# Chapter 30
# Pair-Specific Scents in African Wild Dogs, *Lycaon pictus*, and an Example of a Potential Method to Identify Signals Within Complex Mixtures

**Neil R. Jordan, Peter J. Apps, Krystyna A. Golabek, and J. Weldon McNutt**

## 30.1 Introduction

Olfactory communication plays a major role in the lives of mammals (Ewer 1968), and scent is pivotal in modulating an extremely wide range of behaviors across an equally wide range of taxa. Chemical signals, and an animal's ability to extract this information from scent-marks, are fundamental components of parental (e.g. Poindron et al. 1988), sexual (e.g. Rasmussen et al. 1997), and territorial behavior (e.g. Müller and Manser 2007). A fundamental component of territorial scent marking depends upon discriminating differences between signals from "self" and "other" at the very least, although more complex discrimination such as neighbor–stranger (e.g. banded mongoose, *Mungos mungo*, Müller and Manser 2007) and group-specific discrimination have been described in some species (e.g. raccoon

N.R. Jordan (✉)
Botswana Predator Conservation Trust, Private Bag 13, Maun, Botswana

Centre for Ecosystem Science, School of Biological, Earth and Environmental Sciences, University of New South Wales, Sydney, NSW 2052, Australia

Taronga Conservation Society Australia, Taronga Western Plains Zoo, Wildlife Reproduction Centre, Obley Road, Dubbo, NSW 2830, Australia
e-mail: neil.jordan@unsw.edu.au

P.J. Apps
Paul G. Allen Family Foundation Laboratory for Wildlife Chemistry, Botswana Predator Conservation Trust, Private Bag 13, Maun, Botswana

K.A. Golabek
Botswana Predator Conservation Trust, Private Bag 13, Maun, Botswana

Department of Zoology, Wildlife Conservation Research Unit, University of Oxford, Oxford, UK

J.W. McNutt
Botswana Predator Conservation Trust, Private Bag 13, Maun, Botswana



B.A. Schulte et al. (eds.), *Chemical Signals in Vertebrates 13*,
DOI 10.1007/978-3-319-22026-0_30

dog, *Nyctereutes procyonoides*, Yamamoto 1984; European badger, *Meles meles*, Davies et al. 1988) suggesting that these abilities may be common among mammals. For animals to recognize or discriminate between scents there must be consistent differences between important categories of each unit (e.g. sex and pack). Although unit-specific discrimination has been described in many species (e.g. sex discrimination in domestic dogs, *Canis familiaris*, Dunbar 1977; individual discrimination in Indian mongooses, *Herpestes auropunctatus*, Gorman 1976), in mammalian semiochemistry in particular, the specific distinguishing compound or compounds typically remain unknown (Apps 2013). As mammals commonly use urine and feces in communication, distinguishing active signals from metabolic waste products constitutes a significant part of the challenge in this process.

The most commonly accepted mechanism of territorial defense using scent is by "scent matching" (Gosling 1982). Intruders match the scents of encountered individuals with scents encountered in the environment, and subsequently modulate their behavior in the presence of the identified territory owner(s) due to asymmetry in the costs/benefits of escalating a challenge. In order for such scent matching to occur, individuals must be able to discriminate between scents and identify territory residents as such. In group-living species, territoriality via scent-matching is likely to be facilitated through group-specific odors deployed as scent marks communicating territory residence.

Long-term ranging behaviors of a subpopulation of African wild dogs (*Lycaon pictus*) in northern Botswana followed in this study strongly illustrate characteristic territorial residence, including the use of exclusive areas (Pomilia et al. 2015). Furthermore, observations of daily movements by resident packs indicate that residence may be communicated using scent marks (e.g. Jordan et al. 2013). As part of ongoing research to manipulate the movement and territorial behavior of African wild dogs, we investigated whether the chemical compositions of wild dog scent-marks are pack-specific. We used gas chromatography and mass spectrometry (GC–MS), deconvolution (Chap. 28) and multivariate statistics to separate and identify any pack-specific components that may contribute to a signal of territory residence. Previous work on the scent-marking behavior of African wild dogs showed that the dominant urine overmarks (DUOs) of wild dog pairs—where one member of a pair deposits urine on top of the urine of its partner—are a likely source of semiochemical information (Jordan et al. 2013). Distinct postures were almost uniquely associated with DUO deposition and increased the likelihood that these deposits would be investigated by other dogs (Jordan et al. 2013). Among all scent types, dominant urine was of greatest interest to other wild dogs. Additionally, the dominant pair in a pack may remain stable over many years, whereas the composition of the rest of the pack changes from year to year, as subdominants disperse or die (McNutt 1996). Consequently, DUOs provide the focus for our search for a pack-specific semiochemical signal. If DUOs contain a signal advertising residence, then we would expect the DUOs from each pair/pack to be distinguishable from others (i.e. a pair-specific signature), since residents must be capable of distinguishing their own scents from those of their neighbors.

Searching through complex multicomponent mixtures for signals presents a challenge in analytical semiochemistry. Focusing our investigation on territorial

scent marking in African wild dogs in this study, we attempt to address the following questions: Are the chemical compositions of DUOs of wild dogs pair-specific, and which particular components might impart pair-specificity? In attempting to answer these questions we utilize established statistical techniques in a novel way. We used a multivariate (factorial) approach as a statistical filter to remove components that varied within pairs while retaining components that differed between pairs. This approach uses Principal Components Analysis, and particularly its utility in locating and retaining variables (chemical components) which account for the highest proportion of inter-sample variance. Our approach may be broadly applicable across signaling modalities, as finding signals amongst the noise of multicomponent mixtures is a common problem in animal communication.

## 30.2 Materials and Methods

### *30.2.1 Study Site and Population*

Data were collected between July 2011 and September 2012 from six packs of free-ranging African wild dogs in Botswana. The study area (approx. 2600 km$^2$; 19°31′S, 23°37′E; elevation approx. 950 m) is bordered by the Okavango Delta and includes the Moremi Game Reserve and surrounding Wildlife Management Areas. Further details can be found in McNutt (1996). A pack was defined as a group containing at least one adult male and female which form a potential reproductive unit (cf. Malcolm 1979; McNutt 1996). Mean adult (>1 year) pack size (at the study midpoint) was 7 ± 4.29 (mean ± SD; range; 3–14). Packs were located by radiotracking from the air and from a vehicle, with one to six individuals in each pack fitted with GPS radiocollars (<350 g) with a VHF tracking pinger (Vectronic Aerospace GmbH, Berlin, Germany; Royal Veterinary College, London, UK,) or VHF radiocollars (Sirtrack, Havelock West, New Zealand; <180 g) following procedures described elsewhere (Osofsky et al. 1996). Dominant pairs were easily identified in each pack by their mutual overmarking or "tandem marking" behavior, where one individual places its urine directly on top of (sensu Johnston et al. 1994) the urine of its opposite-sex partner (Jordan et al. 2014). This behavior is associated with sniffing and the adoption of specific leg postures, and was not observed in subdominant wild dogs (Jordan et al. 2013). During this study, breeding attempts were only observed between dominant pairs.

### *30.2.2 Sample Collection*

Packs were observed from a vehicle while resting (at distances of 3–40 m) and traveling (20–200 m). We remained at the DUO deposition site until the dogs had moved away sufficiently to allow the sample to be collected on foot beside the vehicle.

Samples used in this study were collected 9.04 ± 7.33 (Mean ± SD; range 2–26) min post-deposition using clean metal spoons to scoop patches of urine-soaked soil into cleaned glass jars with lids lined with aluminum foil. Jars were double rinsed with boiling distilled water. Caps (80 °C), jars, and liners (220 °C) were baked for at least 24 h at in a forced convection oven, assembled while still warm, and were not opened until the sample was put in them. Samples from the same pair were never collected from the same location, and all samples were frozen (ca. −5 °C) on the day of collection and until analysis. Samples were frozen for 355 ± 17 (Mean ± SD) days (range: 168–606) before analysis.

### *30.2.3 Chemical Analysis*

Samples of urine on soil were spread in stainless steel dishes in a fume cupboard and samples were stirred intermittently to get even drying. Room temperature is air-conditioned to 22–25 °C. Once the sample was dry enough to be free-flowing (which usually took between 2 h and overnight) it was sieved through a stainless steel mesh with holes of about 1 mm to remove fragments of vegetation, and then a weighed subsample was taken from it. The dried, sieved remainder was returned to the sample's original jar, and subsamples were weighed accurately (0.1 mg resolution) into glass Pasteur pipettes with a small pledget of silanized glass wool (Sigma-Aldrich) in the stem. Pipettes were prepared by inserting the glass wool, rinsing with 2 ml of methanol and then baking overnight (at least) at 220 °C. Approximately 2.5 g of sample filled the body of the pipette to about 10 mm below the open end. The pipettes were supported vertically with their tips in 2 ml autosampler vials that had been rinsed with methanol and dried at room temperature. Methanol (Signal-Aldrich Chromasolve Gradient grade) was percolated though the sample under gravity and the first 1 ml (to a mark on the vials) was collected. GC–MS analyses of 2 μL of injected extract were carried out on a Varian 450 GC gas chromatograph with a Varian 1079 programmable inlet, interfaced to a Varian 320 MS single quadrupole mass spectrometer. The inlet was fitted with a fritted liner (LJ-18-24-1) and the column was a 30 m × 0.32 mm × 0.5 μm polyethylene glycol (PEG) (Restek Rtx-Wax #12349), held at 50 °C for 3 min and programmed at 5 °C/min–240 °C. The carrier gas was helium with a constant volume flow of 2 mL/min. Electron impact mass spectra were acquired with electron energy of 70 eV, and compounds were tentatively identified by searches of the NIST 05 mass spectral library. For further details see Apps et al. (2012). Varian Bruker MSWS mass spectral data files from 59 samples (selected from a larger set to provide inter-pack comparisons) were imported to the AnalyzerPro deconvolution software package (SpectralWorks) and processed using the Matrix Analyzer routine with the following settings up to 40 min retention: reject masses 33–45, minimum masses 2, area threshold 800,000, height threshold 1, signal:noise 1, width threshold 0.04 min, resolution high, scan windows 1, smoothing 5, forward fit 550, reverse fit 550, retention window 0.25 min. After 40 min retention m/z 129 was added to the excluded ions, minimum masses

was increased to four, area threshold increased to 1,000,000, smoothing increased to 15, and fits to 650. A known series of methyl-imidazolidinedione isomers, which have a single strong ion at m/z 114, and groups of peaks with shifting retentions and similar spectra were deconvoluted using ion and retention settings specific to the peaks (see also Apps, this volume).

### *30.2.4 Statistical Analysis*

Data were exported to Microsoft Excel for manual cleanup, and then imported to R (R Core Team 2013) for all statistical processing. Statistical analyses were applied to 24 DUOs (four from each of six packs). Raw peak areas were converted to percentages of the total peak area on the chromatogram. As the dataset contained many more components than samples, we first reduced the number of variables using a Principal Component Analysis (PCA). PCA cannot deal with 0 values, and almost all components had peak areas below the minimum threshold in at least one sample, so we assigned an arbitrarily low percentage abundance (0.00001) to all zero values. Following this, all abundance data were log-transformed (using a natural log function) as PCA would otherwise over-inflate the importance of variables (components) with higher values (abundances). As the resultant (post-transformation) variances entered into the PCA reflect proportional changes and not absolute changes, this removes any bias towards components with higher abundances, and gives equal weight to components at low measured abundances; an important consideration in chemical signaling.

#### 30.2.4.1 Are African Wild Dog Duos Pair-Specific?

To determine whether African wild dog DUOs had pair-specific qualities, we conducted Discriminant Function Analyses (DFA) on the dataset of 24 samples (4/pack). Post-hoc "bootstrapping" analyses were conducted to determine the probability that a cross-validated correct assignment value from the DFA was achieved by chance, and followed the methods of Müller and Manser (2008).

#### 30.2.4.2 Which Components Code for Pair?

To determine which components may code for pair, we processed the data to remove intra-pair variation and retain inter-pair variation using PCA; particularly its utility in identifying and retaining maximum variation between samples. PCA is a dimension-reduction tool that can be used to reduce a large set of variables to a small set that still contains most of the variation of the original set. If several samples from each pack were subjected to PCA, then the analysis would retain the maximum variation across the whole data set, therefore resulting in the retention of

both intra- and inter-pair variation. Instead, by conducting PCA with one sample from each pair (six samples total) we were able to maximize the retention of inter-pair variation while excluding intra-pack variation from the resulting principal components (PCs). We conducted PCA using a covariance matrix on 20 different subsets of data (each subset contained a randomly selected single sample from each of the six packs).

From these outputs we identified, by ranking and examining the derived loadings (percentage variance in that PC accounted for by the factor/chemical component, averaged for each component in each PC over the 20 trials), which individual chemical components were consistently associated with this (inter-pair) variance. As the loadings describe the contribution of each of the individual chemical components to each PC, we were then able to estimate the importance of each of the chemical components to inter-pair variation. Loadings were ranked by absolute value (i.e. loadings with high negative values were also included). We identified the component with the highest loading for PC1, and then calculated the proportion of that loading value that the loading for each other component represented (again using absolute values). We identified and extracted any components whose loadings represented >85 % of the loading value of the component with the highest loading. This resulted in six components being extracted for PC1. The initial 85 % cut-off was chosen as a starting point because, with our particular dataset, it would locate a reasonable number of components that could be tested in combination in the field given the constraints of our study system. The relevance of this cut-off was checked by comparing the success of post-hoc DFA classification analyses for combinations which either included additional components, or excluding each of the original components in turn. In this case we could not improve upon the initial cut-off selected.

Since PC1 represented only 31.12 % of the estimated inter-pair variation, we also extracted the highest loading components from PC2 to PC5, as this explained >99 % of the overall variation. The number of components we extracted from these PCs was dependent upon the percentage of variation that each PC accounted for in the inter-pair variation (see Table 30.1). As six components were extracted from PC1 and PC1 represented 31.12 % of the variation, the top four components were extracted from PC2, as this PC represented 22 % of the variation (and so contributed 22 % of the extracted components). The number of components extracted from each PC was proportional to the variation that each PC explained.

Eleven of the 19 components had been previously identified (Apps et al. 2012, shown bold in Table 30.2), and nine of these were considered unlikely to be signals

**Table 30.1** Percentage variance explained by five principal components across 20 trials

| Principal component | PC1 | PC2 | PC3 | PC4 | PC5 | Total |
|---|---|---|---|---|---|---|
| Mean % variance explained | 31.12 | 22.44 | 19.03 | 15.78 | 11.63 | |
| Max. % variance explained | 39.76 | 25.49 | 20.94 | 17.84 | 15.45 | |
| Min. % variation explained | 24.78 | 20.31 | 16.22 | 10.49 | 8.34 | |
| *N* (highest loading) components extracted | 6 | 4 | 4 | 3 | 2 | 19 |

**Table 30.2** Nineteen components potentially contributing to a signal of pair-specificity

| Ret. order | Identity (confirmed in bold) | % detected ($n=24$) | Likely signal? | Comment |
|---|---|---|---|---|
| 1 | **2-methylpropanoic acid** | 62.5 | No | Widespread in mammals |
| 2 | **Quinazoline** | 50 | Yes | Wild dog specific |
| 3 | Unknown | 58.33 | Yes | |
| 4 | ***N*,*N*-dimethylurea** | 54.17 | No | Widespread in mammals |
| 5 | **2-piperidinone** | 75 | No | Ubiquitous urine metabolite |
| 6 | **Glycerine** | 66.67 | No | Ubiquitous |
| 7 | m/z 85/86 unknown compound 1 | 41.67 | Yes | |
| 8 | m/z 85/86 unknown compound 2 | 70.83 | Yes | |
| 9 | **Methyl tridecanoate** | 20.83 | No | Probably from methanol + tridecanoic acid |
| 10 | Unknown | 50 | Unknown | |
| 11 | **Tetradecanoic acid** | 45.83 | No | Ubiquitous |
| 12 | **1-methyl-2,4-imidazolidinedione** | 79.17 | Unknown | Carnivore-specific. 3-methyl- and 5-methyl-isomers also present |
| 13 | Unknown | 4.17 | Unknown | |
| 14 | Pentadecen-1-ol | 29.17 | Unknown | Isomer unknown |
| 15 | **Nonacosane** | 58.3 | Unknown | |
| 16 | **Hexadecanoic acid** | 75 | No | Ubiquitous |
| 17 | **Hexadecenoic acid** | 29.17 | No | Ubiquitous-isomer unknown |
| 18 | 2-aminocarbonyl-1-methylimidazole | 37.5 | Unknown | |
| 19 | **Octadecanoic acid** | 54.17 | No | Ubiquitous |

of pair because they are very widespread in mammal odors (Burger 2005). To test whether the remaining 10 extracted components allowed classification of pairs in the absence of any other chemical information, log-transformed percentage abundance data were extracted for these components. Because multiple components were extracted from each PC, the abundance of each extracted component was likely to be correlated with the abundance of at least one other extracted component. As DFA requires the input variables to be uncorrelated, we ran the dataset ($n=24$, four samples from six pairs) through PCA before assessing the classification of samples using DFA. Using PCA reduced the dimensions of the dataset to produce uncorrelated/orthogonal input variables (PCs) thus removing the common problem of multicollinearity in tests of this nature (Field 2005). Bootstrapping was used to assess whether statistical classification of samples to pair was significantly better than that expected by random assignment.

It is important to point out that our method allows for the possibility of both single components and ratios to be the signal. It also allows for the possibility of binary coding; i.e. that the presence or absence of each component (or combination of components) is what differs among pairs. We explore this possibility further by

conducting an additional DFA on PCs derived from presence absence data. In each of the 24 samples, any of the 10 components that were detected at any abundance were assigned a score of "1," and undetected components were assigned a score of "0."

## 30.3 Results

Deconvolution located 990 components in the 24 samples. Only one component, benzoic acid (Apps et al. 2012) was present in all 24 samples. 27 (2.73 %) components were detected in 75 % or more samples, 66 (6.67 %) components were detected in 50 % or more samples and 135 (13.64 %) components were detected in at least 25 % of samples. The majority of components (703; 71.01 %) were each detected from only one pair, with a large majority of these (669/703, 95.16 %) being found in only one sample. Of the remaining 287 components that were detected in more than one sample, 34 were located exclusively in samples from a single pair, but not in all the samples from a pack. Twenty-six of these components (76.47 %) were found in only two samples from the same pair, and the remaining eight (23.53 %) were found in three of four samples from the same pair. Only one pair had components in ¾ of the samples, and no component found exclusively in one pair was found in all samples from that pair (None of these components were extracted on the basis of their loadings below).

### *30.3.1 Are African Wild Dog DUOs Pair-Specific?*

Fifteen PCs explaining 85.6 % of the measured variance contained in 990 chemical components in 24 samples (four samples/pack from six packs) were input into a DFA which correctly classified 21/24 (87.5 %) of DUOs to the pair that produced them (Table 30.3). This is significantly better than the 16.67 % correct classification

**Table 30.3** A comparison of correct classification of 24 African wild dog dominant urine overmarks (DUOs) to six pairs with that achieved by Discriminant Function Analysis using five Principal Components derived from: (1) log-transformed percentage abundance data on all 990 located chemical components; (2) log-transformed percentage abundance data on 10 extracted chemical components; and (3) presence/absence data on 10 extracted chemical components

| % correct classification results | | |
|---|---|---|
| Log-transformed % abundances (990 components) | Log-transformed % abundances (10 components) | Presence/absence (10 components) |
| 87.5 | 79.17 | 79.17 |
| $p < 0.0001$ | $p < 0.0001$ | $p < 0.0001$ |

*p* values are derived from a bootstrapping analysis and assess whether the classification obtained is relative to that expected at random (16.67 % in all cases)

expected by chance (bootstrapping; $p < 0.0001$), and is evidence that the chemical composition of samples is pair-specific.

### *30.3.2 Which Components Code for Pair?*

With intra-pair variability excluded by analyzing one sample from each pack, in 20 trials, 5 PCs were sufficient to explain 100 % of measured inter-pair variation (Table 30.1). Nineteen of 990 components possibly contributed to a signal of pair specificity (Table 30.2). Of these, nine were unlikely to be signals, due either to their ubiquity (#1, #4, #6, #11, #16, #17, #19) or their known derivation in excretory pathways (#5, #9), and were excluded from further analyses. Of the 10 remaining possible signaling components, quinazoline (#2) has, to our knowledge, so far only been identified in the scent-marks of African wild dogs (Apps et al. 2012) and 1-methyl-2,4-imidazolidinedione (#12) appears to be carnivore-specific and has also been found in the urine of lions, leopards, and spotted hyenas, but not in that of elephants, buffalo, impala, or zebra (Apps, unpublished). Pentadecen-1-ol and 2-aminocarbonyl-1-methylimidazole (#18) were tentatively identified from mass spectra, but the other six components of interest (#3, #7, #8, #10, #13, and #15) are unknowns. Some components are at picogram levels and their identification will present formidable problems.

Seven Principal Components (PCs) encompassing 96.02 % of the variance contained within log-transformed percentage abundance data from 10 components only were sufficient to correctly classify 19/24 (79.1 %) of DUOs to the pair that produced them (Table 30.3). This classification is significantly better than the 16.67 % expected by chance alone (bootstrapping; $p < 0.0001$).

Eight PCs containing 98.39 % of the variation contained within a binary coded (presence/absence) data frame were sufficient to correctly classify 19/24 (79.1 %) of DUOs to the pair that produced them, which is the same percentage correct classification as for analyses including percentage abundance data (Table 30.3). Again this classification is significantly better than the 16.67 % expected by chance alone (bootstrapping; $p < 0.0001$), and may suggest that presence/absence, rather than absolute or relative abundance, may be a sufficient signal.

## 30.4 Discussion

We assessed the chemical composition of scent marks from African wild dogs using multivariate statistics and showed that DUOs have a high degree of pair-specificity in their chemical compositions. From 990 components, we statistically located 10 candidate components based on inter-pack variation that could contribute to a signal of pair-specificity. Information on the presence or absence of these 10 components in combination was sufficient to statistically classify a significantly high proportion

of samples to the correct dominant pair/pack. Overall our results show that African wild dog DUOs contain statistically identifiable characteristics to enable pair-specific discrimination, and that binary-based multicomponent discrimination could be sufficient for pack-discrimination by scent in African wild dogs.

A degree of pack-specificity has previously been shown in the scents of some other species. For example, analytical studies based on gas chromatography showed that sub-caudal secretions from European badgers have inter-group differences in their chemical profile (Gorman et al. 1984), and that the castoreum secretions of Eurasian beavers (Schulte 1998) and North American beavers (Sun and Müller-Schwarze 1998) are family-specific. In general, we would expect territoriality and/or ownership of resources to be advertised by unit-specific signals; in solitary territorial species an individually specific signal may be expected (e.g. Indian mongoose, Gorman 1976), whereas in group-living species territory ownership should be advertised with a group-specific signal (e.g. European badger, Gorman et al. 1984). A high degree of pair-specificity in the DUOs of African wild dogs may be evidence that DUOs play an important role in communicating residence and facilitating territoriality in this species. In common with many cooperative breeders, the dominant breeding pair has most to gain by repelling intruders from their range (e.g. Ethiopian wolves, Sillero-Zubiri and Macdonald 1998), and indeed subdominants may benefit from occasional incursions by intruders, because these can provide dispersal and breeding opportunities (e.g. meerkats, Jordan 2007). The dominant pair in a pack also remains stable over many years, whereas the composition of the rest of the pack changes from year to year as subdominants disperse or die (McNutt 1996). DUOs therefore represent the most likely source of a territorial residence signal, and that we find a specific signature in the combined urine-based scent marks of dominant African wild dog pairs is perhaps unsurprising. However, it should be noted that we have not yet determined whether there are group-wide consistencies in scent, as only DUOs have been compared in this way.

We have shown that 10 components could contribute to a pair-specific signal on the basis of some combination or combinations of their presence and absence. This suggests that African wild dogs may need only to detect the presence of these chemicals (in combination) in order to discriminate between the scents of different packs, which is a similar potential mechanism as shown in the ant (*Formica fusa*), where nine different positional isomers encode the colony signal (Martin et al. 2008). However it is important to keep in mind that, although wild dog scents have pair-specific compositions, this obviously does not demonstrate that wild dogs can discriminate between the scents of different packs, or that if they can, they use that information. Further field work is required to test whether pack-discrimination by scent occurs in African wild dogs, and scent presentation experiments can then test whether we have located the signaling components.

Whether African wild dogs use the pair-specific components in encountered scents to detect and identify intruders is not yet known, and it is theoretically possible that retaining a cognitive map of where (and perhaps when) you scent marked, and comparing this to scents that you encounter could facilitate the detection of intruders on a territory (cf. Bekoff 2001) without the need for a pack/pair-specific

signal. Indeed some species have been suggested to retain the locations of several 100 features of interest throughout their range (e.g. refuge holes in meerkats, Manser and Bell 2004). However, although some components appear to trigger spatial and scent fingerprint memory (e.g. darcin in mouse scent-marks, Roberts et al. 2012), it seems that such "cognitive recall" is an unlikely mechanism of intruder detection, particularly in species with large territories. Given the prevalence of unit-specific signals (e.g. group-specificity in European badgers scents, Gorman et al. 1984) even across modalities (e.g. group-specificity in green woodhoopoe calls, Radford 2005), and the discrimination of these (e.g. partner-specific scent-discrimination in Antarctic prions, *Pachiptila desolata*, Bonadonna and Nevitt 2004; group discrimination in green woodhoopoe, *Phoeniculus purpureus*, Radford 2005), it is expected that detecting differences between self (or group companion) and other scents (e.g. European badger, Palphramand and White 2007), or the scents of different groups (e.g. colony scent-discrimination in big brown bat, *Eptesicus fuscus*, Bloss et al. 2002), is a more likely mechanism for territorial advertisement.

Although we do not test whether pair-specificity results in pair discrimination, the evidence presented here suggests that a multicomponent signal could code for pair in wild dogs. First there were 10 components whose combined abundance or presence/absence reliably discriminated between pairs, and data from each component suggested that any single component alone would be insufficient to distinguish between pairs. Additionally, although we found multiple chemical components that were specific to a pair, most of these were present in only one, two, or three samples (of four) from that pair. Thus single compound signaling of pair did not occur in our dataset, but multicomponent signaling—where the presence of one component and/or another or more in the scent may be used to discriminate between packs—may have been possible.

Although multivariate statistical approaches have often been employed to sift through complex chemical mixtures in search of category-specific information (e.g. Jordan et al. 2010; Safi and Kerth 2003), the validity of these approaches and therefore the results are rarely questioned, and opinions differ as to whether they are fit for the purpose. Martin and Drijfhout (2009) formerly questioned the validity of a multivariate approach, particularly with respect to the analysis of cuticular hydrocarbon profiles of hymenopteran insects. As our specific research problem presents a similar challenge, it is worth considering their concerns, three of which are particularly relevant in this context: (1) compounds have a high chance of being correlated (and should not therefore not be subjected to classification analyses that require independent input variables); (2) minor compounds would have a disproportionately large effect on the analysis; (3) PCA retains all of the variation in a system, not just the variation in the system that the researchers are interested in. In this paper we have attempted to deal with all of these potential pitfalls by specific data processing.

First, while it is true that some chemical compounds within a scent sample have a high chance of being correlated and therefore not being independent of one another (e.g. Martin and Drijfhout 2009), this does not preclude appropriately processed data being used in multivariate statistical analyses for classification (in this case

discriminant function analysis). As potential correlation or multicollinearity problems may lead to erroneous results if multivariate methods are used (Field 2005), we ran potentially correlated data (the abundance of located chemical components) through a PCA which groups correlated variables into PCs which are orthogonal to each other. We then used the resulting principal components for classification. This ensured that our input variables (now principal components) for DFA were independent, but retained the majority of the variation found in the original dataset of potentially correlated variables. Using a PCA in this way is common in studies of scent (e.g. Jordan et al. 2010; Safi and Kerth 2003) and sound (e.g. Golabek and Radford 2013; Radford 2005), where multiple potentially correlated variables are measured, although authors rarely state that producing independent data was a reason for doing so.

Second, the suggestion that minor compounds have a disproportionately large effect on the analysis is important, particularly given that some components we detected were in abundances at picogram levels. This issue derives from the fact that small fluctuations in the abundance of components present at low abundances would result in relatively large changes in their percentage abundance compared to similar raw changes in components present at higher abundances. In such cases, variance—which PCA uses to assign importance—will be larger for variable components at lower abundances and their importance may then be overinflated. With the methods and specific data-processing we have employed in this paper however, this is only a problem if proportional changes are related to abundance (i.e. if low abundance components are more variable in terms of their proportional shifts between samples), which we have no a priori reason to suspect. It is also important to realize that, for example, while components at relatively low and relatively high abundance respectively and displaying the same proportional changes across two samples (e.g. 1, 10, 100 in one sample and 2, 20, 200 in another) will produce different variances, therefore overinflating the importance of the more abundant component in the sample, log-transforming these data (as we did) ensures that the variance is equal for the two series and therefore reflects proportional changes regardless of relative abundances. As the resultant (post-transformation) variances entered into the PCA reflect proportional changes and not absolute changes, this removes any bias towards components with higher abundances, and gives equal weight to components irrespective of their measured abundances. This is an important consideration in chemical signaling generally, and such log transformations are commonplace in problems of this nature (e.g. Jordan et al. 2010; Safi and Kerth 2003). This is primarily because it is rarely possible to use the absolute abundance of a peak, as the total amount of eluted components may differ between samples based on sampling method (e.g. the ratio of sand to urine) and sample dilution introduced both naturally (related to individual hydration levels) and analytically. However, while raw abundances will differ in different dilutions of the same mixture, the ratios of components within them will be constant.

Finally, Martin and Drijfhout (2009) suggest that a multivariate approach may only be valid in the search for potential signaling components if all other variation can be removed. This is a problem since PCA in particular attempts to maximize the

retention of ALL variation between samples, not only the variation (in this case *between* pairs) that the researcher is interested in retaining. For example, even where individually specific signals occur (e.g. Jordan et al. 2011), scent samples from the same individuals are highly variable due to intrinsic and extrinsic factors. In attempting to locate components potential signaling pair identity, including multiple samples from the same pair in a PCA is unhelpful since much within pair variation ("noise" in this context) will also be retained. The key is to attempt to remove as much variation that is unrelated to the factor (pair) of interest and retain the variation describing the factor. In this chapter we have attempted to remove, or at least reduce, the variation that is not attributable to inter-pair differences as much as possible while retaining that variation describing inter-pair differences. We have done so by first running a PCA using only one sample from each pair, repeating this 20 times using different combinations of samples each time, and then using average loadings to locate components of interest (i.e. those describing inter-sample, and therefore inter-pair, variation). This, we believe, is a novel, appropriate and effective method of removing variation in factors other than that of interest for analyses. It may also be important to note that, for confidence in locating the correct components, multiple trials are required using different exemplars from members of each category of interest.

Despite addressing all of the above issues, and successfully reducing the number of components (variables) by nearly two orders of magnitude, we must acknowledge that our final dataset was imperfect for the classification analyses (DFA) we used to evaluate the success of our extraction method. As Mitteroecker and Bookstein (2011) point out, if the sample:variable ratio is less than 5:1, grouping can occur purely as a statistical artifact, and this ratio was 2.4:1 in this study. It is important to stress however that this issue only affects the confirmatory classification analyses at the end of the study, which we have merely used in an attempt to evaluate the efficacy of the described method in locating the components of interest. Further confirmation is therefore required to fully demonstrate the utility of this method in identifying the components contributing to a pair-specific signal in this study, and will take the form of field discrimination experiments ("bioassay"). However, identifying (not just locating) the components of interest is required before the optimal bioassays can be completed, and remains a priority work in progress. It is also important to stress that while in our particular case study we must interpret the final DFA classification results with caution, this issue does not undermine the method described to locate potential signaling components. Sample size is, as always, an important consideration for other studies wishing to use the method. We therefore contend then that the method of deriving these results is sound, and may fruitfully be employed generally where datasets fulfil the above criteria.

Identifying potential signals within complex chemical mixtures is invariably a multivariate problem, and as such it seems appropriate to adopt a multivariate approach to tackling it. While a multivariate *statistical* approach is most commonly used, it is also theoretically possible to take a multivariate *experimental* approach in the field. However, when working with most mammalian secretions, this is clearly impractical as a first step. In this study we located 990 chemical components in our

sample, which would be logistically impossible to test individually, let alone in combinations. A multivariate statistical approach is therefore necessary as a first step, but data input and inference must be undertaken with care to ensure the validity of the results. Having located a manageable number of components (10) using multivariate statistics, we are now in a position to test the relevance of these components (if any) in signaling pair-identity in African wild dogs. Following identification of all of the located components of interest, a carefully designed bioassay-based approach will be the acid test of our method for identifying potential signals in complex mixtures.

In summary, we have used multivariate statistics—and PCA in particular—to filter through data from African wild dog urinary scent samples containing almost 1000 components and reduce the number of candidate signal components by two orders of magnitude. Following this data reduction, experimental tests to determine whether the pair-specific chemical differences identified in these scents are used by the animals themselves in pair discrimination and territory defense are now feasible in the field. Identifying components signaling territory ownership is a critical stage in our research towards developing a synthetic scent-mark which mimics the natural territorial signals and responses to them by African wild dogs. Identifying the specific communication components (signals) is a common problem in the study of animal communication and chemical ecology in particular, and it is suggested that ours or a similar approach may be successfully employed as a first filter in attempts to locate components with communicative value from complex calls or chemical mixtures in particular.

**Acknowledgments** We thank the Paul G. Allen Family Foundation for funding this work. SpectralWorks (www.spectralworks.com/analyzerpro.html) kindly donated the AnalyzerPro software. All work was carried out under a permit from the Botswana Ministry of Environment Wildlife and Tourism, Department of Wildlife and National Parks. Reena H. Walker kindly assisted with data extraction. Briana Abrahms, Geoffrey Gilfillan and Jessica Vitale assisted in collecting the samples used in this study.

## References

Apps PJ (2013) Are mammal olfactory signals hiding right under our noses? Naturwissenschaften 100:487–506

Apps P, Mmualefe L, McNutt JW (2012) Identification of volatiles from the secretions and excretions of African wild dogs (*Lycaon pictus*). J Chem Ecol 38:1450–1461

Bekoff M (2001) Observations of scent-marking and discriminating self from others by a domestic dog (*Canis familiaris*): tales of displaced yellow snow. Behav Processes 55:75–79

Bloss J, Acree TE, Bloss JM, Hood WR, Kunz TH (2002) Potential use of chemical cues for colony-mate recognition in the big brown bat, *Eptesicus fuscus*. J Chem Ecol 28:819–834

Bonadonna F, Nevitt GA (2004) Partner-specific odor recognition in an Antarctic seabird. Science 306:835

Burger BV (2005) Mammalian semiochemicals. In: Schulz S (ed) The chemistry of pheromones and other semiochemicals II (Topics in current chemistry 240). Springer, Heidelberg, pp 231–278

Davies JM, Lachno DR, Roper TJ (1988) The anal gland secretion of the European badger (*Meles meles*) and its role in social communication. J Zool 216:455–463
Dunbar IF (1977) Olfactory preferences in dogs: the response of male and female beagles to conspecific odors. Behav Biol 20:471–481
Ewer RF (1968) Ethology of mammals. Logos Press, London
Field A (2005) Discovering statistics using SPSS, 2nd edn. Sage, London
Golabek KA, Radford AN (2013) Chorus-call classification in the southern pied babbler: multiple call types given in overlapping contexts. Behaviour 150:691–712
Gorman ML (1976) A mechanism for individual recognition by odour in *Herpestes auropunctatus* (Carnivora: Viverridae). Anim Behav 24:141–146
Gorman ML, Kruuk H, Leitch A (1984) Social functions of the sub-caudal scent gland secretion of the European badger *Meles meles* (Carnivora: Mustelidae). J Zool 204:549–559
Gosling LM (1982) A reassessment of the function of scent marking in territories. Z Tierpsychol 60:89–118
Johnston RE, Chaing G, Tung C (1994) The information in scent over-marks of golden hamsters. Anim Behav 48:323–330
Jordan NR (2007) Scent marking investment is determined by sex and breeding status in meerkats. Anim Behav 74:531–540
Jordan NR, Mwanguhya F, Kyabulima S, Rüedi P, Cant MA (2010) Scent marking within and between groups of wild banded mongooses. J Zool 280:72–83
Jordan NR, Mwanguhya F, Furrer RD, Kyabulima S, Rüedi P, Cant MA (2011) Scent marking in wild banded mongooses: 2. Intrasexual overmarking and competition between males. Anim Behav 81:43–50
Jordan NR, Golabek KA, Apps PJ, Gilfillan GD, McNutt JW (2013) Scent-mark identification and scent-marking behavior in African wild dogs (*Lycaon pictus*). Ethology 119:1–9
Jordan NR, Apps PJ, Golabek KA, McNutt JW (2014) Top marks for top dogs: tandem marking and pair-bond advertisement in African wild dogs (*Lycaon pictus*). Anim Behav 88:211–217
Malcolm JR (1979) Social organization and communal rearing in African wild dogs. Dissertation, Harvard University, Cambridge
Manser MB, Bell MB (2004) Spatial representation of shelter locations in meerkats, *Suricata suricatta*. Anim Behav 68:151–157
Martin SJ, Drijfhout FP (2009) How reliable is the analysis of complex cuticular hydrocarbon profiles by multivariate statistical methods? J Chem Ecol 35:375–382
Martin SJ, Helanterä H, Drijfhout FP (2008) Colony-specific hydrocarbons identify nest mates in two species of Formica ant. J Chem Ecol 34:1072–1080
McNutt JW (1996) Sex-biased dispersal in African wild dogs, *Lycaon pictus*. Anim Behav 52:1067–1077
Mitteroecker P, Bookstein F (2011) Linear discrimination, ordination, and the visualization of selection gradients in modern morphometrics. Evol Biol 38:100–114
Müller CA, Manser MB (2007) 'Nasty neighbours' rather than 'dear enemies' in a social carnivore. Proc R Soc B 274:959–965
Müller CA, Manser MB (2008) The information banded mongooses extract from heterospecific alarms. Anim Behav 75:897–904
Osofsky SA, McNutt JW, Hirsch KJ (1996) Immobilization of free-ranging African wild dogs (*Lycaon pictus*) using a ketamine/xylazine/atropine combination. J Zoo Wildl Med 27: 528–532
Palphramand KL, White PCL (2007) Badgers, *Meles meles*, discriminate between neighbor, alien and self scent. Anim Behav 74:429–436
Poindron P, Levy F, Krehbiel D (1988) Genital, olfactory, and endocrine interactions in the development of maternal behavior in the parturient ewe. Psychoneuroendocrinology 13:99–125
Pomilia MA, McNutt JW, Jordan NR (2015) Ecological predictors of African wild dog ranging patterns in northern Botswana. J Mamm DOI:10.1093/jmammal/gyv130.
R Core Team (2013) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna. http://www.R-project.org

Radford AN (2005) Group-specific vocal signatures and neighbor-stranger discrimination in the cooperatively breeding green woodhoopoe. Anim Behav 70:1227–1234

Rasmussen LEL, Lee TD, Zhang A, Roelofs WL, Daves GD (1997) Purification, identification, concentration and bioactivity of (*Z*)-7-dodecen-1-yl acetate: sex pheromone of the female Asian elephant, *Elephas maximus*. Chem Senses 22:417–437

Roberts SC, Davidson AJ, McLean L, Beynon RJ, Hurst JL (2012) Pheromonal induction of spatial learning in mice. Science 338:1462–1465

Safi K, Kerth G (2003) Secretions of the interaural gland contain information about individuality and colony membership in Bechstein's bat. Anim Behav 65:363–369

Schulte BA (1998) Scent marking and responses to male castor fluid by beavers. J Mamm 79:191–203

Sillero-Zubiri C, Macdonald DW (1998) Scent-marking and territorial behavior of Ethiopian wolves *Canis simensis*. J Zool 245:351–361

Sun L, Müller-Schwarze D (1998) Anal gland secretion codes for family membership in the beaver. Behav Ecol Sociobiol 44:199–208

Yamamoto I (1984) Latrine utilization and feces recognition in the raccoon dog, *Nyctereutes procyonoides*. J Ethol 2:47–54