

## Spatial Chemometric Analyses of Essential Oil Variability in *Eugenia dysenterica*

*Eliane C. Vilela,<sup>a</sup> Alessandra R. Duarte,<sup>a</sup> Ronaldo V. Naves,<sup>b</sup> Suzana C. Santos,<sup>a</sup>  
José C. Seraphin<sup>c</sup> and Pedro H. Ferri<sup>\*a</sup>*

<sup>a</sup>Laboratório de Bioatividade Molecular, Instituto de Química,

<sup>b</sup>Setor de Horticultura, Escola de Agronomia e de Engenharia de Alimentos and

<sup>c</sup>Núcleo de Estatística Aplicada, Instituto de Matemática e Estatística,

Universidade Federal de Goiás, CP 131, 74001-970 Goiânia-GO, Brazil

Variações nos óleos essenciais foram utilizadas para o estudo da estrutura química espacial de oito populações de *E. dysenterica* do Cerrado central brasileiro. Previamente, o autocorrelograma multivariado de Mantel e o particionamento da variação da matriz química, utilizando os conjuntos de dados espaciais e ambientais como preditores, sugeriram uma significativa variação espacial nos óleos essenciais. No presente estudo, os métodos quimiométricos espaciais usando variogramas e mapas de probabilidade permitiram a detecção e a caracterização da estrutura químico-espacial entre populações, bem como dos fatores ambientais responsáveis por ela. Todas essas estratégias indicaram que as populações diferem quimicamente a distâncias geográficas superiores a 120 km, um indicador da distância mínima entre amostras necessária para a conservação da diversidade genética das populações. Embora sendo raramente usadas com metabólitos secundários, essas metodologias possuem uma grande aplicação em conservação de espécies e podem permitir uma efetiva integração de perspectivas genética, química e ecológica.

Chemovariations in essential oils were used for studying the spatial chemical structure of eight *E. dysenterica* populations in Central Brazilian Cerrado. Previously, multivariate Mantel autocorrelogram and chemical matrix variation partitioning, using the spatial and environmental data sets as predictors, have suggested a highly significant spatial variation in essential oils. In the present study, spatial chemometric methods using variograms and probability maps detected and characterized the spatial chemical structure among populations, as well as the environmental factors responsible for them. All these strategies indicated that the populations differ chemically whenever the geographical distance exceeds 120 km, an indicator of the minimal distance between samples required for conserving the genetic diversity of populations. Although being scarcely used with secondary metabolites, these methodologies may be used in a wide range of applications in species management and may lead to an effective integration of genetic, chemical and ecological perspectives.

**Keywords:** *Eugenia dysenterica*, essential oil, chemovariations, variogram, spatial variation

### Introduction

Phenotypic variation patterns in secondary plant metabolites have strong ecological significance and are an important factor in understanding the evolutionary history of natural populations, as they affect both intra and interspecific interactions.<sup>1</sup> Variation of chemical phenotypes can be explained by a combination of genetic or ontogenetic and environmental variation sources.<sup>2,3</sup> Spatial

factors can be critical for each of these processes and the levels and spatial structure of phenotypic chemovariations may affect competition,<sup>4</sup> local adaptation to the presence of another plant species,<sup>5</sup> pollination,<sup>6</sup> nutrient cycling,<sup>7</sup> differential freezing resistance,<sup>8</sup> foraging behavior and habitat selection,<sup>9</sup> as well as influence the capacity of herbivores and pathogens to adapt and exert selection on plant chemicals.<sup>10</sup>

Spatial patterns in essential oil variations have been described for *E. dysenterica* from different sites located in South East Goiás state, Brazil.<sup>11</sup> Multivariate Mantel

\*e-mail: pedro@quimica.ufg.br

autocorrelogram and trend surface analysis, with variance partitioning by partial redundancy analyses (pRDA),<sup>12</sup> have enabled the detection and quantification of spatial chemical patterns in the same genetically described *E. dysenterica* populations.<sup>13-15</sup> All genetic and chemical markers showed a similar profile for populations located at around 120 km, according to the isolation-by-distance model. This profile decreases as distance increases, so that populations located more than 190-200 km away become different as far as these descriptors are concerned.<sup>11,13-15</sup>

Here, the spatial chemical patterns of phenotypic variation in previously reported *E. dysenterica* essential oil data sets via spatial chemometric methods were explored.<sup>11</sup> In these techniques, theoretical variograms can be fitted to experimental ones, allowing the comparison of an observed structure with spatial chemical structures derived from chemometric models. Theoretical modeling may be used for understanding and as a predictive tool for interpolation (kriging) or probability mapping. According to this finding, it would be possible to decide which groups of local populations should be given priority in sampling or preserving,<sup>16</sup> as the distribution of genetic (chemical) variation in the geographical space is an essential factor for the conservation and management of wild populations.<sup>17</sup> Nevertheless, models predicting the spatial distribution of secondary metabolites are scarcely studied, mostly restricted to phenolic variations,<sup>18-20</sup> despite their importance in sampling and species management, as well as for individual and community success.

## Experimental

### Plant material and essential oil data

The complete data of the *E. dysenterica* sampling collection was described previously.<sup>11</sup> Briefly, essential oils from 121 *E. dysenterica* trees were extracted from eight populations in South East Goiás state, Brazil. Sampling sites and populations are shown in the Supplementary Information (SI) section (Figure S1). Essential oil data sets were represented by chemical constituents (121 samples  $\times$  49 variables) or oil constituents rearranged according to biosynthetic carbon skeletons (121  $\times$  13).

### Spatial chemometric analyses

In the spatial chemometric analyses, the fitted values on the first extracted axis of the redundancy analysis (RDA) were used to represent essential oil variability in the samples, since summarizes the multidimensional chemical information of every individual. RDA first axes correspond

to the main variance fractions explained by significant spatial and environmental predictor variables.<sup>11</sup> Spatial analyses were also conducted for the percentage values of total oxygenated terpenes (sum of monoterpenes and sesquiterpenes), total oxygenated monoterpenes and total sesquiterpene hydrocarbons, the main biosynthetic classes of essential oils.

To determine the strength and scale of spatial chemical dependence among populations, the variance (formerly the semi-variance)  $\gamma(h)$  was estimated by using the following equation 1:<sup>21</sup>

$$\gamma(h) = \frac{1}{2n(h)} \sum_{x=1}^n (Z_x - Z_{x+h})^2 \quad (1)$$

where  $n(h)$  is the number of lag class pairs at  $h$  distance intervals;  $Z_x$  and  $Z_{x+h}$  are chemical parameter values at location  $x$  and  $x+h$ , respectively. The number of lag classes ( $h = 14$ ) was determined by Sturge's rule,<sup>22</sup> an objective method which avoids arbitrarily inflating the explained spatial variation as a function of lag class size. The plot of  $\gamma(h)$  against  $h$  provides the variogram (Figure S2 in the SI section), which will show either purely random behavior or systematic behavior described by a theoretical model (linear, spherical, gaussian or power law distribution). Most models start with a variance value above zero at the  $y$ -intercept, called the nugget ( $C_0$ ), which is the unexplained variance. This can be attributed to variability at a smaller scale than the sampling resolution or to measurement errors. If there is spatial autocorrelation in the chemical data, the variance in the defined distance intervals increases from the  $y$ -intercept with distance until it reaches its maximum, known as the sill ( $C_0 + C$ ). The variance between the nugget and the sill is called the partial sill ( $C$ ) and is the part of the variance which is explained by the spatial structure of the chemical data. The ratio of the partial sill and the total sill [ $C / (C_0 + C)$ ] determines the strength of the spatial autocorrelation (structural variance,  $Q$ ), i.e., this statistic provides a measure of the proportion of sample variance ( $C_0 + C$ ) that is explained by spatially structured variance ( $C$ ). High values indicate a spatial pattern in the data.

Alternatively, the Cambardella index ( $I_c$ ) was also used to compare the degree of spatial dependence.<sup>23</sup> This index estimates the proportion of variance which is erratic (not spatially structured) and gives a good indication of how data set is spatially arranged, namely:  $I_c < 25$ : strong spatial dependence and small erratic variance;  $25 < I_c < 75$ : moderate spatial dependence; and  $I_c > 75$ : random spatial distribution. The distance in which the model reaches the sill is known as the effective range, which shows the extent of the spatial dependence on the chemical data. Furthermore, the mean correlation distance ( $MCD = \frac{3}{8} \times \text{range} \times Q$ ) was

computed at each variable to compare the distance in which a high spatial dependence occurred in the variogram.<sup>20</sup> If the model shows only the nugget, then there is no spatial structure in the data.

The coefficient of variance (CV) was calculated for all response variables, and the nugget, sill, effective range,  $Q$ ,  $I_c$ , correlation coefficient ( $r$ ), and MCD parameters were obtained from the model with the best fit to the variance data. All sample data were investigated for anisotropy by using directional variograms before calculating omnidirectional variograms. When spatial dependency was present, the spatial distribution of the chemical parameter ( $Z_x$ ) was expressed by a kriged map. This is a linear interpolation technique that provides the best linear unbiased estimation, which minimizes prediction error variances for spatial variables. Although the resulting maps provide a powerful visualization of the spatial pattern, probability maps that exceed the mean value of the response variables were used to avoid excessive smoothing in the kriged map.<sup>18</sup> Probabilities were calculated by performing 1000 conditional simulations on the predicted distribution of values from response variables.

To assess environmental influence on oil variability, leaf nutrients (environmental predictors: P, K<sup>+</sup>, Mg<sup>2+</sup> and Cu<sup>2+</sup>), represented by the first axes of the partial RDA,<sup>11</sup> were submitted to the probability to exceed the mean value maps. Thus, the maps of pure environmental influences were obtained shared with spatial descriptors. All calculations were performed with GS+ (Geostatistics for the Environmental Sciences) and variogram plotting was completed using the R library Gstat<sup>24</sup> version 2.15.0 of the R package.<sup>25</sup>

## Results and Discussion

To identify the existence of spatial structures and describe the spatial variability of response variables, variograms of RDA first axes from each oil data set, percentage values of oxygenated terpenes, oxygenated monoterpene and sesquiterpene hydrocarbons were computed for sampling populations. The study of variograms is important in natural

product chemistry, as its nugget effect provides information on the error made by the measuring instruments and by chemovariations undetected in the sampling neighborhood. In addition, they offer valuable information regarding the spatial dependence of each response variable (nugget/sill ratio) and the effective range of spatial autocorrelation, which supports the establishment of operational chemical units,<sup>11</sup> similarly to genetic operational units for the purposes of conserving and managing populations.<sup>26</sup>

In the present study, the gaussian models fitted the experimental variograms with great success in the majority of oil response variables (Table 1). This confirms the existence of a spatial structure in essential oil contents and rules out the possibility of a random distribution of essential oils throughout population sites. According to the Cambardella index,<sup>23</sup> the spatial dependence of essential oil content, as measured by each oil data set, was the highest among all response variables.

The nugget/sill ratios were 0.04 and 0.05% for oil constituents and oil carbon skeleton, respectively. Thus, at least 99.9% of total variance in essential oil data sets can be explained by spatially structured variance. On the other hand, the proportion of the nugget effect was greater for oxygenated monoterpenes (49.9%) than for the sum of oxygenated terpenes (8.9%), and in general reflected a moderate spatial dependence of the former biosynthetic class. A similar trend was observed for the coefficient of variation. For the oxygenated monoterpenes, the variability (CV = 62.8%) was more than 2.5 fold higher than the variability caused by changes in the oxygenated terpenes (CV = 24.9%). In contrast, the variogram model for sesquiterpene hydrocarbons (CV = 19.2%) showed only a nugget effect, which suggests there is no spatial structure in the data (Figure S3 in the SI section).

For both oxygenated biosynthetic classes, the effective range of the variograms, which is the distance at which a plateau is reached, progressively increased from monoterpenes to the sum of oxygenated terpenes. The MCD value is an informative parameter of the range of distances within which a high spatial dependence exists

**Table 1.** Parameters for the variogram models of the response variables from *E. dysenterica* essential oils

Response variable	CV / % <sup>a</sup>	Model	Nugget	Sill	Effective range / m	$I_c$ / % <sup>b</sup>	$Q$ / % <sup>c</sup>	$r$	MCD / m <sup>d</sup>
Oil constituents <sup>e</sup>	27.1	gaussian	0.0001	0.226	120204.3	0.04	99.96	0.781	45056.7
Oil carbon skeleton <sup>e</sup>	28.9	gaussian	0.0001	0.199	153113.3	0.05	99.95	0.803	57388.7
Oxygenated terpenes	24.9	gaussian	0.0030	0.034	138044.5	8.9	91.1	0.844	47146.9
Oxygenated monoterpenes	62.8	gaussian	0.1852	0.371	77595.9	49.9	50.1	0.901	14588.4
Sesquiterpene hydrocarbons	19.2	linear	135.4	135.4	111301.3	100.0	0.0	0.322	0.0

<sup>a</sup>Coefficient of variation; <sup>b</sup>Cambardella index ( $C_0 / [C_0 + C]$ ); <sup>c</sup>structural variance ( $C / [C_0 + C]$ ); <sup>d</sup>mean correlation distance ( $3/8 \times \text{range} \times Q$ ); <sup>e</sup>values of RDA axis 1.

among populations. According to Table 1, the MCD value tended to expand throughout the oxygenated terpene biosynthesis (from 15 km for monoterpenes to 47 km with oxygenated sesquiterpenes) indicating a progressive increase in the dimension of geographical areas that display similar oxygenated terpene contents.

For oxygenated monoterpenes, the MCD value (14588.4 m) was only twice the lag class distance interval (7823.8 m), leading to a higher proportion of nugget effect with moderate spatial dependence ( $C_0 = 0.1852$ ,  $I_c = 49.9\%$ ). In other words, the variability of the oxygenated monoterpene content in the essential oil occurred at shorter distances than those involving the other response variables, whenever two adjacent sampling sites were taken into account. These results confirm the change in the spatial pattern of the terpene content in the essential oil throughout different biosynthetic classes. The high nugget effect proportion may also reflect a strong degree of intra or interspecimen variation. Furthermore, this effect should be related at least partially to the collection of leaf samples, which occurred in July, during the dry season. During this time, the peak of leafing activities, senescence and emission of new leaves occur,<sup>27</sup> thus requiring large amounts of carbon and macronutrients for protein and RNA synthesis, which is markedly increased in young leaves with a high capacity for biosynthesized essential oils. During leaf growth, leaf volatiles may provide a constitutive defense (by deterring potential herbivores) or an induced response to herbivore damage by attracting predators or parasites.<sup>28</sup> However, as biosynthetic terpenes changed, differences in oxygenated sesquiterpene content decreased, much in the same way that the effective range and the MCD value increase for this class. Presumably, this effect reveals that a plateau has been reached in the essential oil, which results in intra and interpopulation homogeneity. The observed effective range for the oil constituent data set (120.2 km) occurred in a similar geographical distance (117 km) previously determined by multivariate Mantel autocorrelogram.<sup>11</sup> These results contrast with numerical simulations,<sup>29</sup> of which the Mantel test and derived forms could not correctly estimate the proportion of the original data variation explained by spatial structures.

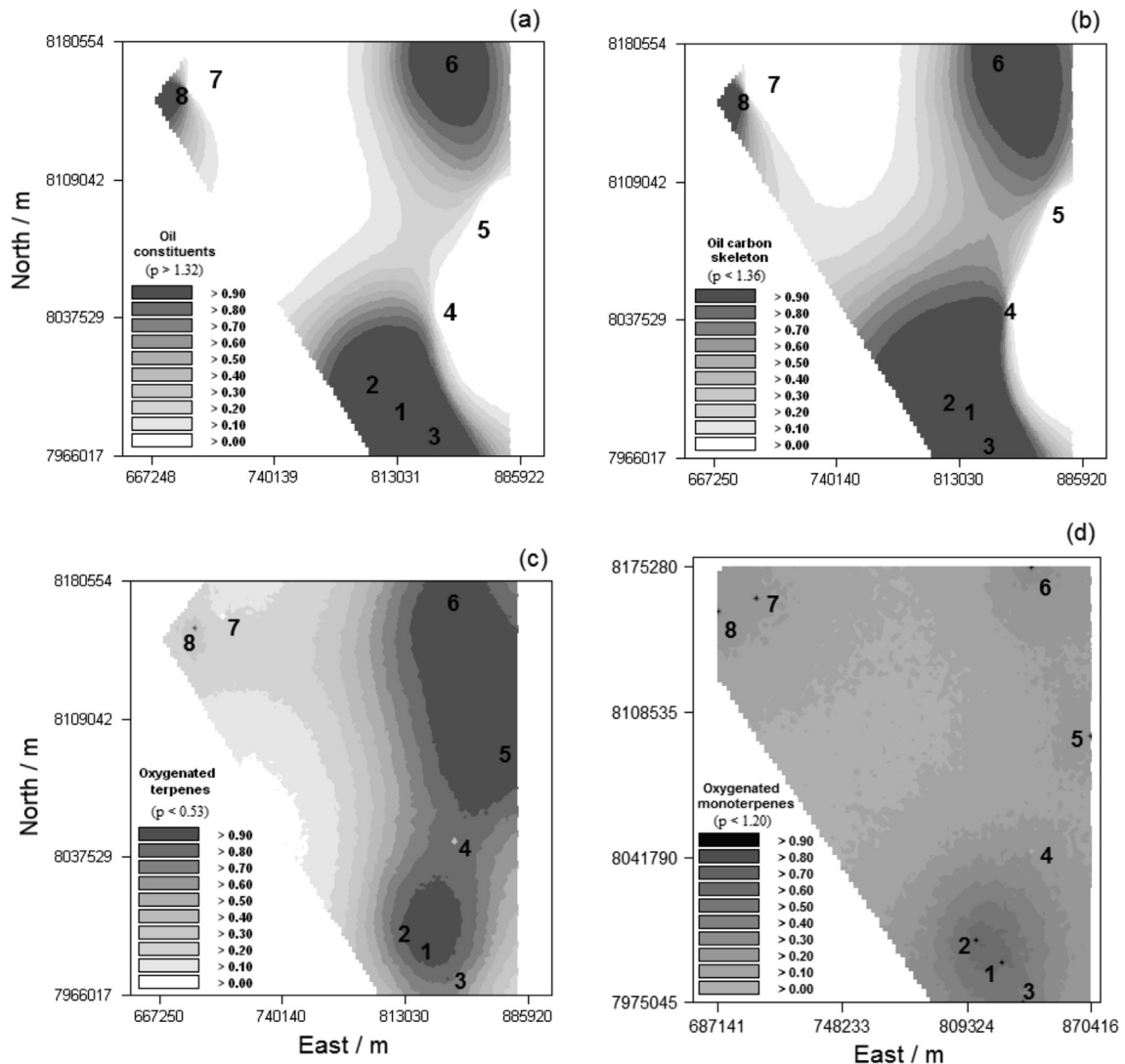
The spatial pattern for *E. dysenterica* essential oils is in agreement with the evolutionary isolation-by-distance model observed for morphological and isozymatic descriptors,<sup>13,14</sup> as well as for simple sequence repeats (SSR) and random amplification of polymorphic DNA (RAPD) genetic markers.<sup>15</sup> All of these genetic descriptors showed a similar genetic profile for populations located at around 120 km. This similarity between the chemical data with genetic markers is consistent with the findings of

other studies which used terpenoids and isozymes, RAPD, amplified fragment length polymorphism (AFLP), SSR and inter-simple sequence repeat (ISSR) molecular markers.<sup>30</sup> Predictive maps (Figure 1) based on fitted models for each spatially structured biosynthetic class and for each essential oil data set were obtained using the probability to exceed threshold value (mean value).

All the maps showed greater variation between populations with certain similarities in the spatial patterns of two well-differentiated areas: an area with high essential oil variation in populations located along the Corumbá River basin (north-south direction of the plot) and another area (east-west) with low oil complexity. This chemical variability may be explained as a result of localized inbreeding effects associated with a low gene migration rate among populations.<sup>15,16,31</sup>

The Corumbá River basin separates the two North West populations, Goiânia (8) and Senador Canedo (7), through a depression formed by the river and its tributaries. This spatial barrier could contribute at least partially to ecological isolation, a pre-requisite for speciation and chemovariation between two sampling sites. In fact, the existence of chemotypic differentiation between populations from these sites could be confirmed by the fact that cultivated plants grown adjacently in the same environment displayed the typical composition of their wild populations.<sup>32</sup> Furthermore, samples from Senador Canedo and Goiânia tend to reveal a certain level of segregation for some response variables (Figures 1a and 1b). These variations were not previously observed and may be attributed to the Meia Ponte River basin that separates these sampling locations. On the other hand, samples from populations 4 (Campo Alegre de Goiás) and 5 (Cristalina) showed low oil variability (Figures 1a and 1b) or corresponded to a transition between the highest variability centers (Figures 1c and 1d). These populations are located between two geographical barriers, the mountain ridges of Cristais and Contraforte Central, which may be an important factor for isolation and speciation.

Another source of oil variation was the nutrient resource variability due to the fact that the environmental predictors were leaf nutrients (P, K<sup>+</sup>, Mg<sup>2+</sup> and Cu<sup>2+</sup>), which account for 8% of the chemovariation.<sup>11</sup> The maps created for pure environmental influences and shared with spatial predictors coincide with the regions of oil variation (Figure S4 in the SI section) and are a relative indication of the divergence in terpenes. These micronutrient effects should be associated with a strict requirement for sesquiterpene synthases with divalent metal ions as cofactors, especially Mg<sup>2+</sup>, which have also influenced the number of by-products obtained from these reactions.<sup>33</sup>



**Figure 1.** Spatial distribution of the essential oil variations in the leaves from eight *E. dysenterica* populations from Central Brazilian Cerrado: oil constituents (a) and oil constituents arranged according to carbon skeleton (b); oxygenated terpene (c) and oxygenated monoterpene (d) biosynthetic classes. The maps show the probability of finding a value higher than the mean content.

Terpenoid accumulation was related to high P and  $K^+$  soil contents or when culture media were supplemented with increased P concentration.<sup>34</sup> On the other hand, populations located around the Meia Ponte River (7 and 8) and the mountainous populations (4 and 5) were discriminated by pure spatial influence and by the fact that they are shared by environmental predictors along the sampling sites (Figure S5 in the SI section).

Unlike the frequent assessment of the geographical distribution of genetic diversity,<sup>35</sup> the description of the spatial structure of essential oil variability has been scarcely studied until now.<sup>11,36</sup> These studies have focused on spatial autocorrelation approaches, such as the Mantel test and other derived forms, which could not correctly estimate the proportion of the original data variation explained by spatial structures.<sup>29</sup> Trend surface analyses have faced

difficulties in interpreting the terms of the polynomials, in addition to the non-independence of the monomials.<sup>37</sup> Furthermore, they have been unable to correctly model fine-scale patterns.<sup>38</sup> Thus, spatial chemometric methods which use variograms of essential oil chemovariations may be used as additional tools to establish *in situ* conservation areas or sampling areas for *ex situ* conservation. In addition, the ecological role of essential oils in neotropical Myrtaceae has not yet been investigated,<sup>39</sup> but the studies referred to here are very important for conservation and management strategies.

## Conclusions

In the present study, spatial chemometric methods detected and characterized the spatial chemical structure

among populations of *E. dysenterica* based on fitted models using variograms and probability maps. In addition to multivariate Mantel autocorrelagram and trend surface analyses,<sup>11</sup> all chemical strategies showed that *E. dysenterica* populations differ chemically whenever the geographical distance exceeds 120 km, an indicator of the minimal distance between samples required for conserving the genetic diversity of populations. Although being scarcely used with essential oil variations, these methodologies show a wide range of applications in species management and may lead to an effective integration of genetic, chemical and ecological perspectives.

## Supplementary Information

Supplementary data (Figures S1-S5) is available free of charge at <http://jbcs.sbq.org.br> as a PDF file.

## Acknowledgements

We are indebted to the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) and to the Programa de Apoio ao Desenvolvimento Científico e Tecnológico (PADCT III) for their financial support, as well as to the Fundação de Amparo à Pesquisa do Estado de Goiás (FAPEG) for granting a fellowship to E. C. V.

## References

- Brenes-Arguedas, T.; Coley, P. D.; *Oikos* **2005**, *108*, 410; Andrew, R. L.; Peakall, R.; Wallis, I. R.; Foley, W. J.; *Ecology* **2007**, *88*, 716.
- Hamilton, J. G.; Zangerl, A. R.; DeLucia, E. H.; Berenbaum, M. R.; *Ecol. Lett.* **2001**, *4*, 86; Laitinen, M.-L.; Julkunen-Tiitto, R.; Tahvanainen, J.; Heinonen, J.; Rousi, M.; *J. Chem. Ecol.* **2005**, *31*, 697; Ballhorn, D. J.; Schiwyl, S.; Jensen, M.; Heil, M.; *J. Chem. Ecol.* **2008**, *34*, 1298.
- Agrell, J.; McDonald, E. P.; Lindroth, R. L.; *Oikos* **2000**, *88*, 259; Silva, F. G.; Oliveira, C. B. A.; Pinto, J. E. B. P.; Nascimento, V. E.; Santos, S. C.; Seraphin, J. C.; Ferri, P. H.; *J. Braz. Chem. Soc.* **2007**, *18*, 990; Costa, D. P.; Santos, S. C.; Seraphin, J. C.; Ferri, P. H.; *J. Braz. Chem. Soc.* **2009**, *20*, 1287.
- Hull-Sanders, H. M.; Clare, R.; Johnson, R. H.; Meyer, G. A.; *J. Chem. Ecol.* **2007**, *33*, 781; Broz, A. K.; Broeckling, C. D.; De-la-Peña, C.; Lewis, M. R.; Greene, E.; Callaway, R. M.; Sumner, L. W.; Vivanco, J. M.; *BMC Plant Biol.* **2010**, *10*, 115; Lages, B. G.; Fleury, B. G.; Hovell, A. M. C.; Rezende, C. M.; Pinto, A. C.; Creed, J. C.; *Mar. Biol.* **2012**, *159*, 1551.
- Ehlers, B. K.; Thompson, J. D.; *Oecologia* **2004**, *141*, 511; Grondahl, E.; Ehlers, B. K.; *J. Ecol.* **2008**, *96*, 981.
- Raguso, R. A.; *Entomol. Exp. Appl.* **2008**, *128*, 196; Hanhineva, K.; Rogachev, I.; Kokko, H.; Mintz-Oron, S.; Venger, I.; Kärenlampi, S.; Aharoni, A.; *Phytochemistry* **2008**, *69*, 2463; Ayasse, M.; Stöckl, J.; Francke, W.; *Phytochemistry* **2011**, *72*, 1667.
- Ushio, M.; Adams, J. M.; *Open Ecol. J.* **2011**, *4*, 18; Cesco, S.; Mimmo, T.; Tonon, G.; Tomasi, N.; Pinton, R.; Terzano, R.; Neumann, G.; Weisskopf, L.; Renella, G.; Landi, L.; Nannipieri, P.; *Biol. Fertil. Soils* **2012**, *48*, 123.
- Amiot, J.; Salmon, Y.; Collin, C.; Thompson, J. D.; *Ecol. Lett.* **2005**, *8*, 370.
- DeGabriel, J. L.; Moore, B. D.; Marsh, K. J.; Foley, W. J.; *Chemoecology* **2010**, *20*, 97.
- Thrall, P. H.; Burdon, J. J.; Young, A.; *J. Ecol.* **2001**, *89*, 736; Macel, M.; Vrieling, K.; Klinkhamer, P. G. L.; *Phytochemistry* **2004**, *65*, 865; Kleine, S.; Muller, C.; *Oecologia* **2011**, *166*, 175.
- Vilela, E. C.; Carvalho, T. C.; Duarte, A. R.; Naves, R. V.; Santos, S. C.; Seraphin, J. C.; Ferri, P. H.; *J. Braz. Chem. Soc.* **2012**, *23*, 1776.
- Legendre, P.; Fortin, M.-J.; *Vegetatio* **1989**, *80*, 107; Borcard, D.; Legendre, P.; Drapeau, P.; *Ecology* **1992**, *73*, 1045.
- Telles, M. P. C.; Diniz-Filho, J. A. F.; Coelho, A. S. G.; Chaves, L. J.; *Rev. Bras. Bot.* **2001**, *24*, 145; Telles, M. P. C.; Silva, R. S. M.; Chaves, L. J.; Coelho, A. S. G.; Diniz-Filho, J. A. F.; *Pesq. Agropec. Bras.* **2001**, *36*, 1387.
- Trindade, M. G.; Chaves, L. J.; *Genet. Mol. Biol.* **2005**, *28*, 407.
- Zucchi, M. I.; Brondani, R. V.; Pinheiro, J. B.; Coelho, A. S. G.; Chaves, L. J.; Vencovsky, R.; *Rev. Bras. Bot.* **2003**, *26*, 449; Zucchi, M. I.; Pinheiro, J. B.; Chaves, L. J.; Coelho, A. S. G.; Couto, M. A.; Morais, L. K.; Vencovsky, R.; *Pesq. Agropec. Bras.* **2005**, *40*, 975.
- Telles, M. P. C.; Coelho, A. S. G.; Chaves, L. J.; Diniz-Filho, J. A. F.; Valva, F. D.; *Conserv. Genet.* **2003**, *4*, 68.
- McCue, K. A.; Buckler, E. S.; Holtsford, T. P.; *Conserv. Biol.* **1996**, *10*, 1425.
- Covelo, F.; Gallardo, A.; *Plant Soil* **2004**, *259*, 267; Covelo, F.; Gallardo, A.; *For. Ecol. Manage.* **2009**, *258*, 2511.
- Cortell, J. M.; Halbleib, M.; Gallagher, A. V.; Righetti, T. L.; Kennedy, J. A.; *J. Agric. Food Chem.* **2005**, *53*, 5798; Wu, D.-C.; Zhou, J.-H.; Fan, Z.-D.; Li, W.-B.; Zhang, Y.-Y.; Yang, C.; Xu, F.-H.; Zhu, K.; Duaj, F.-Y.; Wang, D.-X.; *Acta Agron. Sin.* **2010**, *36*, 141.
- Baluja, J.; Diago, M. P.; Goovaerts, P.; Tardaguila, J.; *Aust. J. Grape Wine Res.* **2012**, *18*, 173.
- Bachmaier, M.; Backes, M.; *Precis. Agric.* **2008**, *9*, 173.
- Legendre, P.; Legendre, L.; *Numerical Ecology*, 2<sup>nd</sup> ed.; Elsevier Science: Amsterdam, 2003.
- Cambardella, C. A.; Karlen, D. L.; *Precis. Agric.* **1999**, *1*, 5.
- Robertson, G. P.; *GS+: Geostatistics for the Environmental Sciences*, version 9.0; Gamma Design Software: Michigan, USA, 2008; Pebesma, E. J.; Wesseling, C. G.; *Comput. Geosci.* **1998**, *24*, 17.

25. R Development Core Team; *R: a Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, 2008.
26. Diniz-Filho, J. A. F.; Telles, M. P. C.; *Conserv. Biol.* **2002**, *16*, 924.
27. Sano, S. M.; Fonseca, C. E. L.; Ribeiro, J. F.; Oga, F. M.; Luiz, A. J. B.; *Pesq. Agropec. Bras.* **1995**, *30*, 5.
28. Dicke, M.; Hilker, M.; *Basic Appl. Ecol.* **2003**, *4*, 3; Jurgens, A.; Feldhaar, H.; Feldmeyer, B.; Fiala, B.; *Biochem. Syst. Ecol.* **2006**, *34*, 97.
29. Legendre, P.; Fortin, M.-J.; *Mol. Ecol. Resour.* **2010**, *10*, 831.
30. Nan, P.; Peng, S.; Shi, S.; Ren, H.; Yang, J.; Zhong, Y.; *Z. Naturforsch., C: Biosci.* **2003**, *58*, 57; Adams, R. P.; Socorro, M.; Elizondo, G.; Elizondo, M. G.; Slinkman, E.; *Biochem. Syst. Ecol.* **2006**, *34*, 205; Trindade, H.; Costa, M. M.; Lima, S. B.; Pedro, L. G.; Figueiredo, A. C.; Barroso, J. G.; *Biochem. Syst. Ecol.* **2009**, *37*, 670; Ben El Hadj Ali, I.; Guetat, A.; Boussaid, M.; *Ind. Crops Prod.* **2012**, *40*, 277.
31. Chaves, L. J.; Vencovsky, R.; Silva, R. S. M.; Telles, M. P. C.; Zucchi, M. I.; Coelho, A. S. G.; *Conserv. Genet.* **2011**, *12*, 569.
32. Duarte, A. R.; Naves, R. V.; Santos, S. C.; Seraphin, J. C.; Ferri, P. H.; *J. Braz. Chem. Soc.* **2009**, *20*, 967; Duarte, A. R.; Naves, R. V.; Santos, S. C.; Seraphin, J. C.; Ferri, P. H.; *J. Braz. Chem. Soc.* **2010**, *21*, 1459.
33. Picaud, S.; Olofsson, L.; Brodelius, M.; Brodelius, P. E.; *Arch. Biochem. Biophys.* **2005**, *436*, 215.
34. Kosínski, I.; *Pol. Ecol. Stud.* **1996**, *22*, 17; Ormeño, E.; Fernandez, C.; Bousquet-Mélou, A.; Greff, S.; Morin, E.; Robles, C.; Vila, B.; Bonin, G.; *Atmos. Environ.* **2007**, *41*, 629.
35. Stich, B.; Haussmann, B. I. G.; Pasam, R.; Bhosale, S.; Hash, T.; Melchinger, A. E.; Parzies, H. K.; *BMC Plant Biol.* **2010**, *10*, 216.
36. Tommasi, L.; Negro, C.; Cerfeda, A.; Nutricati, E.; Zuccarello, V.; De Bellis, L.; Miceli, A.; *J. Essent. Oil Res.* **2007**, *19*, 572; Tommasi, L.; Negro, C.; De Bellis, L.; Miceli, A.; *J. Essent. Oil Res.* **2008**, *20*, 295; Zaouali, Y.; Boussaid, M.; *Biochem. Syst. Ecol.* **2008**, *36*, 11; Ben El Hadj Ali, I.; Zaouali, Y.; Bejaoui, A.; Boussaid, M.; *Chem. Biodivers.* **2010**, *7*, 1276.
37. Jombart, T.; Dray, S.; Dufour, A.-B.; *Ecography* **2009**, *32*, 161.
38. Borcard, D.; Legendre, P.; *Ecol. Modell.* **2002**, *153*, 51.
39. Stefanello, M. E. A.; Pascoal, A. C. R. F.; Salvador, M. J.; *Chem. Biodivers.* **2011**, *8*, 73.

Submitted: February 6, 2013

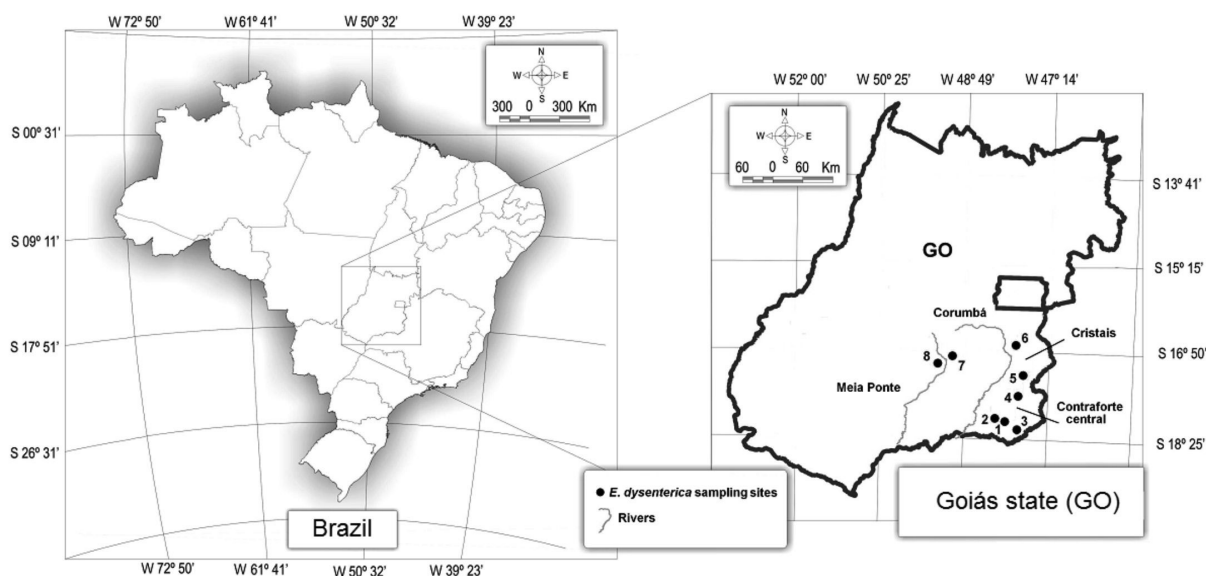
Published online: April 24, 2013

# Supplementary Information

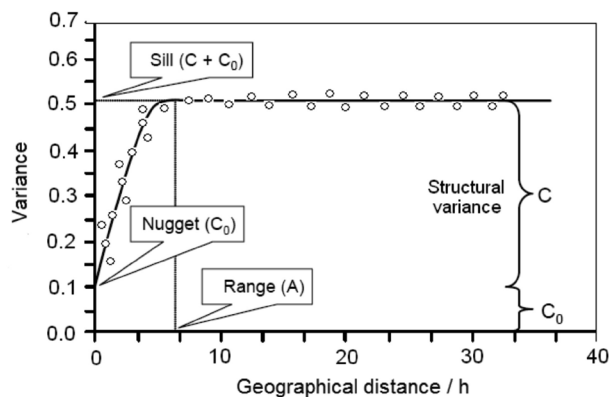
## Spatial Chemometric Analyses of Essential Oil Variability in *Eugenia dysenterica*

*Eliane C. Vilela,<sup>a</sup> Alessandra R. Duarte,<sup>a</sup> Ronaldo V. Naves,<sup>b</sup> Suzana C. Santos,<sup>a</sup>  
José C. Seraphin<sup>c</sup> and Pedro H. Ferri<sup>\*a</sup>*

<sup>a</sup>Laboratório de Bioatividade Molecular, Instituto de Química,  
<sup>b</sup>Setor de Horticultura, Escola de Agronomia e de Engenharia de Alimentos and  
<sup>c</sup>Núcleo de Estatística Aplicada, Instituto de Matemática e Estatística,  
Universidade Federal de Goiás, CP 131, 74001-970 Goiânia-GO, Brazil

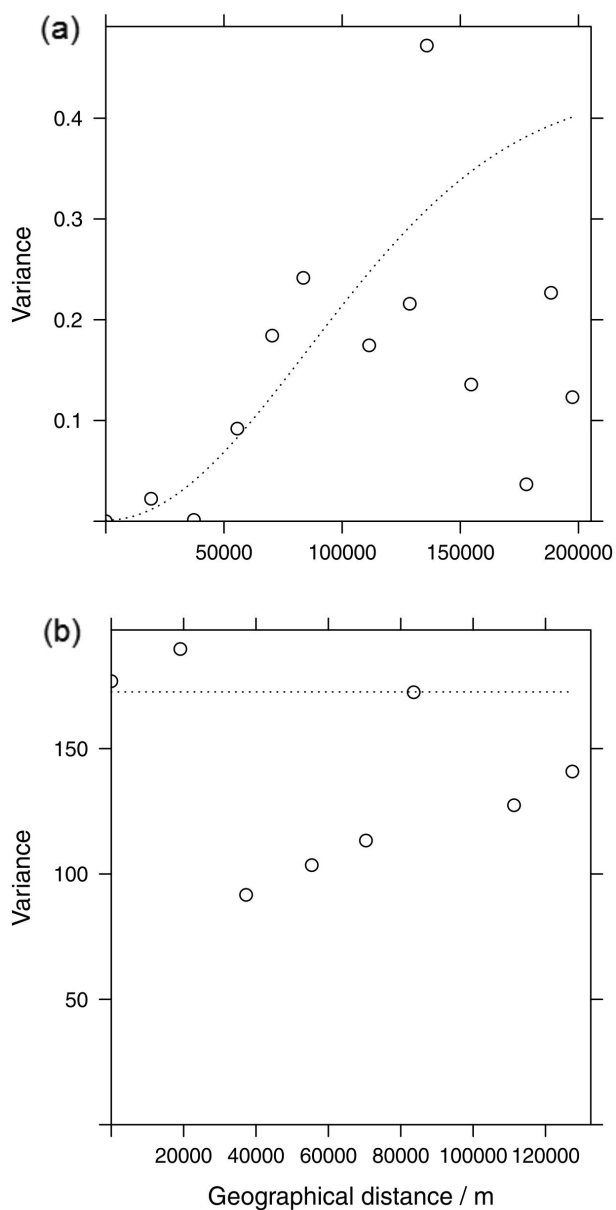


**Figure S1.** Map of Goiás State showing Corumbá and Meia Ponte River basins separating *E. dysenterica* populations 1-6 and 7 from 8, respectively. Cristais and Contraforte Central mountain ridges help to isolate populations 4 and 5. Populations: 1 = Catalão-1, 2 = Catalão-2, 3 = Três Ranchos, 4 = Campo Alegre de Goiás, 5 = Cristalina, 6 = Luziânia, 7 = Senador Canedo, 8 = Goiânia.

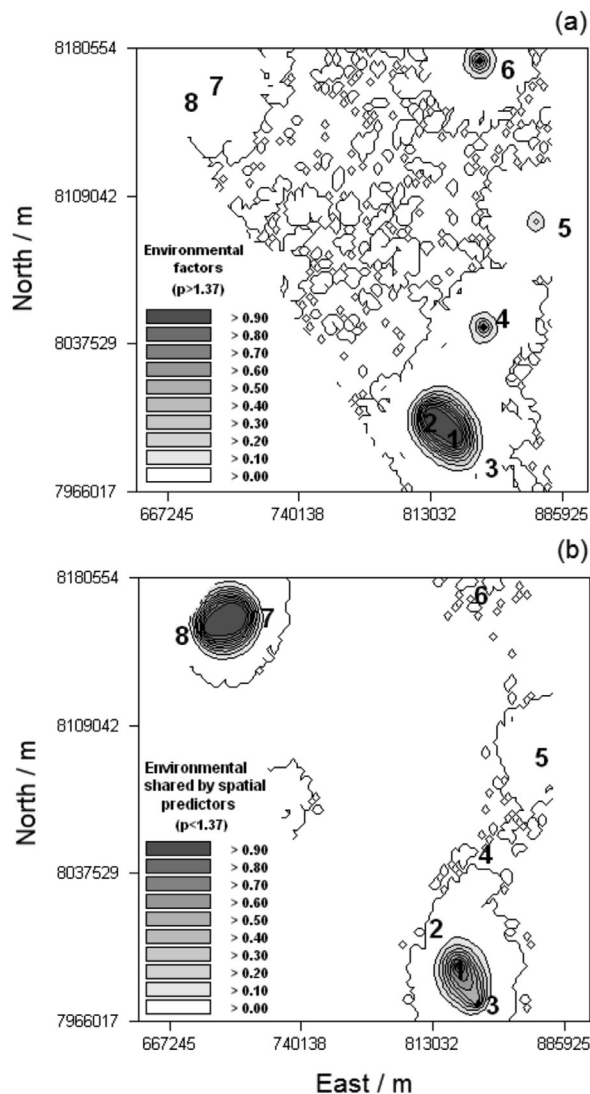


**Figure S2.** Generalized variogram model showing estimate parameters for sill, nugget and effective range. The spatially structured variance is represented by fraction C.

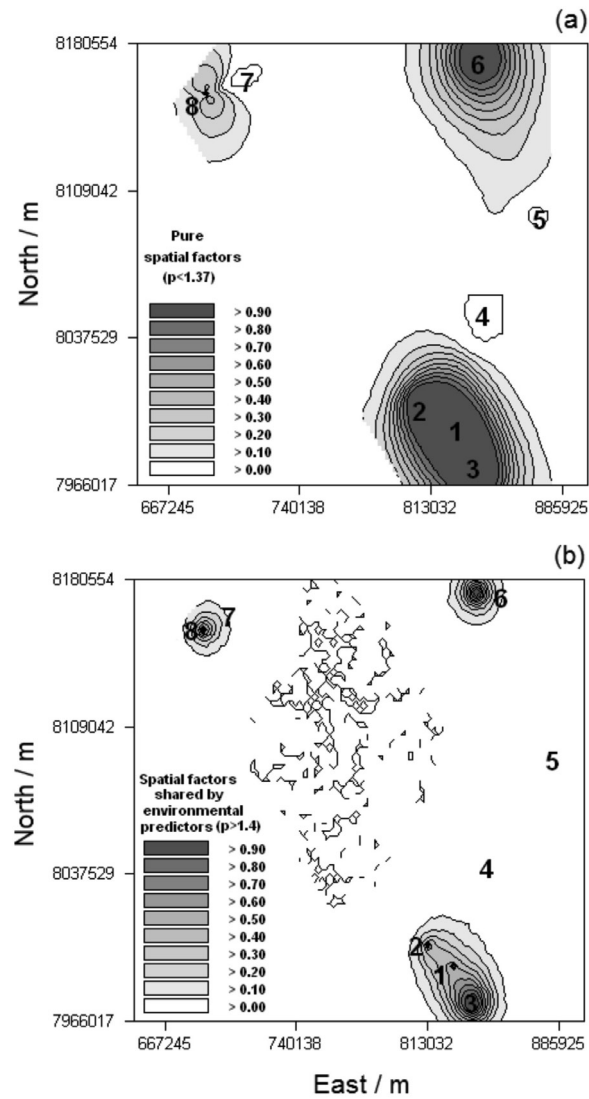




**Figure S3.** Variograms for essential oils from eight *E. dysenterica* populations represented by oil constituent data set (a) and for the sesquiterpene hydrocarbon content (b) along the separation distance of sampling sites. Fitted models show spatially structured variance (a) and nugget effect only (b), with no spatial structure in the data.



**Figure S4.** Distribution of environmental predictors (leaf nutrients: P, K<sup>+</sup>, Mg<sup>2+</sup>, Cu<sup>2+</sup>) of oil constituent variations in the leaves from eight *E. dysenterica* populations from Central Brazilian Cerrado: pure environmental influence (a) and environmental factors shared by spatial predictors (b). The maps show the probability of finding a value higher than the mean content.



**Figure S5.** Distribution of spatial predictors of oil constituent variations in the leaves from eight *E. dysenterica* populations from Central Brazilian Cerrado: pure spatial influence (a) and spatial factors shared by environmental predictors (b). The maps show the probability of finding a value higher than the mean content.