Forensic Science International
Volume 268, November 2016, Pages 81-91

# Age estimation of Calliphorida (Diptera: Calliphoridae) larvae using cuticular hydrocarbon analysis and Artificial Neural Networks

Hannah E. Moore[a*], John B. Butcher[b], Craig D. Adam[a], Charles R. Day[c], Falko P. Drijfhout[a]

[a]*School of Physical and Geographical Sciences, Keele University, Staffordshire, ST5 5BG, UK*

[b]*School of Life Sciences, Keele University, Staffordshire, ST5 5BG, UK*

[c]*School of Computing and Mathematics, Keele University, Staffordshire, ST5 5BG, UK*

**corresponding author: hanamoore@yahoo.co.uk*

- An ageing technique of forensically important larvae using cuticular hydrocarbons

- Analysed using Gas Chromatography – Mass Spectrometry

- Statistically analysed using Principal Component Analysis and Artificial Neural Networks

- Successfully age larvae of *Calliphora vicina* and *Calliphora vomitoria*

Forensic Science International
Volume 268, November 2016, Pages 81-91

# Age estimation of Calliphorida (Diptera: Calliphoridae) larvae using cuticular hydrocarbon analysis and Artificial Neural Networks

Hannah E. Moore, John B. Butcher, Craig D. Adam, Charles R. Day, Falko P. Drijfhout

## Abstract

Cuticular hydrocarbons were extracted daily from the larvae of two closely related blowflies *Calliphora vicina* and *Calliphora vomitoria* (Diptera:Calliphoridae). The hydrocarbons were then analysed using Gas Chromatography-Mass Spectrometry (GC-MS), with the aim of observing changes within their chemical profiles in order to determine the larval age. The hydrocarbons were examined daily for each species from 1 day old larvae until pupariation. The results show significant chemical changes occurring from the younger larvae to the post-feeding larvae. With the aid of a multivariate statistical method (Principal Component Analysis and Artificial Neural Networks), samples were clustered and classified, allowing for the larval age to be established. Results from this study allowed larvae to be aged to the day with at worst, 87% accuracy, which suggests there is great potential for the use of cuticular hydrocarbons present on larvae to give an indication of their age and hence potentially a valuable tool for minimum PMI estimations.

Keywords: cuticular hydrocarbons, GC-MS, PCA, ANN, Calliphoridae, larvae, minimum PMI estimations.

## Introduction

The main contributing factors required to establish the minimal Post-Mortem Interval ($PMI_{min}$)

are species identification and age determination of necrophagous fly larvae [1]. Since Calliphoridae species are known to be the first colonisers of decomposing remains in most circumstances, they are of great forensic importance and have been widely studied [2][3][4]. However, to determine the age of larvae can be challenging and often requires a highly knowledgeable forensic entomologist. The post-feeding larvae also remain to be a problematic stage to age due to the lack of morphological changes occurring with time [5] and therefore the analyst must rely heavily on growth charts. DNA-based analyses have been applied to the field of forensic entomology for over a decade [6]. They have been successfully used to identify and age forensically important species [7-11].

One technique that may have the potential to give the same accuracy as DNA-based techniques for an ageing tool is Cuticular Hydrocarbon (CHC) analysis [12].

CHCs are found in the lipid wax layer of nearly all insects. They have different functions in different species of insects and previous studies have shown their use for age estimation [1, 13-16], sex [17] and species identification [18-20].

The two main factors believed to be influential for the composition of hydrocarbon pools are development/genetic factors and physiological state/environmental conditions [21-23]. The changes observed during the blowflies' development may be affected by the environment they are exposed to. Larvae develop in warm, humid conditions (decomposing remains) and in this stage of their life cycle, they yield profiles consisting of a mixture of low and high molecular weight hydrocarbons.

CHC analysis has the potential to be able to age larvae as they develop through the larval life stages with a good degree of accuracy [1, 13-14, 24]. These characteristics therefore hold a lot of potential in the field of forensic entomology for PMI estimations. CHCs have the advantage of determining the age to the day, rather than to the life stage and compliment the current techniques used in the field. For more information on CHCs, the reader is referred to the following references [25-27].

Artificial neural networks (ANNs) are a machine learning approach that has been utilised

extensively in a variety of applications for pattern recognition and clustering. They have the ability to learn characteristics contained in a dataset and use this to recognise and classify new, previously unseen, data. ANN learning is achieved by altering the weighted connections between each of the artificial neurons contained within the network. This process of weight changes usually occurs until a suitable error (defined as the difference between the actual output of the network and the ground truth target output) is reached. ANNs are well-suited to processing noisy non-linear data and have the ability to learn the relationship between input and output data, making them well-suited to fast processing of large real-world datasets (e.g. speech recognition [28], robotics [29], structural engineering [30] as well as the forensic sciences [31].

Of the many types of ANNs that exist, the Self Organising Map (SOM) [32] is an unsupervised approach that clusters training data based on its underlying characteristics. A SOM consists of an input and an output layer containing artificial neurons, where each neuron in one layer is connected to every neuron in the next. The SOM maps multidimensional data with similar characteristics into topologically co-located clusters in its output layer, generating clusters which represent similar data points in its output layer once training is complete.

Briefly, the unsupervised training of a SOM involves the presentation of every input pattern to the input layer where the incoming weighted connections of the output neuron which best matches the input pattern (known as the winning neuron) are modified. This change in weights results in the formation of clusters within the output layer where similar input patterns cluster in close proximity in the output layer. Readers are referred to [32] and [33] for further details on SOM training.

The topological ordering capabilities of a SOM make it well suited to the analysis of high dimensional data such as that collected from blowfly species. This was shown in a previous study where the classification of hydrocarbon data collected from *Lucilia sericata* pupae was classified using a SOM providing classification accuracy that exceeded 89% [16].

The novelty of this study is to examine the CHC profiles of the larvae from two forensically

important blowfly species found in the UK, *Calliphora vicina* and *Calliphora vomitoria*, with the aim of determining whether there are distinguishable chemical changes occurring over time and to test the reliability of the method by investigating whether this technique is more generally applicable across species. Gas Chromatography-Mass Spectrometry (GC-MS) was used to analyse the HCs and methods of statistical and artificial intelligence data analysis were applied to the compiled datasets to cluster and automatically classify the data as well as aid data visualisation.

## Material and Materials

### *Insect materials*

A colony of *Calliphora vicina* and *Calliphora vomitoria,* (geographical origin, University of Birmingham campus, UK), kindly supplied by the Scott Hayward's research group at the University of Birmingham, was reared in the laboratory and maintained in separate rearing cages under standard environmental conditions (22□1 $^{o}$C). They were fed with sugar and water on a need-to basis and pigs liver, which was used as an oviposition medium. Once the eggs were laid they were separated into plastic containers containing approximately 700 eggs. This ensured there would be enough larvae for the completion of the life cycle with daily extractions and to resolve the problem of overcrowding and food competition. The larvae were fed daily with minced beef (approximately 50 g) and were kept in an incubator at a set temperature of 22±1 °C. Under the rearing temperature of 22±1 °C larvae reached the puparial stage after 11 days for C. vicina and 14 days for C. vomitoria. The hydrocarbons were extracted daily from the larvae upon hatching, until they pupariated.

### *Sample Preparation*

Liquid extraction with hexane was used to extract the hydrocarbons. For each sample ($n$=10) a number of larvae (Table 1) were pooled together to ensure the concentration was sufficient for the GC-MS to detect the hydrocarbons [12]. Preliminary experiments carried out at the rearing

temperature showed that approximately 20-30 first instar larvae yielded a sufficient concentration to produce a reliable chromatogram on the GC-MS. As the larvae became older, and therefore larger in size, fewer larvae were used for the extractions (Table 1). The larval sample numbers were reduced as they aged due to the hydrocarbon concentration increasing with larval size and therefore preventing the GC-MS from overloading. The larvae were placed into a GC vial with hexane ensuring that they were submerged. It was left for 10 to 15 minutes after which the hexane was removed and passed through a silica gel column. The column was made by plugging a Pasteur pipette with glass wool followed by a small amount of silica gel. The larval extract from the GC vial was transferred to the column and an additional 500 µl of hexane was added. The eluted hexane was collected into a clean GC vial and left until completely dry. The extracts were redissolved in 10 µl (for manual injections) and 20 µl (for autosampler injections) of hexane and a 2 µl aliquot was injected into the GC-MS. For the first 3 days extracts were injected manually and thereafter with the autosampler.

**Table 1**: Number of larvae used daily during *C. vicina* and *C. vomitoria* larval extractions

| Species | Day | Instar | Number of larvae added to each sample *n=10* |
|---|---|---|---|
| *C. vicina* | 1 | 1st instar | ~20 |
| | 2 | 2nd instar | 10 |
| | 3 | 3rd instar | 5 |
| | 4 & 5 | 3rd instar | 3 |
| | 6 to 11 | 3rd instar & post- feeding | 2 |
| *C. vomitoria* | 1 | 1st instar | ~30 |
| | 2 | 2nd instar | ~15 |
| | 3 | 2nd/3rd instar | 7 |
| | 4 & 5 | 3rd instar | 3 |
| | 6 & 10 | 3rd instar | 2 |
| | 11 | Post-feeding | 2 |
| | 12 to 14 | Post-feeding | 1 |

*Chemical analysis of extracts*

Chemical analysis of all extracts was carried out on an Agilent Technologies 6890N Network GC with a split/splitless injector at 250 $^{o}$C, a Restek Rxi-1MS capillary column (30m x 0.25 mm ID, 0.25µm film thickness) and coupled to an Agilent 5973 Network Mass Selective Detector. The GC was coupled to a computer and data processed with Agilent Chemstation software. Elution was carried out with helium at 1mL/min. The oven temperature was programmed to be held at 50 $^{o}$C for 2 minutes then ramped to 200 $^{o}$C at 25 $^{o}$C/min, then from 200 $^{o}$C to 260 $^{o}$C at 3$^{o}$C/min and finally from 260 $^{o}$C to 320 $^{o}$C at 20 $^{o}$C/min where it was held for 2 minutes. The mass spectrometer was operated in Electron Ionisation mode at 70 eV, scanning from 40 – 500 amu at 1.5 scans $s^{-1}$. Hydrocarbons were identified using a library search (NIST08), the diagnostic fragmented ions and the Kovats indices.

*Statistical analysis:*

*Principal Component Analysis*

The chromatograms were initially interpreted using Principal Component Analysis (PCA) in order to ease visualization of trends that may be present within the dataset. For further details the reader is referred to [12, 34]. PCA has already been used by the authors to age *Lucilia sericata* [13] and the same experimental design was used for this study. Methyl branched alkanes and alkenes with a percentage peak area greater than 0.5% were used for statistical analysis.

131 *Artificial Neural Networks*

132 The same training and testing approach reported in [16] is used in the current study to
133 automatically classify the larvae of both *Calliphora vicina* and *Calliphora vomitoria*,
134 where readers are referred to for further details. Briefly, the data was preprocessed using
135 PCA to reduce the dimensionality of the data, with six principal components that contained
136 the most variance used (percentage sum of Eigenvalues for PC1 to PC 6 - *C. vicina* 96%
137 and *C. vomitoria* 97%) as input data for the neural network. The data was then normalized
138 between the range -1 and +1 by:

$$v_n = \frac{v - \min(v)}{\max(v) - \min(v)} \times (U_{lim} - L_{lim}) - L_{lim}$$

139 where $v_n$ is the normalised data, $v$ is the original data, $max(v)$ and $min(v)$ are the maximum
140 and minimum data values respectively and $_{Ulim}$ and $_{Llim}$ are the desired upper and lower limits
141 of the normalised data which in this study are set to +1 and -1 respectively. Training data
142 was presented to the SOM by averaging five hydrocarbon profiles for both spec ies' larvae
143 and adult flies. Testing the generalization performance of the SOM was estimated using
144 two approaches:

145 1) Presenting the remaining unseen individual hydrocarbon profiles to the SOM
146 2) Presenting the average of the remaining unseen hydrocarbon profiles to the SOM

147 Training was performed using 10-fold cross-validation and an average and standard
148 deviation across the ten folds calculated. For each fold, a random subset for each day's

profile was chosen for the training and testing data, with a different subset chosen for each fold (note: the same subsets were chosen for both testing approaches). In this study, two important SOM training parameters, the neighbourhood size and learning rate, were updated during training after a set number of elapsed epochs as described in Day et al [33]. A number of candidate output layer sizes were systematically evaluated to find the output layer size that delivered the most effective clusters.

## Results

*GC-MS analysis:*

*C. vicina*

*C. vicina* yielded a profile of 46 different identifiable compounds with some co-eluting resulting in a total of 40 resolvable peaks from day 1 to 11 (Table 2). All compounds were hydrocarbons consisting of *n*-alkanes (33%), alkenes (17%), and methyl branched hydrocarbons (50%) for day 1. The chain lengths ranged from C20:H to C33:H.

The lower molecular weight compounds in the profile (C20:H to C22:H) are mainly made up of volatile compounds which are chemically less stable and variable, therefore they were not used for subsequent PCA and ANN analysis as they showed little significance in ageing the larvae. The middle region of the chromatogram consists of straight chain *n*-alkanes, alkenes and methyl branched alkanes (ranging from C23:H to C27:H). The higher end of the chromatogram is dominated by high boiling point *n*-alkanes (ranging from C29:H to C33:H) and methyl branched alkanes which are at their most abundant in the 1$^{st}$ and 2$_{nd}$ instar larvae.

**Table 2** List of all compounds extracted from the larvae of *C. vicina* and their calculated Kovats Indices to aid identification. Compounds in bold were used for subsequent PCA analysis (peak numbers refer to numbers in Figure 1)

| Peak number | Peak Identification | Kovats iu |
|---|---|---|
| 1 | **Eicosene**[1] | 1990 |
| 2 | Eicosane | 2000 |
| 3 | Heneicosane | 2100 |
| 4 | **Docosene**[1] | 2190 |
| 5 | Docosane | 2200 |
| 6 | Tricosane | 2300 |
| 7 | **7-Methyltricosane** | 2342 |
| 8 | **5-Methyltricosane** | 2351 |
| 9 | **3-Methyltricosane** | 2373 |
| 10 | Tetracosane | 2400 |
| 11 | **2-Methyltetracosane** | 2465 |
| 12 | **Pentacosene**[1] | 2471 |
| 13 | **Pentacosene**[1] | 2479 |
| 14 | Pentacosane + Phthalate | 2500 |
| 15 | **11-Methylpentacosane** | 2536 |
| 16 | **9-Methylpentacosane** | 2538 |
| 17 | **7-Methylpentacosane** | 2544 |
| 18 | **5-Methylpentacosane** | 2552 |
| 19 | **3-Methylpentacosane** | 2574 |
| 20 | Hexacosane | 2600 |
| 21 | **x,12-Dimethylhexacosane**[2] **+ Heptacosene**[1] | 2666 |
| 22 | **Heptacosene**[1] | 2676 |
| 23 | **Heptacosene**[1] | 2679 |
| 24 | Heptacosane | 2700 |
| 25 | **11+13-Methylheptacosane** | 2735 |
| 26 | **7-Methylheptacosane** | 2743 |
| 27 | **5-Methylheptacosane** | 2753 |
| 28 | **3-Methylheptacosane** | 2775 |
| 29 | Octacosane | 2800 |
| 30 | **2-Methyloctacosane** | 2871 |
| 31 | Nonacosane | 2900 |
| 32 | **11+13-Methylnonacosane** | 2936 |
| 33 | **9-Methylnonacosane** | 2941 |
| 34 | **7-Methylnonacosane** | 2947 |
| 35 | **3-Methylnonacosane** | 2977 |
| 36 | Triacontane | 3000 |
| 37 | **2-Methyltricontane** | 3067 |
| 38 | **2,6/2,8/2,10-Dimethyltriacontane**[2] | 3097 |

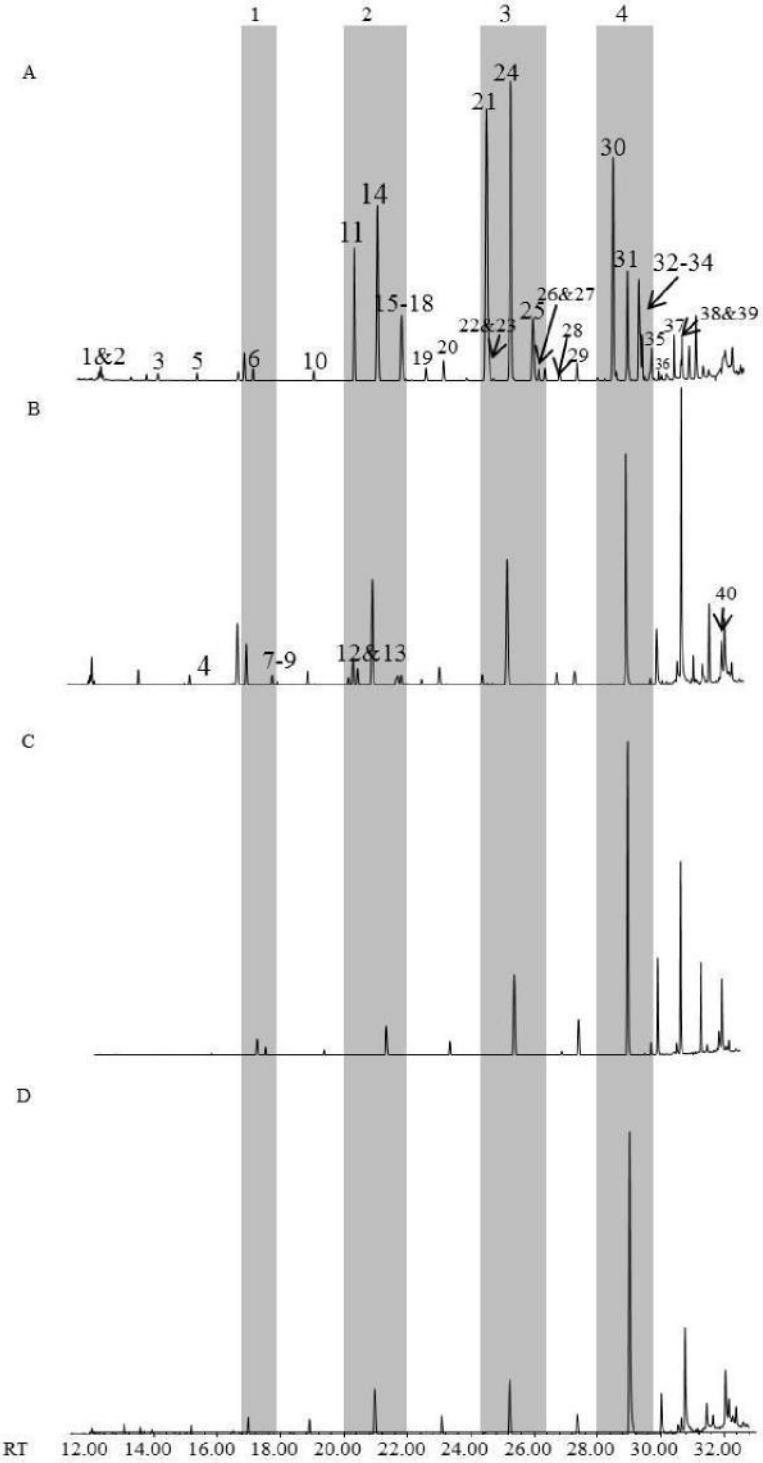
| | | | |
|---|---|---|---|
| 39 | Hentriacontane | 3100 |
| 40 | Tritriacontane | 3200 |

[1]Double bond position assumed but not assigned to specific peaks

[2]Tentative Identification based on calculated Kovats Indicies values and match with NIST08 Library database

Figure 1 shows the stacked GC chromatograms of a single sample (Table 1) from larvae extracted at days 1, 5, 7 and 11. The shaded bars highlight areas of contrast within the profiles showing potential for ageing the larvae of *C. vicina* because of the distinguishable chemical changes occurring with time.

The profile of the 1st instar (day 1) possesses no peaks specific to this stage. The 2nd instar, represented in day 2, also reveals no age specific compounds but there are eight peaks specific to both the 1st and 2nd instar stage (peak 26, 27, 32-34, 37, 38). The number of methyl branched compounds is greatly reduced from the immature larvae stages (1st and 2nd) to the 3rd instar. This can be seen in Figure 2 in the highlighted areas on the two chromatograms of A (day 1-1st instar) and B (day 5 – 3rd instar).




187

188 **Figure 1**: GC chromatograms of *C. vicina* larvae at four different ages, A: Day 1, B: Day 5, C: Day 7 and

189 D: Day 11. Shaded bars illustrate distinctive changes over time indicating specific areas of interest

Day 3 to day 5 reveal very similar chromatograms but they can be aged to the day rather than to the instar (3rd instar) because of the varying peak area ratios (Table 4). This instar has three compounds that are specific to this phase, 7-MeC23:H, 5-MeC23:H and 3-MeC23:H. This group of MeC23:H isomers (peak 7 to 9 in Figure 1) could be a very good age indicator for the 3rd instar stage, with the 7-MeC23:H (peak 7) also increasing further with age during the 3rd instar.

Peak 21 (x,12-DimeC26:H) co-elutes with C27:1 in the younger larvae (up to day 2), making it a good age indicator. C21:1 and C22:1 (peaks 1 & 4) both increase gradually as the larvae age (up to 23% and 16% respectively), before decreasing in the late post-feeding stage at day 11.

*GC-MS analysis: C. vomitoria*

*C. vomitoria* exhibited a profile of 57 identifiable compounds with some co-eluting giving a total of 51 resolvable peaks (Table 5). Of these 51 resolvable peaks, 96% were hydrocarbons, with day 1 consisting of *n*-alkanes (28%), alkenes (29%) and mono-methyl alkanes (43%). The chain length of all hydrocarbons range from C16:H to C33:H.

**Table 3:** List of all compounds from C20:H extracted from the larvae of *C. vomitoria* and their Kovats Indices to aid identification. Compounds in bold were used for subsequent PCA (peak numbers refer to numbers in Figure 3)

| Peak Number* | Peak Identification | Kovats iu |
|---|---|---|
| 1 | Eicosane | 2000 |
| 2 | Heneicosene[1] | 2066 |
| 3 | Heneicosene[1] | 2074 |
| 4 | Heneicosane | 2100 |
| 5 | **3-Methylheneicosane** | 2172 |
| 6 | Docosane | 2200 |
| 7 | **2-Methyldocosane** | 2264 |
| 8 | Tricosene[1] | 2271 |
| 9 | Tricosene[1] | 2278 |
| 10 | Tricosane | 2300 |
| 11 | **9+11-Methyltricosane** | 2339 |
| 12 | **7-Methyltricosane** | 2343 |
| 13 | **5-Methyltricosane** | 2348 |
| 14 | **3-Methyltricosane** | 2374 |
| 15 | Tetracosane | 2400 |
| 16 | **2-Methyltetracosane** | 2464 |
| 17 | Pentacosene[1] | 2471 |
| 18 | Pentacosene[1] | 2478 |
| 19 | Pentacosane | 2500 |
| 20 | **9+11-Methylpentacosane** | 2536 |
| 21 | **7-Methylpentacosane** | 2539 |
| 22 | **5-Methylpentacosane** | 2549 |
| 23 | **3-Methylpentacosane** | 2574 |
| 24 | Hexacosane | 2600 |
| 25 | **2-Methylhexacosane** | 2665 |
| 26 | Heptacosene[1] | 2669 |
| 27 | Heptacosene[1] | 2679 |
| 28 | Heptacosane | 2700 |
| 29 | **9+11-Methylheptacosane** | 2735 |
| 30 | **3-Methylheptacosane** | 2774 |
| 31 | Octacosane | 2800 |
| 32 | **2-Methyloctacosane** | 2871 |
| 33 | Nonacosene[1] | 2879 |
| 34 | Nonacosene[1] | 2886 |
| 35 | Nonacosane | 2900 |
| 36 | **11+13-Methylnonacosane** | 2937 |
| 37 | **7-Methylnonacosane** | 2948 |

| | | |
|---|---|---|
| 38 | **5-Methylnonacosane** | 2957 |
| 39 | **Dimethylnonacosane** | 2966 |
| 40 | **3-Methylnonacosane** | 2978 |
| 41 | Tricontane | 3000 |
| 42 | Cholesterol + Hentriacontene[1] | 3070 |
| 43 | Hentriacontene[1] | 3077 |
| 44 | Hentriacontene[1] | 3085 |
| 45 | Hentriacontane | 3100 |
| 46 | **11+15-Methylhentriacontane** | 3131 |
| 47 | **3-Methylhentriacontane** | 3175 |
| 48 | Dotriacontane | 3200 |
| 49 | Tritriacontene[1] | 3263 |
| 50 | Tritriacontene[1] | 3276 |
| 51 | Tritriacontane | 3300 |

[1] Double bond position assumed but not assigned to specific peaks

Figure 2 shows the chromatograms of a single sample (Table 1) from days 1, 4, 8 and 13. Chemical distinctions can be made between the different ages from the chromatogram comparison over time. The shaded bars highlight some regions of contrast within the profiles of the different ages.

Again, the lower retention time compounds of the profile are mainly made up of volatile compounds which are less stable and therefore excluded from PCA and ANN analysis. As with the chemical profile of *C. vicina*, the middle section of the chromatogram consists of *n*alkanes, alkenes and methyl branched alkanes (ranging from C23:H to C27:H). The higher end of the chromatogram is dominated by high molecular weight *n*-alkanes (ranging from C29:H to C33:H) and methyl branched alkanes which are at their most abundant in the 3rd instar larvae.

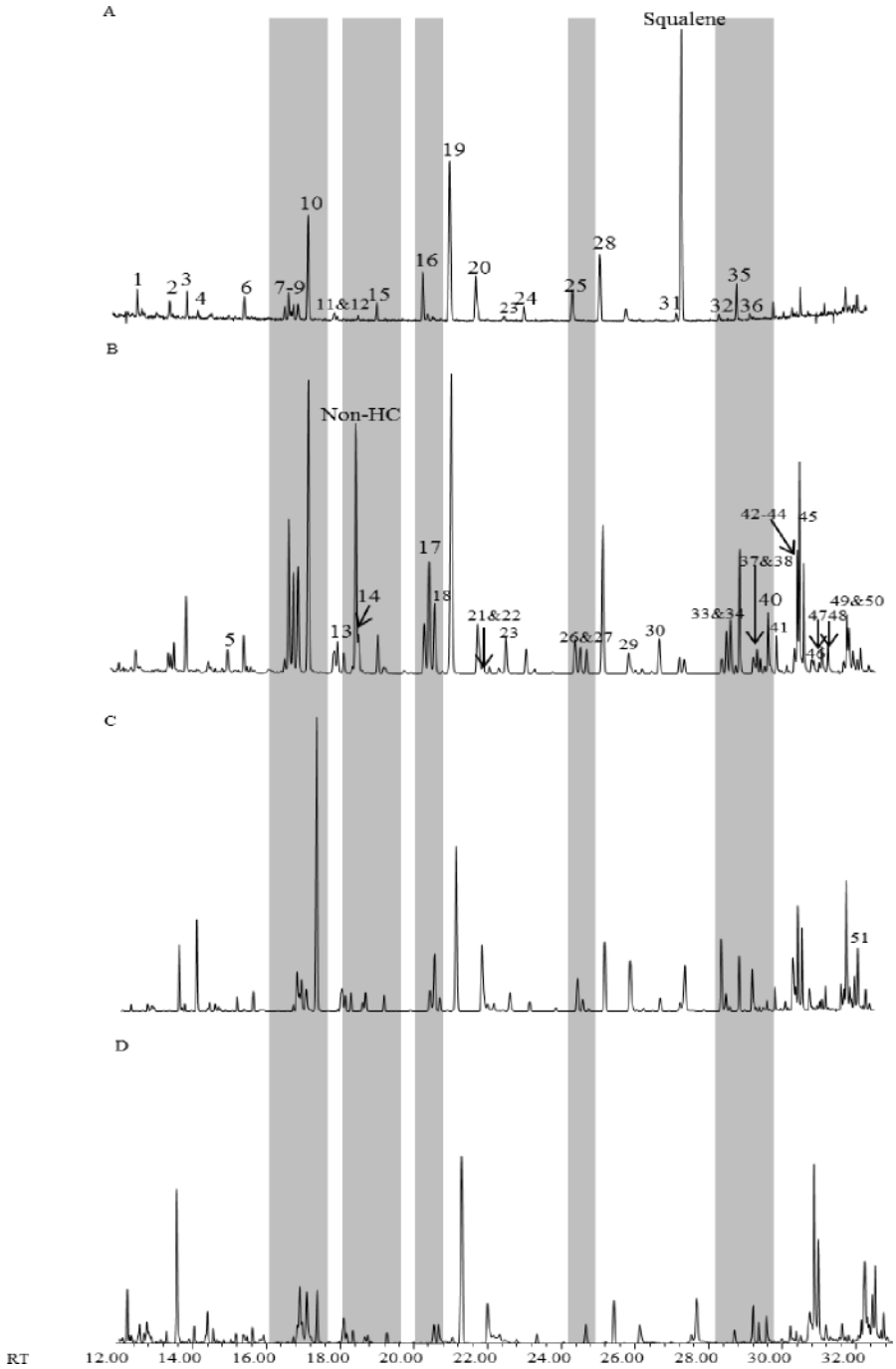



**Figure 2.** GC chromatograms of *C. vomitoria* larvae at four different ages, A: Day 1, B: Day 4, C: Day 8 and D: Day 13. Shaded bars illustrate distinctive changes over time indicating specific areas of interest

There are no compounds exclusive to the profile of day 1 larvae (1st instar). However, there are a few compounds specific to certain larval ages, such as 3-MeC29:H and 3-MeC31:H, which are only seen with a relative peak area percentage greater than 0.5% in day 4. 3-MeC27:H and an unidentified DimeC29:H are specific to day 9 larvae (in a peak area exceeding 0.5%). These compounds are therefore likely to be good age indicators for these two larval ages.

2-MeC22:H is only present in the early larval life stages (days 1 to 3) and decreases with age so the presence of this compound could be used to determine the early larval life age (1st and 2nd instar). 7-MeC23:H and 3-MeC23:H (Table 4) are present in a high concentration in day 3. 2-MeC24:H is present in substantial concentrations in days 1 and 2 then decrease with age, before an increase is seen in the late post-feeding stage. 9+11-MeC25:H and 2-MeC26:H also appear in a very high concentration in the immature larvae stages (days 1 and 2). Three peaks are absent (below the peak area threshold of 0.5%) from the 1st and 2nd instar larvae (3-MeC21:H, 9+11-MeC27:H and 11+15-MeC31:H) and three compounds are absent from the 1st instar larvae alone (compounds 5-MeC23:H, 3-MeC23:H and 7-MeC25:H). The absence of these compounds from the immature life stages makes them good age indicators.

A group of MeC23:H compounds (9+11-MeC23:H, 5-MeC23:H and 3-MeC23:H) increases significantly in the post-feeding stage (days 10 to 14), implying that they could act as a good post-feeding stage indicator. 7-MeC23:H is not present in a detectable concentration in the late post-feeding stage (days 10 and 11), which again could potentially be a useful age indicator for this life stage. 11+15-MeC31:H is relatively stable over the first three instars after which a considerable increase in the peak area is observed in the post-feeding stage.

*Principal Component Analysis:*

*C. vicina:*

PCA analysis was initially carried out including all three classes of hydrocarbons (*n*-alkanes, alkenes and methyl branched alkanes). However, this gave a lot of scatter and cluster overlap such that an accurate age could not be determined for individuals. A dataset containing just the alkenes and methyl branched alkanes was therefore compiled.

Of the 40 resolvable peaks extracted from the cuticle of *C. vicina* larvae, 27 hydrocarbon peaks were used for PCA analysis, of which 76% were methyl branched and 24% were alkenes. Table 4 shows the compounds used for PCA analysis, along with the total percentage of each compound present, the percentage standard deviation for each day and the calculated Kovats Indices.

**Table 4:** List of the compounds extracted and used for subsequent PCA analysis from the larvae of *C. vicina*, with the total percentage of each compound present, the percentage standard deviation for each day and the calculated Kovats Indices to aid identification.

| Peak number | Peak Identification | Kovats iu | Day 1 *n*=10 % | Day 2 *n*=10 % | Day 3 *n*=10 % | Day 4 *n*=10 % | Day 5 *n*=10 % | Day 6 *n*=10 % | Day 7 *n*=10 % | Day 8 *n*=10 % | Day 9 *n*=10 % | Day 10 *n*=10 % | Day 11 *n*=10 % |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | Eicosene[1] | 1990 | 1.79±0.68 | 3.52±1.54 | 5.48±2.47 | 3.55±0.64 | 4.55±2.12 | 7.19±3.07 | 6.61±1.36 | 9.62±2.60 | 23.23±17.18 | 23.77±5.38 | 14.59±12.54 |
| 4 | Docosene[1] | 2190 | tr | tr | tr | tr | 3.26±1.51 | 5.52±1.35 | 5.79±1.24 | 8.59±2.47 | 17.03±10.91 | 16.43±4.43 | 11.93±9.49 |
| 7 | 7-Methyltricosane | 2342 | tr | tr | 4.43±2.13 | 4.76±1.49 | 5.60±3.23 | tr | tr | tr | tr | tr | tr |
| 8 | 5-Methyltricosane | 2351 | tr | tr | 2.44±1.23 | 2.21±0.76 | 2.92±1.77 | tr | tr | tr | tr | tr | tr |
| 9 | 3-Methyltricosane | 2373 | tr | tr | 2.48±1.25 | tr | 2.79±1.55 | tr | tr | tr | tr | tr | tr |
| 11 | 2-Methyltetracosane | 2465 | 12.55±4.83 | 8.12±5.46 | 5.07±3.02 | 4.40±1.40 | 5.05±3.13 | 4.75±1.16 | 8.14±2.93 | 13.23±9.07 | 20.76±13.98 | 15.78±5.01 | 18.20±15.51 |
| 12 | Pentacosene[1] | 2471 | tr | tr | 13.88±7.34 | 14.89±5.04 | 12.35±6.91 | 13.06±5.15 | tr | tr | tr | tr | tr |
| 13 | Pentacosene[1] | 2479 | tr | tr | 11.24±6.25 | 11.79±4.36 | 8.94±5.27 | 8.99±4.20 | tr | tr | tr | tr | tr |
| 15 | 11-Methylpentacosane | 2536 | 9.90±3.93 | 5.34±1.26 | 4.59±2.93 | 4.17±1.37 | 4.53±2.49 | 4.63±1.63 | 6.78±1.62 | 8.87±2.63 | tr | tr | tr |
| 16 | 9-Methylpentacosane | 2538 | 2.49±1.08 | 2.13±0.44 | 3.74±2.15 | 4.42±1.69 | 5.62±3.04 | 5.33±2.03 | 4.88±0.78 | 5.78±1.78 | tr | tr | tr |
| 17 | 7-Methylpentacosane | 2544 | 1.48±0.76 | 2.66±0.46 | 4.80±3.12 | 6.15±2.70 | 8.43±4.54 | 7.25±2.50 | 6.21±1.73 | tr | tr | tr | tr |
| 18 | 5-Methylpentacosane | 2552 | 1.16±0.60 | 1.63±0.68 | 2.09±1.47 | 2.49±1.55 | 3.80±2.98 | 4.42±1.98 | 3.38±1.73 | tr | tr | tr | tr |
| 19 | 3-Methylpentacosane | 2574 | 2.19±1.21 | 2.41±0.87 | 4.06±2.68 | 4.97±1.97 | 4.92±2.66 | 4.90±1.78 | 5.24±1.21 | 5.76±1.95 | tr | tr | tr |
| 21 | x,12-Dimethylhexacosane[2] + Heptacosene | 2666 | 37.91±26.98 | 38.39±25.62 | 17.51±11.18 | 14.61±4.73 | 8.87±4.84 | 9.41±2.44 | 16.12±5.84 | 29.33±15.37 | 38.97±22.46 | 44.02±15.38 | 55.28±39.23 |
| 22 | Heptacosene[1] | 2676 | 1.97±0.88 | tr | tr | 4.79±1.70 | 4.30±2.50 | tr | tr | tr | tr | tr | tr |
| 23 | Heptacosene[1] | 2679 | 0.58±0.35 | 1.95±0.31 | 4.32±2.31 | 4.56±1.96 | 3.12±1.73 | tr | tr | tr | tr | tr | tr |
| 25 | 11+13-Methylheptacosane | 2735 | 8.29±3.64 | 6.45±4.78 | 3.98±2.00 | tr | tr | tr | tr | tr | tr | tr | tr |
| 26 | 7-Methylheptacosane | 2743 | 1.47±0.86 | 1.67±1.59 | tr | tr | tr | tr | tr | tr | tr | tr | tr |
| 27 | 5-Methylheptacosane | 2753 | 1.37±0.84 | 1.63±1.52 | tr | tr | tr | tr | tr | tr | tr | tr | tr |

| | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 28 | 3-Methylheptacosane | 2775 | 1.03±0.86 | 2.08±1.44 | 6.15±3.65 | 8.10±3.06 | 8.04±4.85 | 11.47±4.61 | 17.47±1.83 | 8.76±3.51 | tr | tr | tr |
| 30 | 2-Methyloctacosane | 2871 | 8.53±7.04 | 11.28±16.23 | 3.74±2.65 | 4.14±2.32 | 2.91±1.67 | 4.46±1.21 | 4.95±1.35 | tr | tr | tr | tr |
| 32 | 11+13-Methylnonacosane | 2936 | 2.75±2.87 | 3.78±5.39 | tr | tr | tr | tr | tr | tr | tr | tr | tr |
| 33 | 9-Methylnonacosane | 2941 | 1.01±0.97 | 1.48±1.97 | tr | tr | tr | tr | tr | tr | tr | tr | tr |
| 34 | 7-Methylnonacosane | 2947 | 1.25±1.22 | 1.85±2.58 | tr | tr | tr | tr | tr | tr | tr | tr | tr |
| 35 | 3-Methylnonacosane | 2977 | tr | tr | tr | tr | tr | 8.63±4.57 | 14.44±3.31 | 10.05±7.41 | tr | tr | tr |
| 37 | 2-Methyltricontane | 3067 | 1.46±1.15 | 2.77±2.34 | tr | tr | tr | tr | tr | tr | tr | tr | tr |
| 38 | 2,6/2,8/2,10-Dimethyltriacontane[2] | 3097 | 0.81±0.92 | 0.88±0.63 | tr | tr | tr | tr | tr | tr | tr | tr | tr |

[1]Double bond position assumed but not assigned to specific peaks

[2]Tentative identification based on calculated Kovats Indices values and match with NIST08 Library database

tr = Trace amounts detected <0.5%

PCA was carried out using the first 6 principal components which described 96.9% of the variation within the data set with the first three principal components, comprising 60.5%, 20.3% and 8.9% respectively.

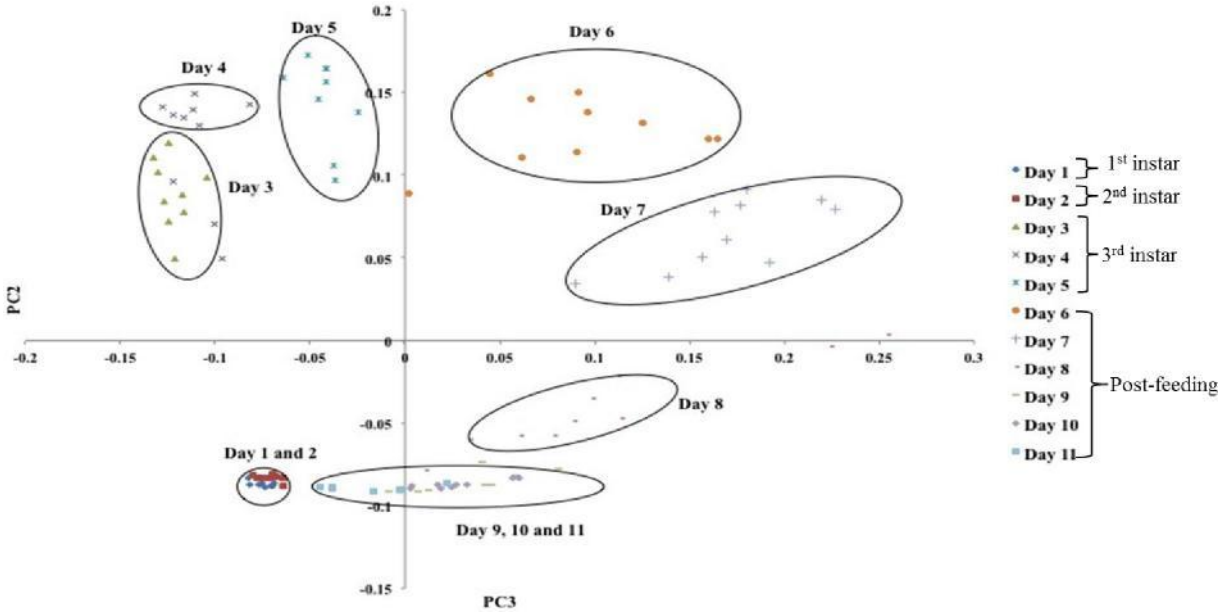



**Figure 3**: PCA plot showing PC3 against PC2 for *C. vicina* larvae using alkenes and methyl branched alkanes only, with clustering days circled

Figure 3 shows the PCA plot of PC3 vs PC2 for data gathered from day 1 to day 11 of larvae extractions of *C. vicina*. There are eight clusters within the plot allowing for the larvae to be aged down to the day with the exception of days 1 and 2 and the late post-feeding stage (days 9 to 11). The clustering groups follow a systematic pattern clockwise around the plot. Although day 1 has a single compound specific to that day, day 1 and day 2 have eight compounds detectable only in those two life stages (1st and 2nd instar), hence the likely reasoning for these two stages clustering together in the PCA plot. There is a substantial change within the PCA plot from day 2 to day 3, which represents the transition between the 2nd and 3rd instar.



20

The main compounds which have substantial PCA score values are x,12-diMethylhexacosane which co-elutes with heptacosene in the early larval stages. Another compound exhibiting a high score is 2-Methyltetracosane. Both these methyl branched hydrocarbons are present throughout the larval development of *C. vicina* and this indicates that the methyl branched alkanes are influential for ageing this species.

*C. vomitoria:*

Of the 51 compounds extracted from the cuticle of *C. vomitoria*, 29 of them were used for subsequent PCA analysis. Table 5 shows the compounds used for PCA analysis, along with the total percentage of each compound present, the percentage standard deviation for each day and the calculated Kovats Indices.

292 **Table 5**: List of the compounds extracted and used for subsequent PCA analysis from the larvae of *C. vomitoria*, along with the total percentage of each compound present,
293 the percentage standard deviation for each day and the calculated Kovats Indices to aid identification
294

| | | | Day 1 | Day 2 | Day 3 | Day 4 | Day 5 | Day 6 | Day 7 |
|---|---|---|---|---|---|---|---|---|---|
| **Peak** | **Peak** | **Kovats** | ***n*=10** | ***n*=10** | ***n*=10** | ***n*=10** | ***n*=10** | ***n*=10** | ***n*=10** |
| **no.** | **Identification** | **iu** | **%** | **%** | **%** | **%** | **%** | **%** | **%** |
| 5 | 3-Methylheneicosane | 2172 | tr | tr | 3.97±0.68 | 3.16±0.61 | 5.47±2.49 | 4.27±0.51 | 3.70±1.00 |
| 7 | 2-Methyldocosane | 2264 | 5.44±2.86 | 5.06±1.56 | 2.58±1.05 | tr | tr | tr | tr |
| 11 | 9+11-Methyltricosane | 2339 | 5.32±4.90 | 6.31±2.67 | 7.47±2.17 | 4.96±1.65 | 8.30±3.96 | 7.93±1.13 | 7.47±2.21 |
| 12 | 7-Methyltricosane | 2343 | 2.53±1.13 | 4.83±1.72 | 7.31±1.51 | 4.78±0.89 | 5.15±2.20 | 4.33±0.48 | 3.66±0.84 |
| 13 | 5-Methyltricosane | 2348 | tr | 2.11±0.72 | 4.53±0.93 | 3.24±0.70 | 4.58±1.95 | 4.37±0.58 | 3.98±0.90 |
| 14 | 3-Methyltricosane | 2374 | tr | 3.19±1.08 | 7.15±1.76 | 5.36±1.26 | 4.81±2.19 | 4.77±1.08 | 4.18±1.03 |
| 16 | 2-Methyltetracosane | 2464 | 22.30±24.82 | 22.76±11.56 | 13.06±9.31 | 6.34±1.38 | 7.65±4.12 | 7.87±1.46 | 7.17±2.01 |
| 20 | 9+11-Methylpentacosane | 2536 | 31.70±44.99 | 30.31±18.40 | 19.89±13.24 | 10.63±1.90 | 17.14±7.75 | 17.94±2.32 | 17.56±4.49 |
| 21 | 7-Methylpentacosane | 2539 | tr | 3.19±1.47 | 3.63±0.91 | 3.82±1.77 | 5.14±2.78 | 3.24±0.91 | 3.28±1.04 |
| 22 | 5-Methylpentacosane | 2549 | tr | tr | tr | tr | 4.35±2.52 | 2.49±0.96 | 3.48±1.03 |
| 23 | 3-Methylpentacosane | 2574 | 5.45±6.96 | 4.44±2.04 | 5.87±1.99 | 7.69±1.25 | 7.27±3.24 | 6.41±0.79 | 6.14±1.57 |
| 25 | 12,22-Dimethylhexacosane[1], 12,20-Dimethylhexacosane[1] | 2665 | 15.26±22.60 | 12.44±6.84 | 9.61±6.20 | 5.58±1.39 | 9.86±4.40 | 11.84±1.56 | 13.45±2.92 |
| 29 | 9+11-Methylheptacosane | 2735 | tr | tr | 2.61±0.96 | 10.56±2.76 | 6.31±2.72 | 4.31±0.63 | 4.22±1.18 |
| 30 | 3-Methylheptacosane | 2774 | tr | tr | tr | tr | tr | tr | tr |
| 32 | 2-Methyloctacosane | 2871 | 5.98±4.83 | 2.36±0.86 | 3.56±2.47 | 3.67±0.99 | 8.98±3.96 | 11.69±1.06 | 12.46±2.63 |
| 36 | 11+13-Methylnonacosane | 2937 | 6.02±5.39 | 2.99±1.35 | 4.27±2.98 | 3.82±1.08 | tr | 4.38±0.51 | 4.81±2.14 |
| 37 | 7-Methylnonacosane | 2948 | tr | tr | tr | 3.31±1.80 | tr | tr | tr |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 38 | 5-Methylnonacosane | 2957 | tr | tr | tr | 3.39±1.12 | 0.10±0.32 | tr | tr |
| 39 | Dimethylnonacosane | 2966 | tr | tr | tr | tr | tr | tr | tr |
| 40 | 3-Methylnonacosane | 2978 | tr | tr | tr | 9.62±2.57 | tr | tr | tr |
| 46 | 11+15-Methylhentriacontane | 3131 | tr | tr | 4.50±5.66 | 4.26±1.97 | 4.88±2.35 | 4.16±0.49 | 4.44±1.28 |
| 47 | 3-Methylhentriacontane | 3175 | tr | tr | tr | 5.82±1.96 | tr | tr | tr |

295

| | | | Day 8 | Day 9 | Day 10 | Day 11 | Day 12 | Day 13 | Day 14 |
|---|---|---|---|---|---|---|---|---|---|
| **Peak** | **Peak** | **Kovats** | ***n*=10** | ***n*=10** | ***n*=10** | ***n*=10** | ***n*=10** | ***n*=10** | ***n*=10** |
| **no.** | **Identification** | **iu** | **%** | **%** | **%** | **%** | **%** | **%** | **%** |
| 5 | 3-Methylheneicosane | 2172 | 2.97±1.13 | 2.87±1.03 | 2.83±0.43 | 3.92±0.89 | 5.54±1.93 | 7.10±2.82 | 4.55±1.97 |
| 7 | 2-Methyldocosane | 2264 | tr | tr | tr | tr | tr | tr | tr |
| 11 | 9+11-Methyltricosane | 2339 | 6.93±1.36 | 5.67±2.14 | 7.12±2.12 | 9.19±1.14 | 10.11±4.36 | 10.66±7.84 | 9.86±10.06 |
| 12 | 7-Methyltricosane | 2343 | 3.33±0.62 | 2.57±0.92 | 3.11±0.67 | 4.48±0.77 | 4.89±1.95 | tr | tr |
| 13 | 5-Methyltricosane | 2348 | 4.01±0.83 | 3.68±1.19 | 3.82±0.88 | 5.21±0.84 | 6.37±2.18 | 6.12±3.13 | 7.51±10.51 |
| 14 | 3-Methyltricosane | 2374 | 4.05±1.03 | 3.90±1.41 | 4.00±1.00 | 4.83±0.86 | 4.97±2.17 | 5.32±2.00 | 6.88±10.70 |
| 16 | 2-Methyltetracosane | 2464 | 6.76±1.48 | 6.25±2.17 | 7.86±6.63 | 6.19±1.05 | 5.32±2.40 | 9.71±6.52 | 14.59±9.64 |
| 20 | 9+11-Methylpentacosane | 2536 | 16.23±2.94 | 14.32±5.08 | 18.29±12.81 | 16.82±1.87 | 16.33±7.39 | 18.98±13.12 | 15.29±9.74 |
| 21 | 7-Methylpentacosane | 2539 | 2.78±0.36 | 2.08±0.85 | 3.22±1.22 | 4.37±1.14 | 3.75±2.88 | tr | tr |
| 22 | 5-Methylpentacosane | 2549 | 3.12±0.37 | 2.81±1.04 | 3.66±1.12 | 4.60±1.20 | 4.39±3.29 | tr | tr |
| 23 | 3-Methylpentacosane | 2574 | 5.86±0.74 | 5.55±1.81 | 5.21±2.07 | 5.01±0.95 | 4.09±1.22 | tr | tr |
| 25 | 12,22-Dimethylhexacosane[1], 12,20- | 2665 | 13.96±2.49 | 14.18±4.56 | 14.66±7.08 | 11.48±1.12 | 10.57±4.14 | 14.00±8.26 | 13.51±9.57 |

296

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | Dimethylhexacosane[1] | | | | | | | | |
| 29 | 9+11-Methylheptacosane | 2735 | 4.13±0.64 | 4.58±1.12 | 4.32±0.96 | 3.71±0.60 | tr | tr | tr |
| 30 | 3-Methylheptacosane | 2774 | tr | 2.05±0.63 | tr | tr | tr | tr | tr |
| 32 | 2-Methyloctacosane | 2871 | 12.95±2.97 | 13.66±3.43 | 9.00±1.24 | 7.00±0.79 | 6.17±2.60 | 6.81±4.27 | 7.98±10.56 |
| 36 | 11+13-Methylnonacosane | 2937 | 6.71±1.19 | 9.36±2.47 | 8.13±2.00 | 7.53±0.90 | 8.16±2.94 | 8.91±5.54 | 9.83±10.14 |
| 37 | 7-Methylnonacosane | 2948 | 0.77±0.31 | tr | tr | tr | tr | tr | tr |
| 38 | 5-Methylnonacosane | 2957 | 1.09±0.49 | tr | tr | tr | tr | tr | tr |
| 39 | Dimethylnonacosane | 2966 | tr | 2.05±0.79 | tr | tr | tr | tr | tr |
| 40 | 3-Methylnonacosane | 2978 | tr | tr | tr | tr | tr | tr | tr |
| 46 | 11+15-Methylhentriacontane | 3131 | 4.37±1.81 | 4.42±1.86 | 4.80±0.90 | 5.65±0.73 | 9.36±3.59 | 12.38±5.99 | 10.00±10.05 |
| 47 | 3-Methylhentriacontane | 3175 | tr | tr | tr | tr | tr | tr | tr |

PCA analysis was preliminarily carried out including all three classes of hydrocarbons (*n*-alkanes, alkenes and methyl branched alkanes), but as with *C. vicina*, ageing could not be determined from the plot due to substantial scatter. A dataset containing just the methyl branched compounds provided the best PCA plot, shown in Figure 4.

As with *C. vicina*, PCA was carried out using the first 6 principal components which described 97.2% of the variation within the data sets with the first four principal components, comprising 73.3%, 8.6%, 6.6% and 4.1% respectively. PC3 and PC4 were used to plot the relevant PCA scores (Figure 4).

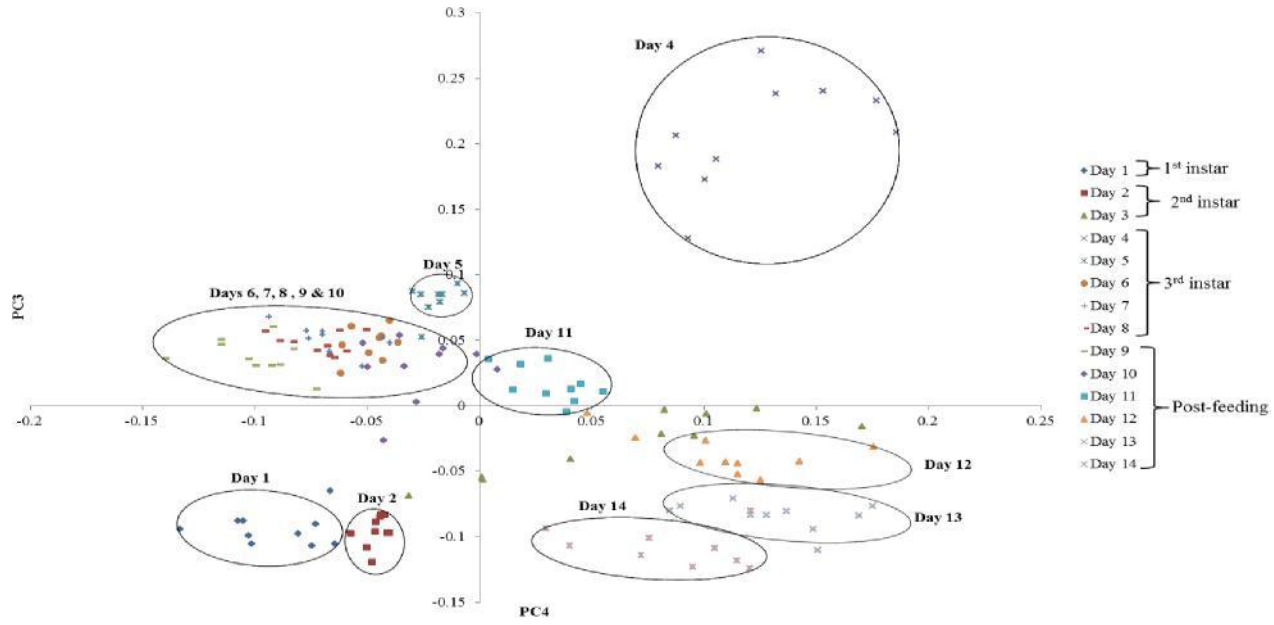



**Figure 4**: PCA plot showing PC4 against PC3 for *C. vomitoria* larvae using methyl branched alkanes only, with clustering days circled to catch the majority of that days points.

The PCA plot in Figure 4 gives significantly enhanced clustering within the PCA plot, allowing for ageing to be established to a much higher degree of accuracy, with the exception of the third instar which clusters days 6 to 10 in the same group. There are still some outliers that must not be overlooked, for example day 1. However, the post-feeding stage can be aged to the individual day, which is highly advantageous as this is usually a problematic stage to age using current ageing techniques.

The compound that exhibited the largest PCA loading and therefore was the most significant in the changes seen in the PCA plot was 11+9-Methylpentacosane.

*Neural network analysis*

Table 6 shows the generalization performance of trained SOMs for each dataset below with the average test performance and standard deviation shown in parenthesis. As the table shows, the best test performance achieved by the SOM for both *C. vicina* and *C. vomitoria* was when the average of five samples was used as the test data, where results are improved by 11 and 17% respectively. This is to be expected as a result of the variation between hydrocarbon profile samples of individual larvae and was shown to be the case when classifying *L. sericata* previously [16]. The majority of the test errors given by each SOM are likely to be a result of similarities between hydrocarbon profiles of larvae which are of a similar age as was shown by the PCA analysis in Figures 3 and 4 where certain ages were within the same clusters. For example, when testing using the average of the remaining five hydrocarbon profiles the majority of errors were misclassifications of +/- 1 day, as shown by the confusion matrices in Table 7.

**Table 6:** The overall test performance of each SOM when classifying the larvae of *C.vicina* and *C. vomitoria* hydrocarbon profiles.

| **Test approach** | **% correct (SD)** | |
|---|---|---|
| | ***C. vicina*** | ***C. vomitoria*** |
| Average of five samples | 89.09 (15.04) | 87.86 (16.26) |
| Individual samples | 78.4 (19.58) | 70.714 (20.19) |

**Table 7:** Confusion matrices showing the performance of each SOM when classifying for each fold of cross-validation as well as the overall classification performance for each day when tested using the average of the remaining five input patterns of *C.vicina* (top) and *C. vomitoria* (bottom) hydrocarbon profiles.

| SOM classification | Input pattern tested | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | D1 | D2 | D3 | D4 | D5 | D6 | D7 | D8 | D9 | D10 | D11 |
| D1 | 9 | 1 | | | | | | | | | |
| D2 | 1 | 9 | | | | | | | | | |
| D3 | | | 10 | | | | | | | | |
| D4 | | | | 9 | | | | | | | |
| D5 | | | | 1 | 10 | | | | | | |
| D6 | | | | | | 9 | | | | | |
| D7 | | | | | | 1 | 10 | | | | |
| D8 | | | | | | | | 10 | | | |
| D9 | | | | | | | | | 5 | 1 | |
| D10 | | | | | | | | | 4 | 7 | |
| D11 | | | | | | | | | 1 | 2 | 10 |
| % correct | 90 | 90 | 100 | 90 | 100 | 90 | 100 | 100 | 50 | 70 | 100 |

| SOM classification | Input pattern tested | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | D1 | D2 | D3 | D4 | D5 | D6 | D7 | D8 | D9 | D10 | D11 | D12 | D13 | D14 |
| D1 | 8 | | | | | | | | | | | | | |
| D2 | 2 | 10 | | | | | | | | | | | | |
| D3 | | | 9 | | | | | | | | | | | |
| D4 | | | | 10 | | | | | | | | | | |
| D5 | | | | | 10 | | | | | | | | | |
| D6 | | | | | | 10 | 1 | | | | | | | |
| D7 | | | | | | | 6 | 2 | | | | | | |
| D8 | | | | | | | 3 | 5 | | | | | | |
| D9 | | | | | | | | 1 | 10 | | | | | |
| D10 | | | | | | | | 2 | | 10 | | | | |
| D11 | | | | | | | | | | | 9 | 2 | | |
| D12 | | | | | | | | | | | 1 | 8 | | |
| D13 | | | 1 | | | | | | | | | | 10 | 2 |
| D14 | | | | | | | | | | | | | | 8 |
| % correct | 80 | 100 | 90 | 100 | 100 | 100 | 60 | 50 | 100 | 100 | 90 | 80 | 100 | 80 |

# Discussion

Roux and co-workers [14] examined the cuticular hydrocarbons of three forensically important blowflies (*C. vicina*, *C. vomitoria* and *P. terraenovae*). They examined the ontogenetic study of these three species from egg through to 8 day old adult flies. Similar observations were noted in comparison to the larvae results presented in this study. The chemical profiles of larvae and post-feeding larvae contain short chain hydrocarbons which evolve into long chain compounds in the pupae and adult flies. The methyl branched alkanes

were also seen to be more abundant in the immature stages of the larvae, with a substantial decrease as they became post-feeding [12].

The PCA plots (Figures 3 & 4) show the potential to age the larvae down to the day, with the exception of the first and second instar larvae (days 1 and 2) and late post-feeding larvae (days 9 to 11) for *C. vicina* and mid-aged third instar larvae (days 6-10) for *C. vomitoria*, which cluster together into one group within the plot.

The very tight clustering observed in *C. vicina* days 1 and 2 are maybe due to the eight methyl branched compounds these two days alone share. Day 3 has a few principal components within its clustering region from day 4, indicating some of the extracted larvae were developing at a slightly faster rate than others. The clusters seen in the PCA plot (Figure 3) form a systematic sequence that tracks the chemical changes. Starting from days 1 and 2 on the bottom left, the changes can be followed in a largely clockwise direction. The large jump from days 1 and 2 to day 3 corresponds to the chemical change going from the 1$^{st}$ and 2$_{nd}$ instar to the 3$^{rd}$ instar.

For *C. vomitoria*, the late post feeding stage can be aged to the day (Figure 4) with days 11 to 14 all individually clustering. This stage is often extremely difficult to age and there are currently no publications able to age this larval stage accurately. This technique therefore shows very promising results for this particular life stage of *C. vomitoria*.

As the larvae age, the higher boiling point alkanes become more abundant. The heavier long chain hydrocarbons are believed to be involved in waterproofing [35-37] which could explain why the larvae of the *Calliphora* species exhibit an increase in these compounds at a later age. When the larvae become older and gradually move into the post-feeding stage of the life cycle they move away from the source of food and seek a site for pupariation, exposing them to a drier environment. Therefore they have a greater need for extra waterproofing compared to their younger age where they are usually at least partly submerged within their food source which is warm

and moist. The display of higher boiling point *n*-alkanes mixed in with alkenes and methyl branched hydrocarbons has also been linked to flexibility of the cuticle [38]. To help flexibility in the larvae's cuticle it will need a composition of methyl branched alkanes and alkenes, which have lower melting points compared to the straight chain alkanes. Further analysis using ANNs allows for the automated classification of newly collected samples without the need for analysis of principal component plots while maintaining a high level of performance. A high performance was achieved for both *C. vicina* and *C. vomitoria* where 89 and 87% accuracy was achieved respectively when testing using an average of five hydrocarbon profiles. Further analysis of the performance of each SOM for both species using confusion matrices showed that a large proportion of test error was caused by the misclassification of larvae by +/- 1 day. This was true for all test cases when classifying *C. vicina* and for all but 3 cases when classifying *C. vomitoria*. Larger datasets for training and testing are likely to improve these results further.

## Conclusion

This study has successfully performed preliminary CHC analysis with the aid of statistical analysis to determine the age of larvae from two forensically important *Calliphora* species under controlled laboratory conditions.

Early results show great potential to utilise this technique and to develop it into a highly useful automated ageing tool using principal component analysis and artificial neural networks where test accuracy scores of 78% and 70% were obtained for *C. vicina and C. vomitoria* respectively when using individual hydrocarbon profiles. This performance was improved when testing using an average of the remaining five hydrocarbon profiles (i.e. those that were not used to train the neural network) to provide 89% and 87% for *C. vicina and C. vomitoria* respectively. The majority of errors given by the trained networks were

misclassifications of larvae by +/- 1 day. Further analysis and larger datasets are required to verify these results and assess the suitability of such algorithms for the automated classification of hydrocarbon profiles for accurate $_{PMImin}$ estimations.

Further work is needed to develop this technique. Results presented in this study were executed in a controlled laboratory environment; however experiments need to be carried out in the field to look at the effects that weathering may have on the stability of the hydrocarbons. On-going method development will test if practical implications will be an issue for hydrocarbon analysis, for example, pooled samples collected from the crime scene of young and old larvae.

This study has shown the great potential of hydrocarbon use coupled with statistical techniques for accurate and automated $PMI_{min}$ estimations. This work should now be repeated and validated in the field to test the stability of the hydrocarbons as well as the practicalities of the proposed techniques shown within this paper.

## References


[1] G. Zhu, G. Ye, C. Hu, X. Xu, and K. Li, Development changes of cuticular hydrocarbons in *Chrysomya rufifacies* larvae: potential for determining larval age, Medical and Veterinary Entomology 20 (2006) 438–44.

[2] J. Amendt, R. Krettek, and R. Zehner, Forensic entomology, Naturwissenschaften 91 (2004) 51–65.

[3] M. Benecke, A brief history of forensic entomology, Forensic Science International 120 (2001) 2–14.

[4] A. Hart, A. Whitaker, M. Hall, The Use of Forensic Entomology in Criminal Investigations: How it can be of benefit to SIOs, The Journal of Homicide and Major Incident Investigation 4 (2008) 37–48.

[5] B. Greenberg, Flies as Forensic Indicators, Journal of Medical Entomology 28 (1991) 565–577.

[6] F. Sperling, G. Anderson, and D. Hickey, A DNA-based approach to the identification of insect species used for postmortem interval estimation., Journal of Forensic Sciences 39 (1994) 418–27.

[7] C. Ames, B. Turner, and B. Daniel, Estimating the post-mortem interval (I): The use of genetic markers to aid in identification of Dipteran species and subpopulations, International Congress Series 1288 (2006) 795–797.

[8] Y. Malgorn and R. Coquoz, DNA typing for identification of some species of Calliphoridae. An interest in forensic entomology, Forensic Science International 102 (1999) 111–9.

[9] M. Harvey, I.. Dadour, and S. Gaudieri, Mitochondrial DNA cytochrome oxidase I gene: potential for distinction between immature stages of some forensically important fly species (Diptera) in Western Australia, Forensic Science International 131 (2003) 134–9.

[10] A. Tarone and D. Foran, Gene expression during blow fly development: improving the precision of age estimates in forensic entomology, Journal of Forensic Sciences 56 (2011) S112–S122.

[11] R. Zehner, J. Amendt, and P. Boehme, Gene expression analysis as a tool for age estimation of blowfly puparia, Forensic Science International: Genetics Supplement Series 2 (2009) 292–293.

[12] H. Moore, Analysis of cuticular hydrocarbons in forensically important blowflies using mass spectrometry and its application in Post Mortem Interval estimations, Ph. D Thesis, (2012), Keele University.

[13] H. Moore, C. Adam, and F. Drijfhout, Potential Use of Hydrocarbons for Aging *Lucilia sericata* Blowfly Larvae to Establish the Postmortem Interval., Journal of forensic sciences 58 (2012) 404–412.

[14] O. Roux, C. Gers, and L. Legal, Ontogenetic study of three Calliphoridae of forensic importance through cuticular hydrocarbon analysis, Medical and Veterinary Entomology 22 (2008) 309–17.

[15] J. Pechal, H. Moore, F. Drijfhout and E. Benbow, Hydrocarbon profiles throughout adult Calliphoridae aging: A promising tool for forensic entomology, Forensic Science International 245 (2014) 65-71.

[16] J. Butcher, H. Moore, C. Day, C. Adam and F. Drijfhout, Artificial Neural Network analysis of hydrocarbon profiles for the ageing of *Lucilia sericata* for Post Mortem Interval estimation, Forensic Science International 232 (2013) 25-31.

[17] W. Brown, R. Morton, J. Spradbery, Cuticular hydrocarbons of the Old World screw-worm fly, *Chrysomya bezziana* Villeneuve (Diptera: Calliphoridae). Chemical characterization and quantification by age and sex, Comp. Biochem. Phys. B 101 (4) (1992) 665–671.

[18] R. Urech, G. Brown, C. Moore, P. Green, Cuticular hydrocarbons of buffalo fly, *Haematobia exigua*, and chemotaxonomic differentiation from horn fly, *H. irritans.*, J. Chem. Ecol. 31 (2005) 2451–2461.

[19] G. Ye, K. Li, J. Zhu, G. Zhu, C. Hu, Cuticular hydrocarbon composition in pupal exuviae for taxonomic differentiation of six necrophagous flies, Journal of Medical Entomology 44 (3) (2007) 450–456.

[20] H. Moore, C. Adam, F. Drijfhout, Identifying 1st instar larvae for three forensically important blowfly species using "fingerprint" cuticular hydrocarbon analysis, Forensic Science International 240 (2014) 48-53.

[21] G. Blomquist, D. Nelson, and M. De Renobales, Chemistry, biochemistry, and physiology of insect cuticular lipids, Archives of Insect Biochemistry and Physiology 6 (1987) 227–265.

[22] J. Espelie, K., Payne, Characterization of the cuticular lipids of the larvae and adults of the pecan weevil, *Curculio caryae*, Biochem. Sys. Ecol 19 (1991) 127–132.

[23] G. Zhu, X. Xu, X. Yu, Y. Zhang, and J. Wang, Puparial case hydrocarbons of *Chrysomya megacephala* as an indicator of the postmortem interval, Forensic Science International 169 (2007) 1–5.

[25] G. Blomquist and L. Jackson, Chemistry and biochemistry of insect waxes, Progress in Lipid Research 17 (1979) 319–345.

[26] G. Blomquist, D. Nelson, and M. De Renobales, Chemistry, biochemistry, and physiology of insect cuticular lipids, Archives of Insect Biochemistry and Physiology 6 (1987) 227–265.

[27] F. Drijfhout, Cuticular Hydrocarbons: A New Tool in Forensic Entomology?, in: J. Amendt, C.P. Campobasso, M.L. Goff, and M. Grassberger, eds., Current Concepts in Forensic Entomology, Springer, 2010, pp. 179–204.

[28] J. Butcher, D. Verstraeten, B. Schrauwen, C. Day and P. Haycock, "Pruning reservoirs with random static projections", IEEE workshop on Machine Learning for Signal Processing, (2010) 250-255

[29] M. Beyeler, N. Oros, N. Dutt, and J Krichmar., 2015. A GPU-accelerated cortical neural network model for visually guided robot navigation. Neural Networks, 72, pp.75-87.

[30] J. Butcher, D. Verstraeten, B. Schrauwen., C. Day,and P. Haycock, Defect detection in reinforced concrete using Reservoir Computing and Extreme Learning Machines, Computer-Aided Civil and Infrastructure Engineering, 29 (3) (2014) 191-207.

[31] A. Bianconi, C. Von Zuben, A. Serapião, J. Govone, Artificial neural networks: A novel approach to analysing the nutritional ecology of a blowfly species, *Chrysomya megacephala*, Journal of Insect. Science 10 (58) (2010).

[32] T. Kohonen, The self-organising map, Proceedings of the IEEE 78(9) (1990) 1464–1480.

[33] C. Day, J. Austin, J. Butcher, P. Haycock, Element-specific determination of x-ray transmission signatures using neural networks, NondisT&E Int. 42 (5) (2009) 446–451.

[34] C. Adam, S. Sherratt, V. Zholobenko, Classification and individualisation of black ballpoint pen inks using principal component analysis of UV-vis absorption spectra. Forensic Sci Int 2008;174(1) (2008) 16-25.

[35] T. Tregenza, S. Buckley, V. Pritchard, and R. Butlin, Inter- and Intra-Population Effects of Sex and Age on Epicticular Composition of Meadow Grasshopper,*Chorthippus parallelus*, Journal of Chemical Ecology 26 (2000) 257–278.

[36] R. Toolson, E. Kuper-Simbron, Laboratory evolution of epicuticular hydrocarbon composition and cuticular permeability in *Drosophila pseudoobscura*: effects on sexual dimorphism and thermal- acclimation ability, Evolution 43 (1989) 468–473.

[37] J. Ferveur, Cuticular hydrocarbons: their evolution and roles in *Drosophila* pheromonal communication., Behavior genetics 35 (2005) 279–95.

[38] D. Morgan, *Biosynthesis in Insects*, The Royal Society of Chemistry, UK, 2010.