Electronics

SCAD file: **full_view("elec")**.

Needed parts:

- printed parts: `elec_holder`, `camera_holder`
- Raspberry Pi
- Arduino mega
- Ramsp 1.4
- 2x A9688 motor drivers
- LED strip
- Prototype soldering plate
- 3x NPN transistor 2n3904h33
- Breadboard Jumper Wires Ribbon Cable
- 6x M3x20 screws
- 8x M2x10 screws
- 8x M2 nuts
- 255 nm LEDs

Solder the inkshield board as shown on the website.

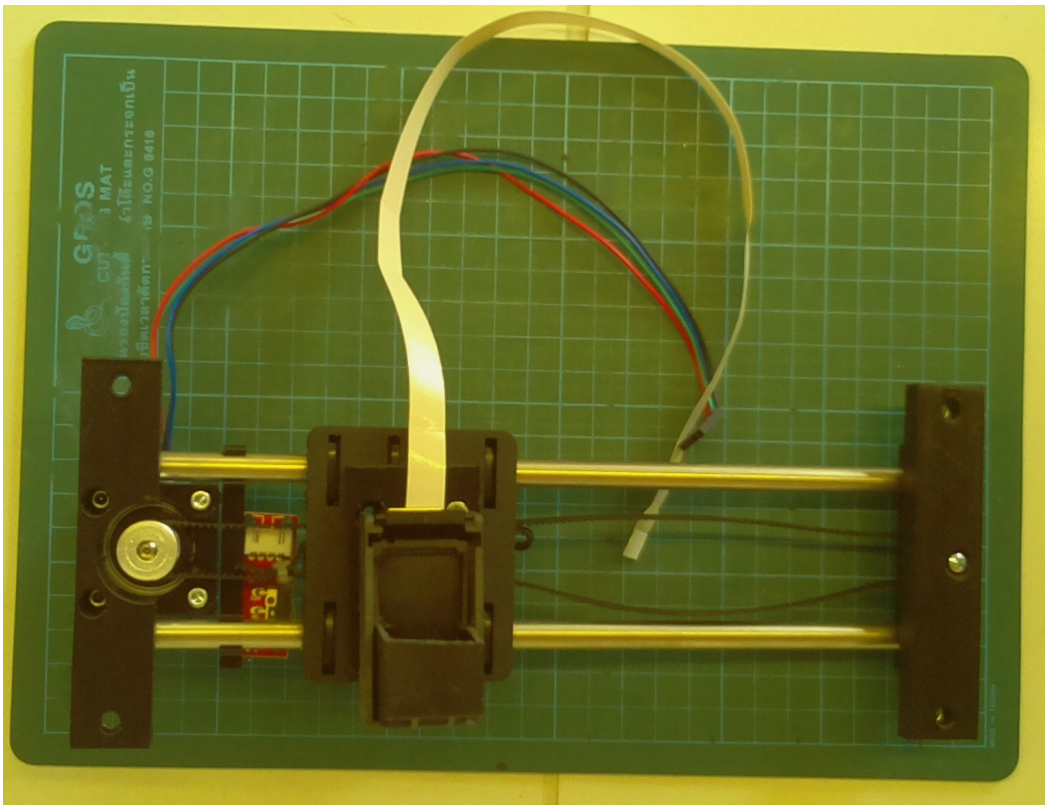


Figure 7: X axis after.

Mount the UV LED to an appropriate board (screw distance 35.5mm, hole for camera lens centered in the middle) and connect it to the `camera_holder`.

Follow this instructable to make the electronics for the RGB LED strips. The LED strips must be attached with double sided tape to the `camera_holder` part.

Screw the Arduino/Ramps and inkshield boards to the `elec_holder` part and mount it in the frame as in Figure 9. Plug the motor driver in the X and Y axis positions.

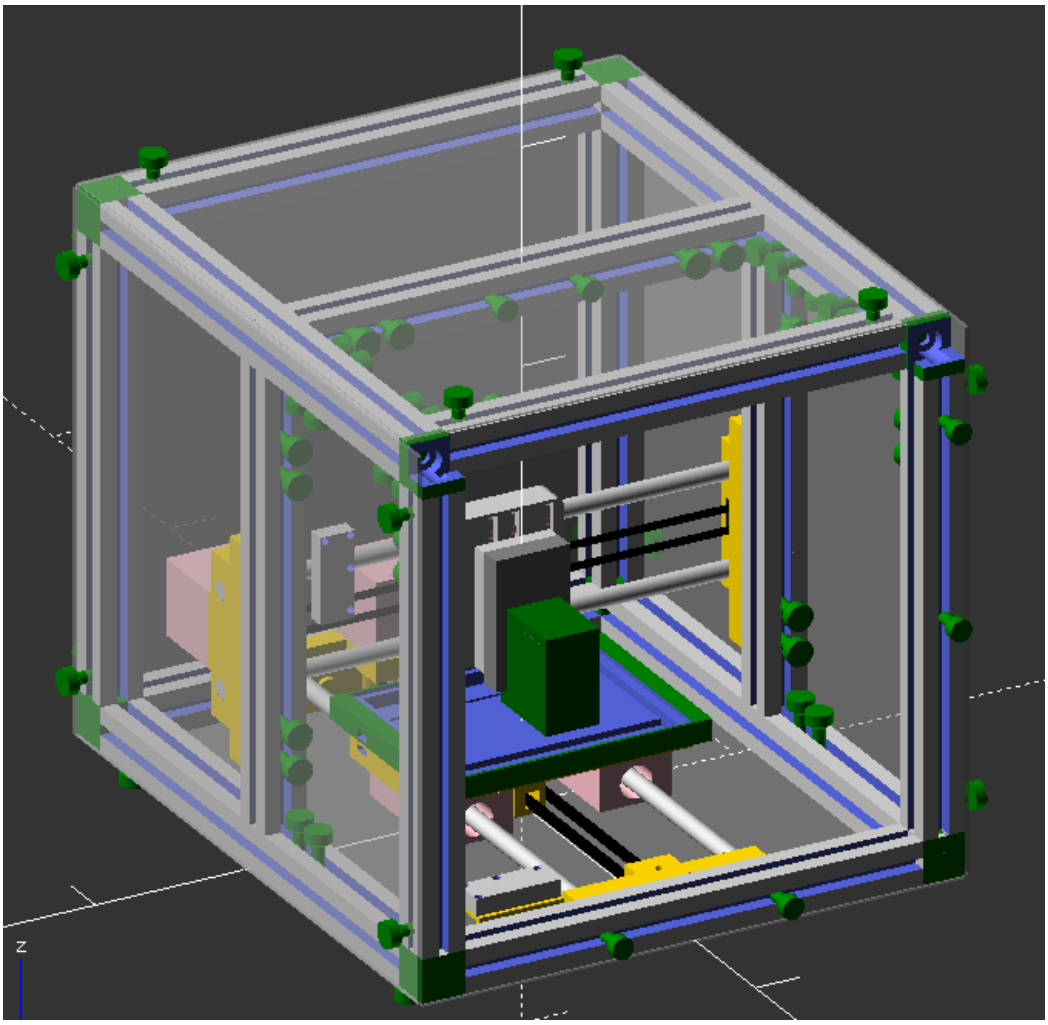


Figure 8: Assembled.

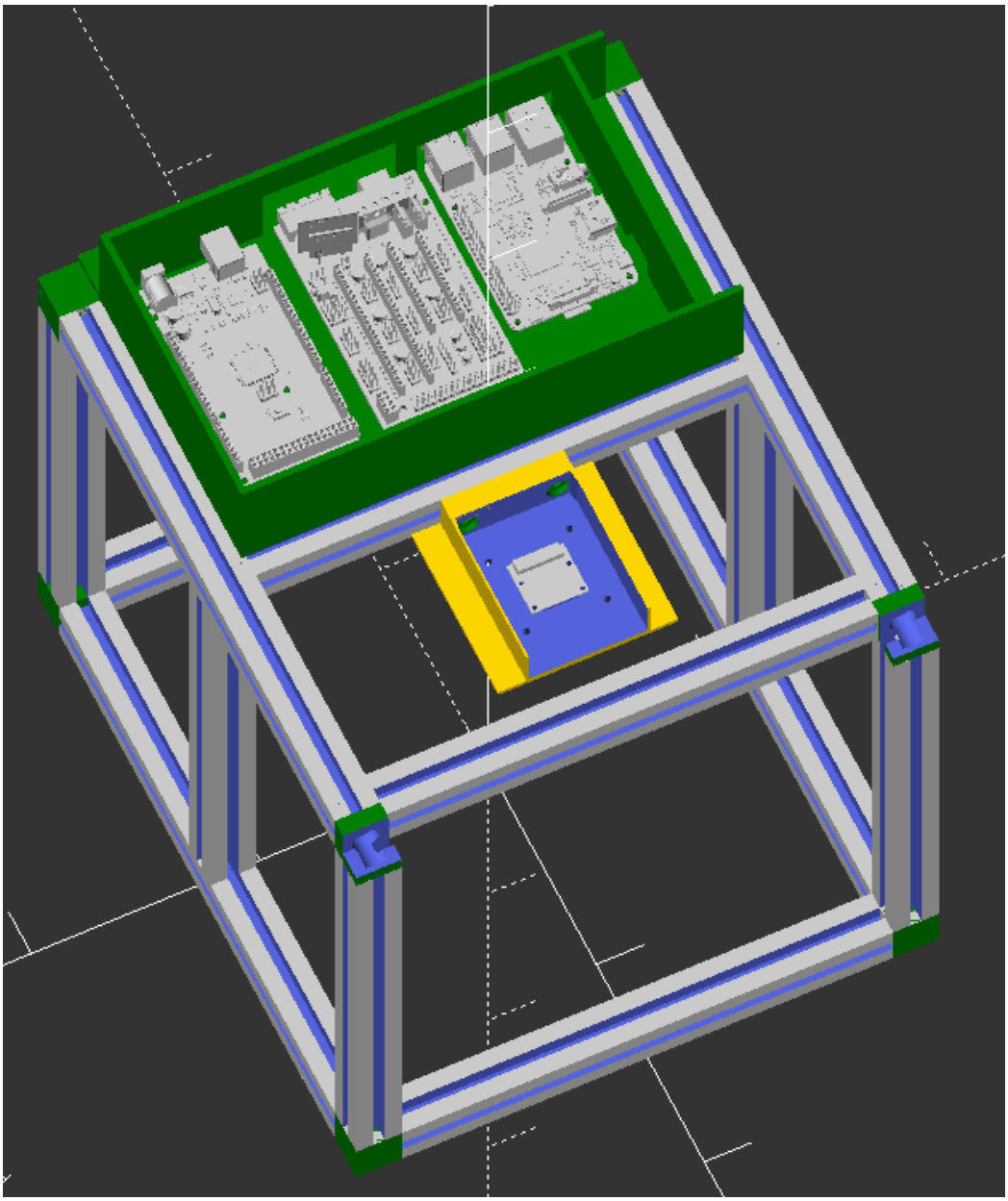


Figure 9: Electronic SCAD.

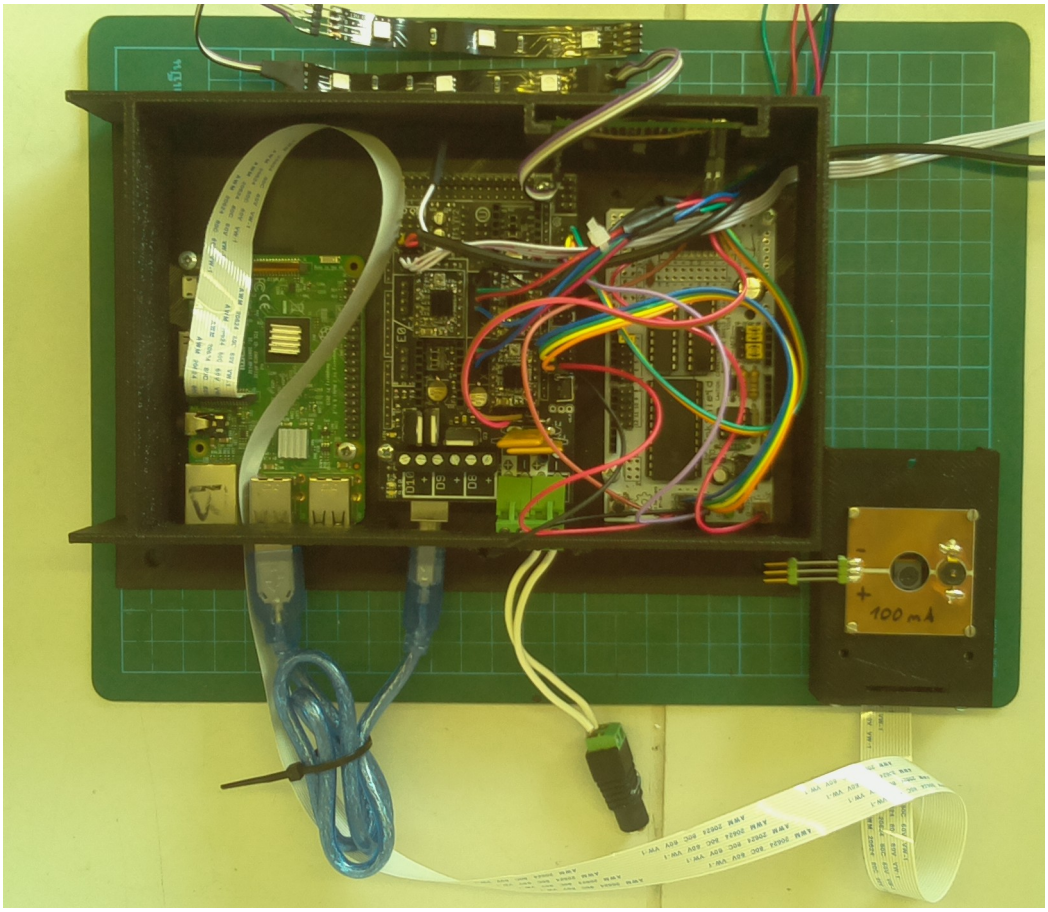


Figure 10: Electronic.

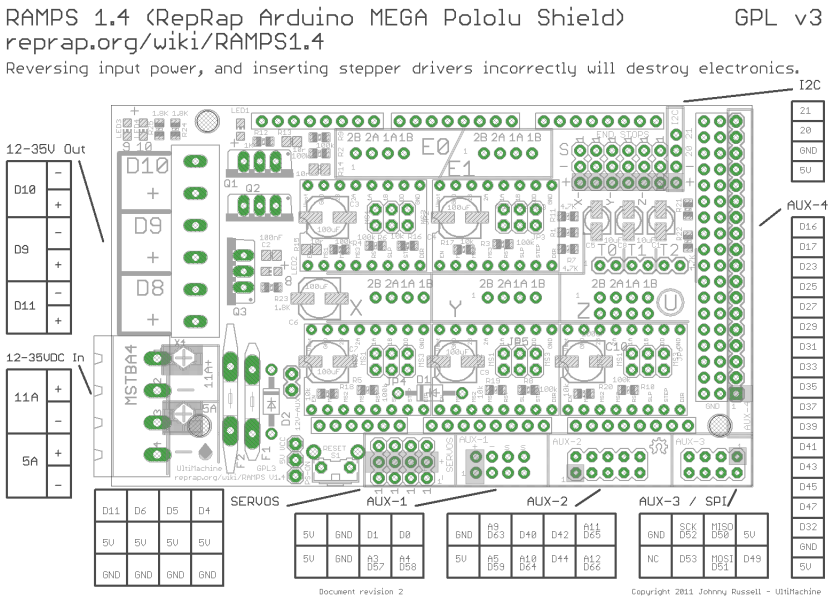


Figure 11: Ramps 1.4 schematic.

Wiring

The schematic of the Ramps 1.4 is presented in Figure 11.

Motors and endstop

Wire the Ramps 1.4 board as in the Figure 12. Only the X and Y motor are used, including their respective endstops. Note that this figure is extracted from a 3D printer wiring.

LEDs

Wire the RGB LED as in Figure 13 with the difference that “Red” = 44,“Green”=66,“Blue”=64. Again, go back to the instructable if needed and refer to Figure 11.

For the UV LEDs, a separated power supply is necessary and it cannot be controlled by the software. If a solution is found, the pin 59 is reserved for the 255 nm LED.

Inkshield board

Similarly to the RGB LEDs, connect the inkshield board with ribbon female-female cable. The auxiliary input must be used and the connections are: “A”= 11,“B”= 6,“C”= 5,“D”= 4,“pulse”=63. The 12V, 5V and ground can be taken from the Ramps and RGB LED board.

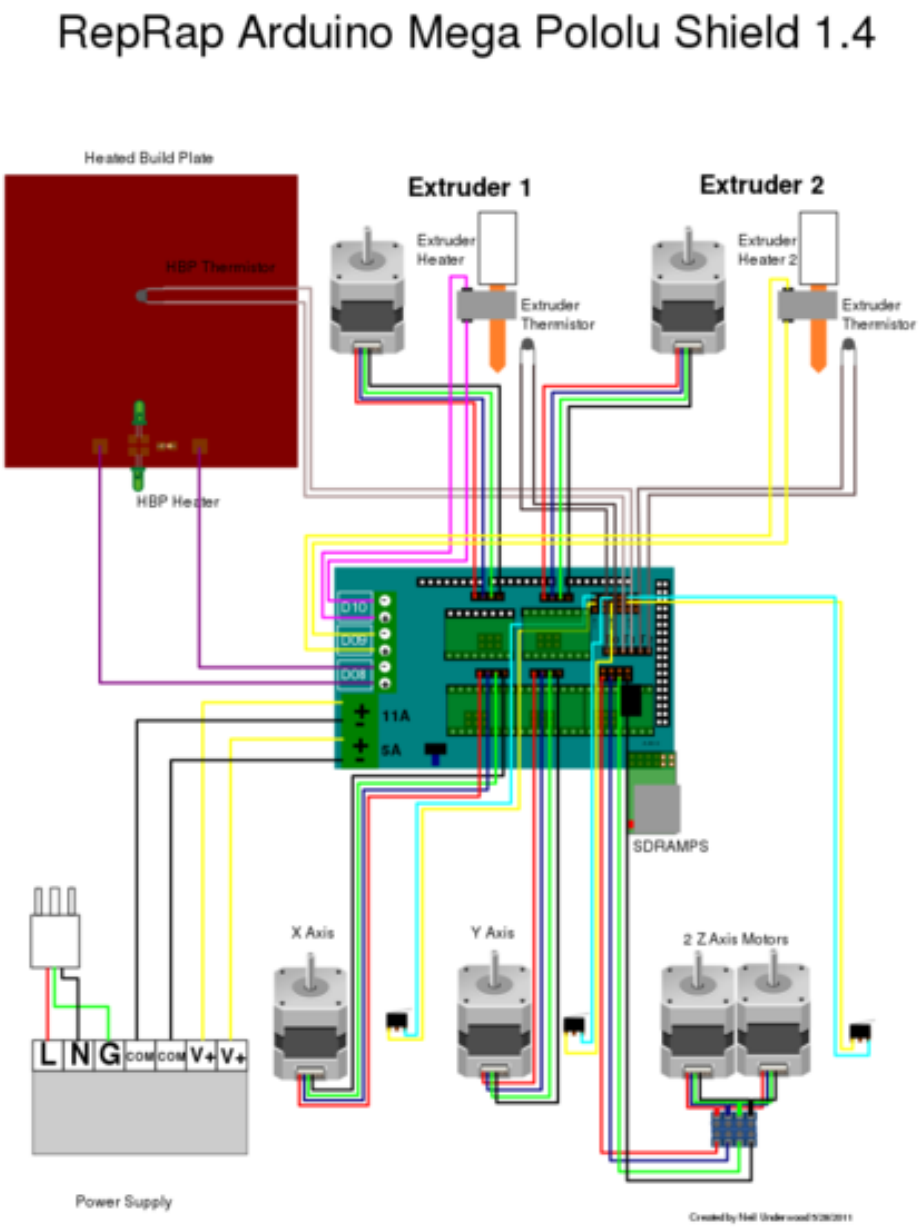


Figure 12: Motor wiring.

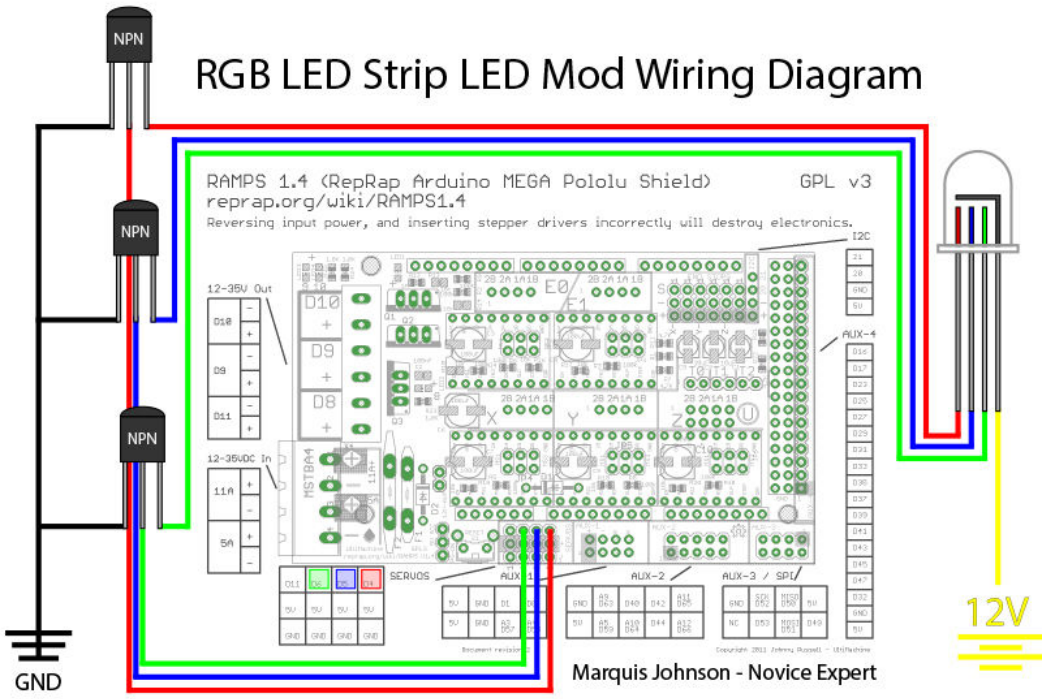


Figure 13: LED wiring.

Software set up

Uploading the firmware

Install the 1.0.x release of the Arduino IDE on your PC. The version of Marlin used is only compatible with this IDE.

Upload the Marlin firmware present in the github repository of OCLab to control the board. Open the `Marlin.ino` file and flash the firmware on the board while connected via the USB cable that you normally received with the Arduino (Figure 14).

OC manager

Follow the instruction on the github repository of OC_manager. Steps could be missing and feedbacks are more than welcome. As a rule of thumb, a rapid google search with the error message can solve most problems.

You can choose to install it from scratch or to use the prepared SD card to be use directly with a Raspberry Pi, in both cases, you will need to set-up a static IP to access it from the network.

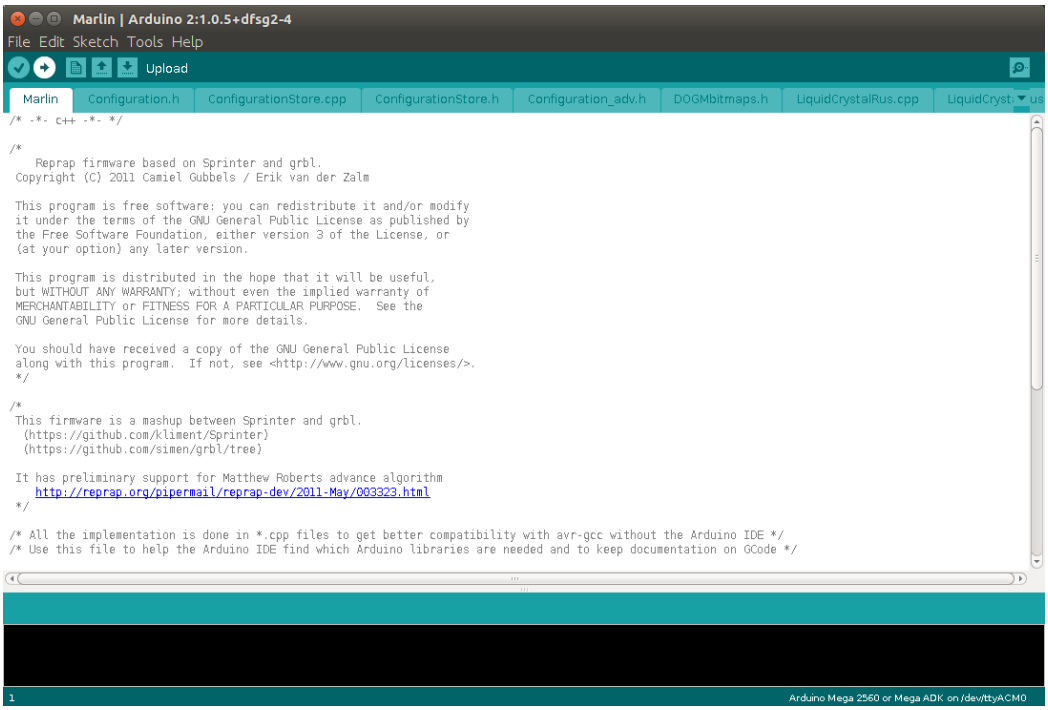


Figure 14: Marlin upload.

Calibration

The only calibration necessary is to position correctly the Y axis. To do so:

- Open OC_manager by going in your browser on the IP address of the Raspberry Pi,
- Connect the board
- In the Fine control tab (Figure 15): home the X axis and the Y axis, (if the motor are not in the good direction, turn off the full system and simply reverse the cable on the ramp)
- With an ink cartridge in the cartridge holder and a plate in the 10x10 cm plate holder, click on `fire selected nozzles`. The 12 ink droplets must be on the plate in the top left corner. Move the Y axis if necessary.

This calibration procedure is not enough and is a known problem, ideally, a complementary procedure should be inserted in the code to add a bias in X and Y to every movement. The 3D printing environment was done to have reproducible movement **during the print** and not between print.

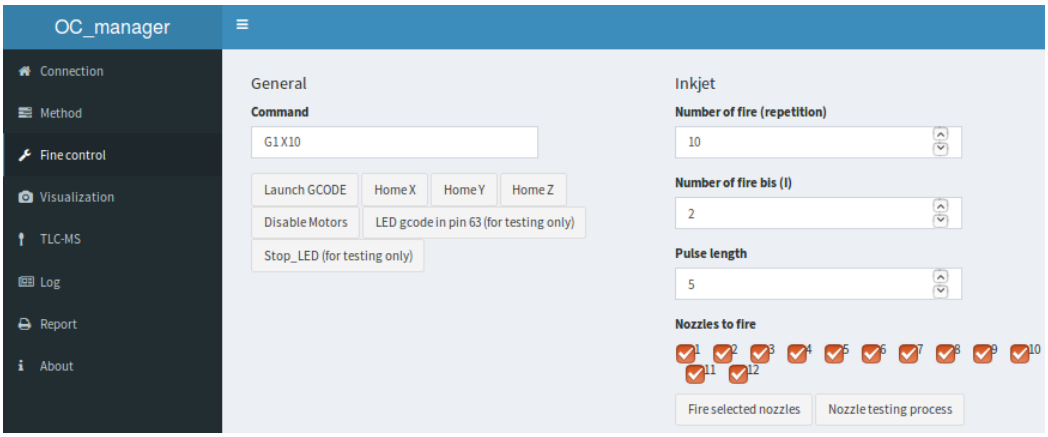


Figure 15: Fine tuning OC manager.

Instruction S-2: Instruction for the use of the OC_manager software.

For preview, a version of OC_manager without device connected can be accessed online at 134.176.7.66/OC_manager.

Prerequisite

- The device had been installed properly (github repository).

- OC manager had been installed on a linux system (github repository).
- OC manager is hosted on a raspberry pi with static IP and that OC manager is launch at reboot via crontab or equivalent method.
- OC_manager can be accessed on a modern web browser via the static IP on port 80.

Turning ON the device

Those connections can be made with any order but this one is preferable.

1. If the raspberry pi is accessed via the local network, connect the ethernet cable to the raspberry pi.
2. Connect the 5V 2A power supply to the raspberry pi.
3. Check that the USB cable is connected between the RAMPS and the raspberry pi.
4. Connect the 12V power supply to the RAMPS.
5. From a web browser, go to the static IP of the rapsberry pi to access OC manager.

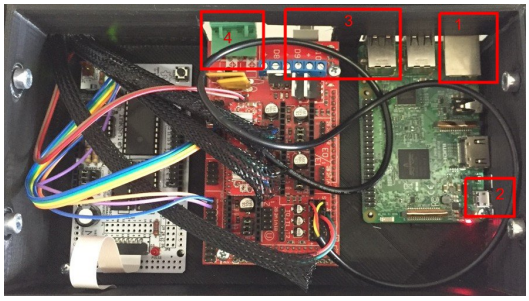


Figure 1: Electronics

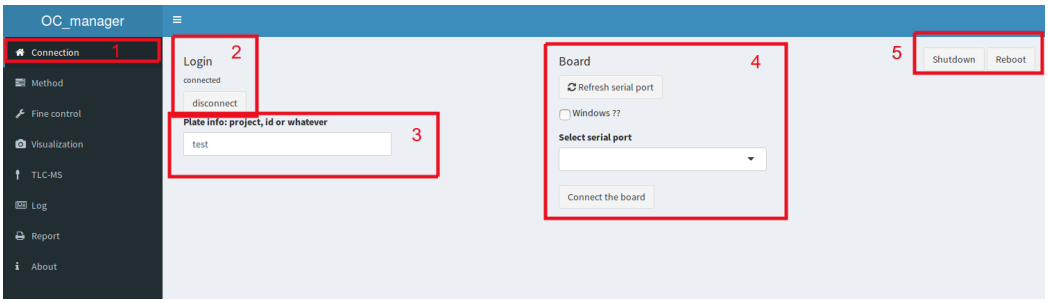


Figure 2: Connection tab

Tabs presentation

Connection

The device does not connect automatically.

1. Location of the tab
2. Login if enable in the **config.R** file, this function is very basic and not encrypted but at least it is not in a text document. When logged in with the **admin** user, you can add and delete user and modify their password, the default password for the admin user is “raspberry”.
3. Plate info, give it a specific name and this name will appear in the log as well as in the pictures folder files.
4. Connect to the board, choose the good port, refresh if needed and connect the board. When the button change to **disconnect the board**, you are connected, it is always good to test it in **fine control**. If there is connection problem, reboots fix it most of the time.
5. Reboot and shutdown, if the directory of OC_manager is /**home**/**pi**/**OC_manager**/ you can reboot and shutdown the system from here, useful for headless system.

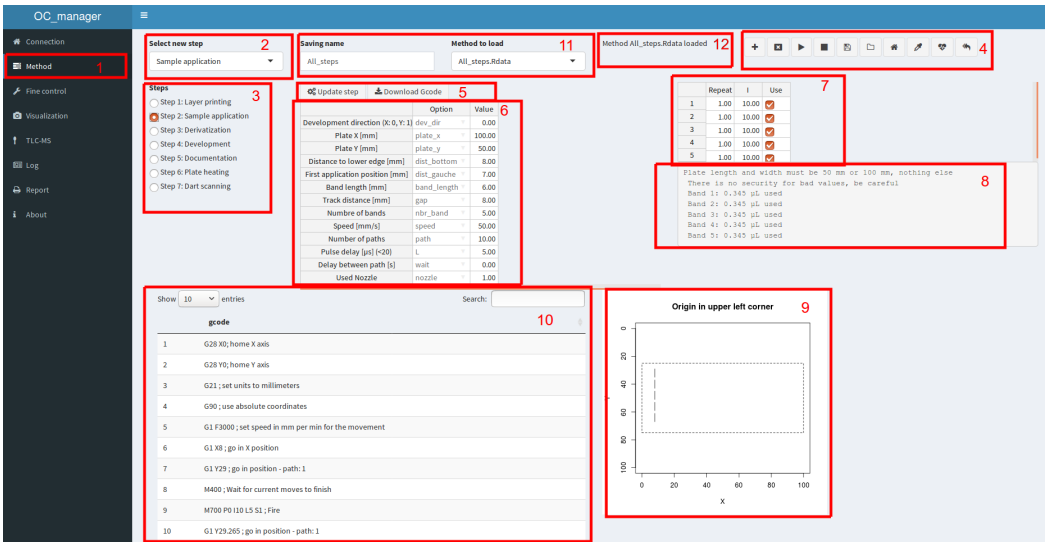


Figure 3: Method tab

Method

The core of OC_manager, this is where GCODE sequences are created, updated and sent to the machines.

1. Location of the tab
2. List of available steps. Select one and use the “+” button (4) to add it to the list of selected step (3).
3. List of selected step. Select one to update (5) and launch with the “play” button (4).
4. Action buttons. Add a step, delete a step, run a step, save, load etc... Note that help appear when each button it hovered, USE IT.
5. Specific button for the current step. Update the step and download the GCODE. When a step is added, no GCODE is associated with it, the step must be updated to do so, then GCODE (10), text feedback (8) and plotting feedback (9) will appear.
6. Step table. This is where you can modify the parameters, no details here, use the feedbacks for this (8-9).
7. Complementary step table: For some step, a complementary table is necessary. Generally, the number of row in this table is set in the main table, it may be necessary to first update after the main table is set, then update again with this complementary table set.
8. Text feedback after update
9. Plot feedback after update. This is may be the most important as it gives an idea of what the device will do.
10. GCODE table. Not available in classic machines of analytical chemistry, this table is here to give more details to the user.
11. Saving and loading method name. By using the “load” and “save” buttons (4), it is possible to access the method in an other session.
12. Text feedback. Another feedback which changes at each user action

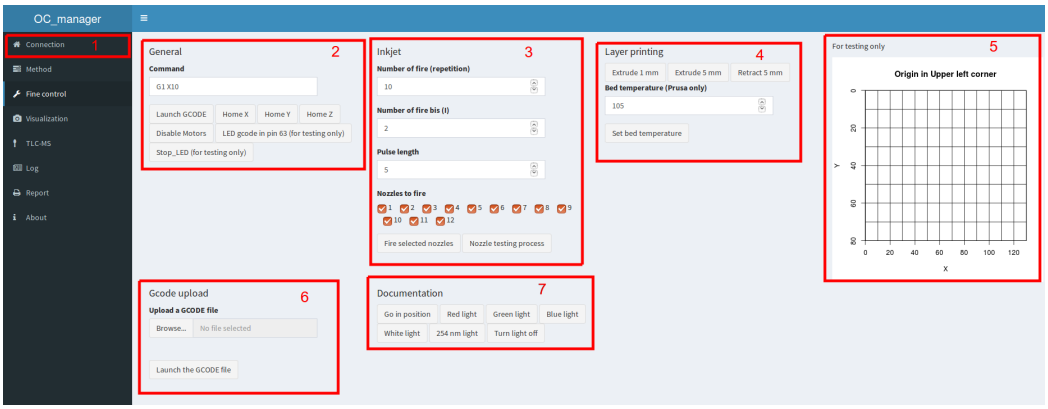


Figure 4: Fine control tab

Fine control

As this device is a prototype, it is often necessary to control it more in detail.

1. Location of the tab
2. General section, should apply to all device, you can launch a specific GCODE or home specifically an axis and release the motors. Look at the GCODE page of the reprap website to learn your gcode.
3. Inkjet section made specifically to fire drop with the inkjet, you can fire on selective nozzles, fire different amount of drop with different pulse width. The best here is to use the **nozzle testing process** which will fire each nozzles one after another in different position allowing to know which one are working.
4. Layer printing section, deprecated
5. Position section, can be used but not very useful anymore
6. Use this section to launch a hand made or modified GCODE.
7. Visualization section to turn on and off specific LEDs

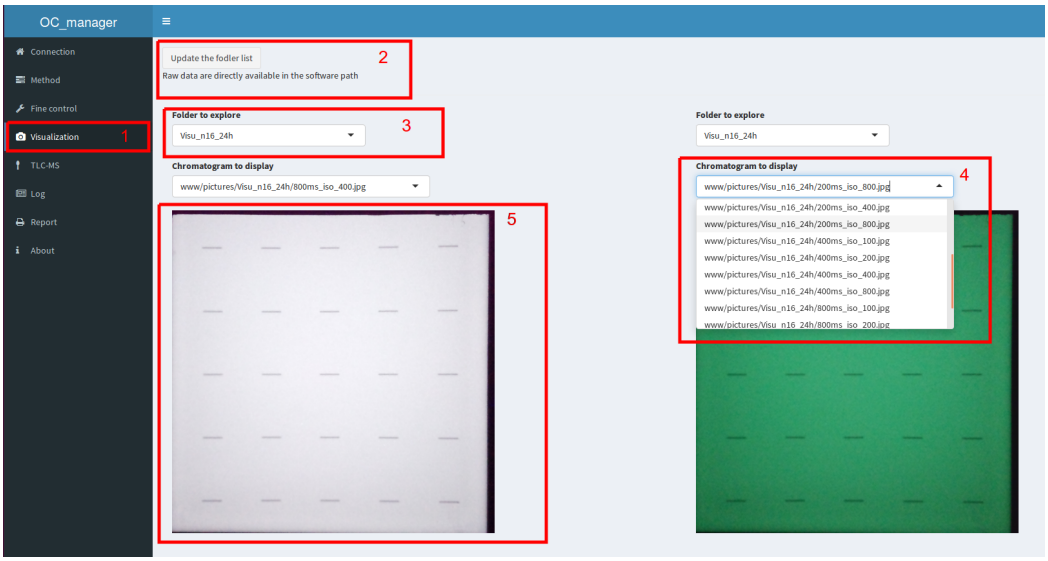


Figure 5: Visualization tab

Visualization

This tab allows to access the pictures chromatograms that were captured. Those files are also available in the pictures folder in the OC_manager folder.

1. Location of the tab
2. Update the folder to access the latest data.
3. Select a subfolder, note that the name of the plate set in the connection tab is present here.
4. Select one of the image in this folder, the choice is between light, ISO and exposure time.
5. Observe the picture.

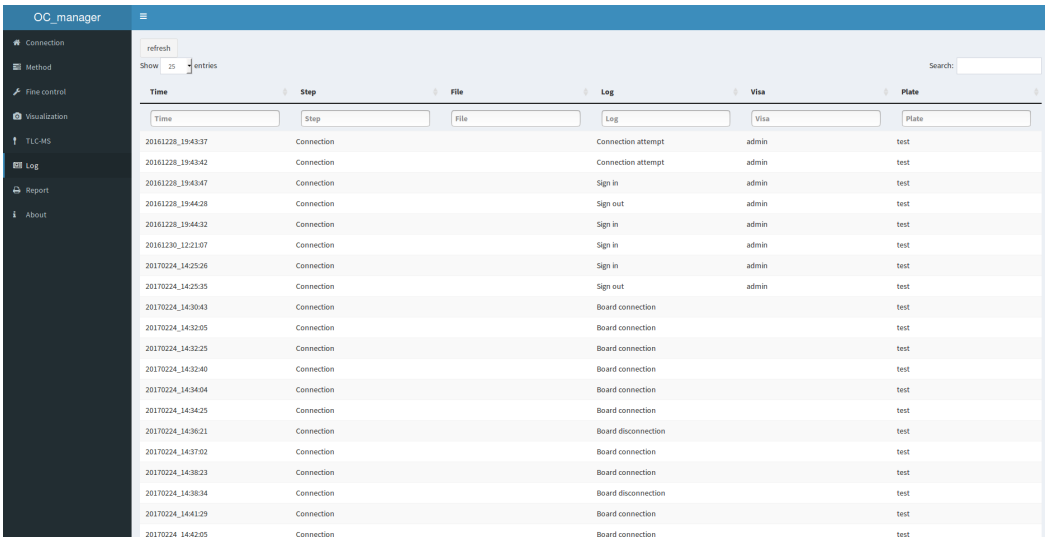


Figure 6: Log tab

Log

In OC_mannager, all actions are logged in a csv file which can be observed in this tab.