

# Identification of Honey from Different Producing Areas Based on Principal Component Analysis

Ying Zhang\*, Zhanfeng Ma, Xin Liu

Institute of Product Quality Supervision and Inspection, Harbin Heilongjiang  
Email: \*571414827@qq.com

Received: Jul. 2<sup>nd</sup>, 2020; accepted: Jul. 20<sup>th</sup>, 2020; published: Jul. 27<sup>th</sup>, 2020

## Abstract

According to the differences of components in honey, the method of identifying honey producing area was established. We took 215 batches of honey samples from 3 areas in northeast China as the research object, the components of honey from different areas were determined and analyzed. The differences of components in honey in different producing areas were conducted based on one-way ANOVA and principal component analysis using R studio. The results showed that the honey components among different producing areas were significantly different and the significant level was  $P < 0.01$ . The cumulative contribution rate of the first two principal components in principal component analysis was 82.1%, which could represent the main information of honey components. The samples from the same place were clustered together in the plot of principal components, while the samples from different places were separated from each other. This method can effectively distinguish the producing areas of honey samples. The honey components such as glucose, fructose, sucrose, maltose and amylase contributed more to the first principal component (PC1), which indicated that these honey components was the main basis to distinguish the producing areas of honey samples. Principal component analysis of honey provides the experimental basis for the quality control of honey, and the identification method of honey producing area.

## Keywords

Honey, Principal Component Analysis, Honey Components, Honey Producing Area

# 基于主成分分析对不同产地蜂蜜的鉴别研究

张莹\*, 马占峰, 刘欣

哈尔滨市产品质量监督检验院, 黑龙江 哈尔滨  
Email: \*571414827@qq.com

\*通讯作者。

## 摘要

本文利用蜂蜜中各个成分含量的差异，建立鉴别蜂蜜产地的方法。以我国东北地区3个产地的215批次蜂蜜样品为研究对象，对不同产地蜂蜜成分的指标进行测定及分析，并利用R语言软件，对不同产地蜂蜜成分的指标数据进行单因素方差分析和主成分分析。结果显示，不同产地蜂蜜成分差异显著性水平为 $P < 0.01$ ，蜂蜜成分在不同产地间具有极显著差异。主成分分析的前2个主成分累积贡献率达到82.1%，能表达蜂蜜成分的主要信息，主成分分散点图中，同一产地样品成簇聚集，不同产地样品分布于相对独立的空间，能有效区分不同产地的蜂蜜样品。葡萄糖、果糖、蔗糖、麦芽糖、淀粉酶等蜂蜜成分对第一主成分(PC1)贡献率较大，表明这些蜂蜜成分含量的差异是区分不同产地蜂蜜样品的主要依据，本实验通过对蜂蜜成分的主成分分析，为蜂蜜的质量控制提供试验依据，提供蜂蜜产地鉴别手段。

## 关键词

蜂蜜，主成分分析，蜂蜜成分，蜂蜜产地

Copyright © 2020 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

蜂蜜是蜜蜂采集植物的花蜜、分泌物或蜜露，与自身分泌物混合后，经充分酿造而成的天然物质，主要的营养成分包括碳水化合物、蛋白质、矿物质、维生素和多酚类物质等[1]，现已发现大约有320种不同种类的蜂蜜，它们来自不同的植物。蜂蜜中的各种成分和含量随产地的气候条件、产地的土壤性质及养蜂管理技术的不同而存在差异，因此蜂蜜的成分常常与蜂蜜产地息息相关。但是，在蜂蜜的销售和流通环节，一些不法商贩将蜂蜜胡乱贴标签，使得蜂蜜产地混乱，扰乱市场秩序，为了减少人们购买假冒蜂蜜的损失，保证食品的安全，一些发达国家实施食品安全追溯制，有了食品安全追溯制，便可以成为消费者购买食品时选择的标准。本文通过蜂蜜成分指标的测定并结合主成分分析，对不同产地蜂蜜样品成分含量的数据进行了研究，建立了3个不同产地蜂蜜的聚类，为不同产地蜂蜜鉴别和质量控制提供手段。

## 2. 材料与方方法

### 2.1. 实验材料

直接从蜂农处购得215批次椴树蜂蜜(在室温条件下密封贮藏)，分别来自A、B、C三个地区。

### 2.2. 蜂蜜成分指标的测定方法

蜂蜜成分各指标按照现有实行的国标方法的步骤进行测定，如表1所示。

### 2.3. 数据分析

#### 2.3.1. 单因素方差分析法

一个复杂的事物受多个相互制约的因素影响，通过统计分析中的方差分析来科学的解决在科学实验

过程中,不同实验条件下样本间平均值的差异情况,从而了解不同的试验条件对试验结果的影响情况[2]。本研究经方法测定蜂蜜中八项成分指标,对这些指标利用单因素方差分析分析各产地蜂蜜成分对产地变异结果的贡献情况,从而分析出不同产地蜂蜜成分的显著性差异。

**Table 1.** Determination methods of various indicators of honey components

**表 1.** 蜂蜜成分各指标的测定方法表

序号	测定指标	测定方法
1	果糖	GB 5009.8-2016
2	葡萄糖	GB 5009.8-2016
3	蔗糖	GB 5009.8-2016
4	麦芽糖	GB 5009.8-2016
5	淀粉酶值	GB 18932.16-2003
6	水分	SN/T 0852-2012
7	电导率	GB 18932.15-2003
8	酸度	GH/T 1141-2017

### 2.3.2. 主成分分析法

主成分分析法是一种做数据筛选的有效的统计分析方法,它借助于一个正交变换,将其分量相关的原随机向量转化成其分量不相关的新随机向量,这在代数上表现为将原随机向量的协方差阵变换成对角形阵,在几何上表现为将原坐标系变换成新的正交坐标系,使之指向样本点散布最开的多个正交方向。主成分分析是一种无监督的降维方法,通常用 2~3 个主成分就可以对指标原始变量进行解释说明,适合于高维数据的定量及定性分析[3]。本研究基于 R 语言软件,利用主成分分析对蜂蜜成分指标进行分析,具体以显示出不同产地蜂蜜成分上的差异,并分析出对不同产地蜂蜜鉴别有较大贡献的蜂蜜成分,以达到蜂蜜产地的鉴别。

## 3. 结果与分析

### 3.1. 蜂蜜各成分指标单因素方差分析结果

对蜂蜜八个指标进行单因素分析,由表 2 可知,八个指标在不同产地间的差异显著性水平为  $P < 0.01$ ,说明不同产地蜂蜜成分间具有极显著差异。

**Table 2.** The results of one-way ANOVA of components of honey

**表 2.** 蜂蜜各成分指标单因素方差分析计算结果

		自由度	离差平方和	均方	F 值	P 值
果糖	组间	2	305.805	152.902	117.334	<0.01
	组内	212	276.264	1.303		
葡萄糖	组间	2	727.205	363.603	301.991	<0.01
	组内	212	255.251	1.204		
蔗糖	组间	2	682.078	341.039	93.404	<0.01
	组内	212	774.062	3.651		
麦芽糖	组间	2	52.983	26.491	267.862	<0.01
	组内	212	20.967	0.099		
果糖与葡萄糖之和	组间	2	1841.540	920.770	201.800	<0.01
	组内	212	97.310	4.563		

Continued

酸度	组间	2	14744.720	7372.360	133.268	<0.01
	组内	212	11727.839	55.320		
水分	组间	2	194.862	97.431	140.980	<0.01
	组内	212	146.513	0.691		
电导率	组间	2	1063894.966	531947.483	766.211	<0.01
	组内	212	147182.569	694.257		

### 3.2. 蜂蜜成分的主成分分析

对 72 批次 A 产地蜂蜜样品、83 批次 B 产地蜂蜜样品和 60 批次 C 产地蜂蜜样品的成分指标进行主成分分析, 前 3 个主成分的特征值及方差贡献率如表 3 所示。其中, 第 1 主成分(PC1)对方差的贡献率为 57.3%, 第 2 主成分(PC2)对方差的贡献率为 24.8%, 第 3 主成分(PC3)对方差的贡献率为 8.1%, 前 2 个主成分的累计方差贡献率为 82.1%。其中, 特征值大于 1 的主成分有 2 个, 因此, 取前两个主成分进行可视化处理。

Table 3. Eigenvalue and variance contribution rate

表 3. 特征值及方差贡献率

主成分	特征值	各因子贡献率/100%	总方差贡献率/100%
1	2.209	57.3	57.3
2	1.409	24.8	82.1
3	0.854	8.1	90.2

### 3.3. 蜂蜜成分的主成分分析可视化处理

不同产地蜂蜜成分指标的前两个主成分的可视化结果如图 1 所示, 蜂蜜成分相似的样品在主成分得分图中聚在较近的区域, 反之, 蜂蜜成分差异越大在主成分得分图中距离越远。

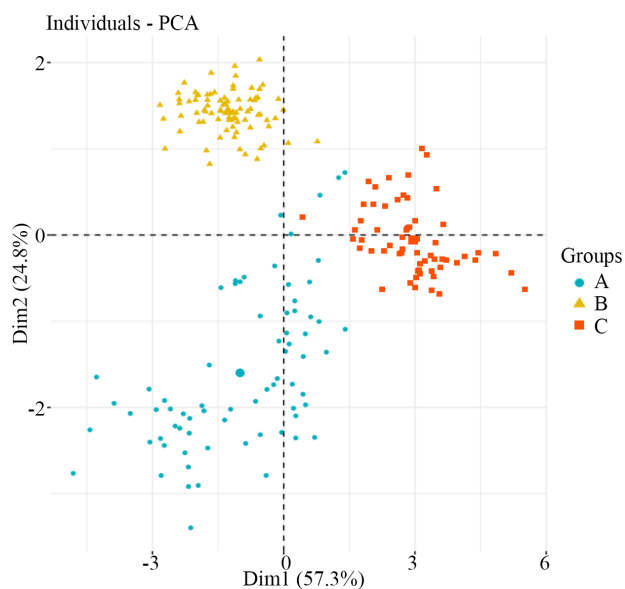


Figure 1. Scatter diagram of principal component analysis of honey composition indicators of different producing area

图 1. 不同产地蜂蜜成分指标的主成分分析散点图

由图 1 可看出, 同一产地的蜂蜜样品能够成簇聚集, B 产地蜂蜜样品与靠近第二轴上端的 C 产地蜂蜜样品完全分离, 两个产地蜂蜜样品分离率达到 100%; B 产地蜂蜜样品与 A 产地蜂蜜样品沿第一轴分离, 分离率达到 100%。A 产地蜂蜜样品与 C 产地蜂蜜样品基本分开, A、C 两个产地按蜂蜜来源基本聚集在不同的区域, 两个产地的蜂蜜样品分离率达到 94%。A 产地蜂蜜样品、B 产地蜂蜜样品和 C 产地蜂蜜样品的指标能够按照蜂蜜产地的不同, 聚集在不同的区域, 反映出不同产地的蜂蜜样品具有差异性, 表明通过对蜂蜜各种成分指标分析能够有效地区分蜂蜜产地来源。

### 3.4. 不同产地蜂蜜各主成分的变量载荷

主成分载荷表示主成分分析中原始变量与主成分之间的相关系数[4] [5], 如表 4 所示, 在第 1 主成分(PC1)中变量因子对其贡献率较大的有葡萄糖、果糖、酸度、淀粉酶, 贡献率分别为 0.432、0.400、0.304、0.264, 表明这些蜂蜜成分的含量可以作为区分不同产地蜂蜜样品的主要依据。这可能是由于不同产地蜂蜜因土壤背景、地质地貌, 植物种类等生长环境的差异所积累的果糖、葡萄糖等成分含量存在差异, 因此可利用这种差异对不同产地蜂蜜进行鉴别。

Table 4. Variable load factor of three principal components

表 4. 三个主成分的变量载荷

蜂蜜成分的指标	PC1	PC2	PC3
果糖(G)	0.400	-0.254	0.141
葡萄糖(P)	0.432	0.078	0.114
蔗糖(Z)	-0.404	0.108	-0.120
麦芽糖(M)	-0.359	-0.287	0.074
果糖和葡萄糖含量之和(GP)	0.437	-0.069	0.132
水分(SF)	0.003	0.623	0.164
电导率(DD)	-0.007	-0.659	-0.009
酸度(SD)	0.304	-0.068	0.193
淀粉酶(DFM)	0.264	0.026	-0.929

## 4. 结论

本文通过对我国东北地区三个地点随机采取的蜂蜜样品的成分测定数据进行主成分分析和单因素方差分析, 发现不同产地间蜂蜜成分指标有显著的差异; 3 个产地间蜂蜜相互分离, 相同产地能够各自聚集, 分离率最高可达到 100%, 达到了很好地鉴别目的, 从而确定了蜂蜜成分指标对产地鉴别的有效性和可行性; 葡萄糖、果糖、酸度、淀粉酶的贡献率较大, 是区分不同产地蜂蜜样品的主要依据。本研究为蜂蜜产地鉴别提供检测和鉴别手段。

## 参考文献

- [1] 中华人民共和国卫生部. 蜂蜜: GB14963-2011 [S]. 北京: 中国标准出版社, 2011.
- [2] 何成. SPSS 中单因素方差分析的运用[J]. 农业网络信息, 2018(1): 92-94.
- [3] 雷鸣, 何晋浙, 孙培龙. 掺假蜂蜜检测技术的研究综述[J]. 食品科技, 2012, 37(7): 283-287.
- [4] 杨天伟, 张霖, 李涛, 王元忠, 刘鸿高. 基于主成分分析和聚类分析的不同产地绒柄牛肝菌红外光谱鉴别研究[J]. 光谱学与光谱分析, 2016, 36(6): 1726-1730.
- [5] 李杨, 李昕光, 刘吉爽, 徐犇, 李纳, 张雅蓉, 陈新. 基于主成分分析对于不同产地黄芩茎叶提取物紫外指纹图谱的鉴别研究[J]. 药学研究, 2018, 37(3): 134-138.