

CHAPTER II

LITERATURE REVIEW

2.1 History of Date Fruits (*Phoenix dactylifera*)

About 4000 years ago, date has been cultivated commercially in Southern part of Iraq. In the ancient world, East and African region had grown large quantities of dates. The botanical name of the palms tree, *Fenix dactylifera*. Fenix derived from the Phoenician supposed "Phoenix", which means palms, and "dactylifera" is derived from the Greek word "dactylos" which means finger (Barrow, 1998). The fruits of palms (*phoenix dactylifera*) are a vital crop in the arid and semi-arid areas of the universe since the dawn of history. It has played a principal role the in economic development social life of the inhabitants in the areas. These forms a fundamental part of their diet daily, so dates are basic food. History is well known about date fruit that have been consumed throughout the world for long periods. It is one the earliest cultivated plants of mankind, that is used for feeding 6000 years ago since the anathematized. The history of one of the most vital fruit trees in all regions of southern Iran (Zohary and Hopf, 2000). **Figure 1** show dates palm.

Figure 1: Dates palm



2. 2 Productions

The main date cultivating nations in the world are Iran, Iraq, Egypt, Saudi Arabia, Pakistan, U.A.E., Algeria, Morocco, Libya, Tunisia, USA, and in minor quantities in Spain, Mexico, Yemen, Israel. FAO report in 2013 revealed that Iraq is the highest producer in the universe. World production of dried grapes was estimated at 1,361,000 metric tons in 2014/2015, which means an increase of 35 percent in the last decade. Turkey, with 328,167 MT in 2014 / 2015, was the country with the largest crop, followed by USA (283,000 MT), Iran (175,000 MT), China (160,000 MT) and India (135,000 MT). Turkey has been the largest exporter of dried grapes in the last ten years. Exports have increased by 18 percent, which reflects the positive trend of dried grapes' production. The United Kingdom and Germany were the top destinations of dried grapes in 2013. Actually, UK and Germany imported the 30 and 17 percent of the total exports from Turkey respectively. On the other hand, Japan and UK were the top destinations of USA dried grapes. The date fruit production in the agro industry in the world produce around five million tonnes of fruit. The fruit of history, which significantly is cultivated in the arid barren Southwest Asian and North Africa the area, and commercialization all over the world as a highly valuable desserts, fruits and crops remains great important subsistence crop in most of the desert areas (Barreveld, 2010). There is reduction in production in the past decade, due to the economic, social, technical and political. This has opened chances for other regions of the world, including the sub-continent south of Africa and subcontinent of Asia.

World output from date fruit from around 760.011 million tons, in 2012. There is an increase from 115.211 million tonnes since 2014 representing about 5 % annual

increase. Commerce numbers suggest that about 93 % of the harvest are consumed domestically in history, and that even now most of these palm species of are not known to export (Zohary & Hopf, 2015). It is just in the last few years that export have been made on the palm as of modern farms in the United States and Israel and in the southern hemisphere (Australia, Namibia and South Africa).

According to Zaid, 2013. The main date-producing countries in the world are Iran, Iraq, Egypt, Saudi Arabia, Pakistan, U.A.E., Algeria, Morocco, Libya, Tunisia, USA, and in small quantities in Spain, Mexico, Yemen, Israel. As can be seen in **Table 1** based on FAO report in 2013 Iraq is the first producer in the world.

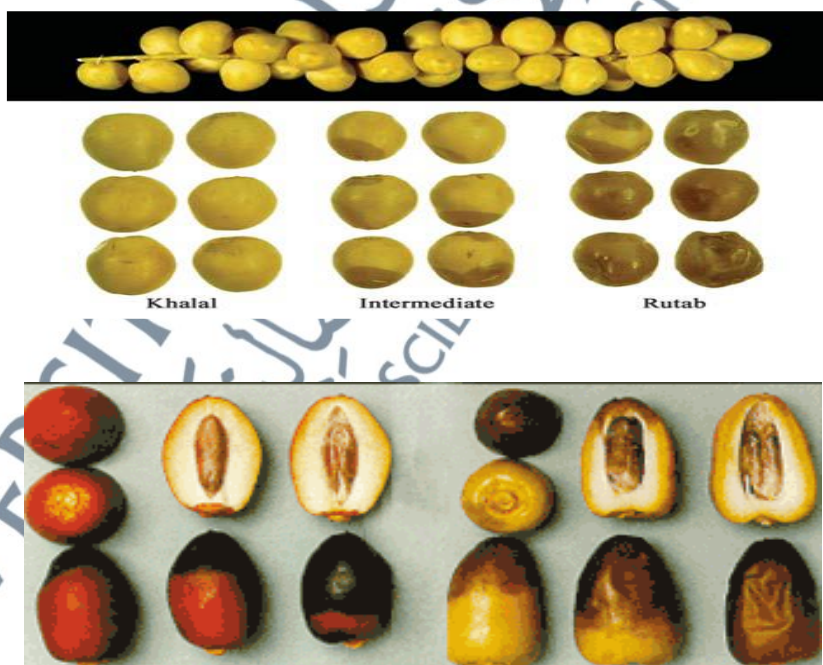
Table 1: Top ten dates producers in 2013 (1000 tonnes) based on FAO report.

Egypt	1502
Iran	1084
Saudi Arabia	1065
Algeria	848
Pakistan	527
Sudan	438
South Sudan	432
Oman	269
UAE	245
World Total	6410

2.3 Stages of Date Ripening

The dates mature in four phases, which is well identified worldwide through the Arab names kimri (immature), Khalal (full-size, crunchy), Rutab which is ripe, soft, and tamar that is mature, and dried. Kimri phase there in rapid increase in size and weight, and reduction of sugars. It is the highest of acid activity and the moisture content (up to 85 %). All the factors level off the end of this phase when the fruit begins to turn yellow (or red according to the wide variety). At that point the date seeds can sprouting and already have mature fruit vegetatively (Kchao et al., 2014) **Figure 2** show stages of date ripening.

Figure 2: Stages of Date Ripening



The procedure for certain varieties of date is developing quickly, making them more pleasant in Khalal phase and one could say commercial maturity of this variety of fruit (Al Shehayeb & Marshall, 2003). With information from the fruit, one can convert and characterized the Rutab phase by a reduction in the weight because of loss of moisture, partial groups (degrees depending on the diversity) in the nature of the sucrose and invert sugar maturation of the skin and lubricate the tissues. The moisture content is reduced to around 35 %, and is sold. The dates at this point are fresh fruit. **Table 2** shows the moisture content of date fruit during its maturation from Khalal to Tamar stage.

Table 2: Moisture content of a date fruit during its maturation from Khalal to Tamar stage.

Stage	Water content (%)
Kimri and Early Khalal	85 %
Late Khalal	50 %
Early Rutab (tip browning)	45 %
50 % Rutab	40 %
100 % Rutab	30 %
Tamar	24 and less

Date ripening is ready for eating when it reaches the value of water 24 while increases value of sugar in dates. This feature enables dates are relatively stable over a long period of time during storage Kevers et al., (2007).

2.3.1 Storage of Date

The Tamar is the harvest season, and certain industries receive the fruit in sums far exceeding immediate ability of markets. Therefore, majority of dates are kept and then sold to the market when they are being demanded for. The industries that produces

date fruits typically stockpile dates at 30 °C in the period of up to a year. After packaging and releasing to the market, the date fruit is expected to have a shelf life of for up to 2 years at room temperature (25 °C) and fruits. Al-Redhaimani (2005) recognized the impact of storage on dates, the author concluded that the storage period for 'Barhi' date studied in his research could be significantly improved by exposure to elevated CO₂ concentrations up to 20 % CO₂. Consequently, it is imperative to study the impact of the current methods of storage on the quality of various the physical attributes of date fruit. Thus, the development of ideal storage conditions suitable for the fruit is important. Fruit of date received a little interest, their standing in Iran and several other countries and other areas are recognized

2.4 Chemical Composition of Date Fruit

Dates are one perfect food and vital component of nutrition with vitamins and minerals with the most percentages. There are 25 % more potassium in date fruit than bananas, while fat, sodium, and cholesterol content is negligible. These components in date fruit show important role in the diet and treatment of obesity and is the most vital basis of energy and food in the producing nations in the fruit.

According to diversity of date fruit, the growth conditions (pass) differ in their shape, size and weight and they are usually the rectangular, they although have certain varieties up to nearby round shape. The dimensions may differ from 18 and 8 to 60 mm and 32 mm, but the averages at 40 mm and 20 mm consecutively. The mean weights in the fruit is around 7 to 10 grams (El-Sohaimy & Hafez., 2010). Studies for the characterization of values of biochemistry feeding from the dates fruits of

Egyptian history components. The date palms containing a variety of complex vitamins B1, B2, nicotinic acid and vitamin A. Except of date palm that contains a 13.80 % moisture and a total of 86.50 % solid. The ash and fiber contents of the coagulate 2.13 % and 5.20 % straight. Protein, carbohydrate and fat contents are 3.00 %, 73.00 % and 2.90 % respectively. Decline in the level of fatty content 2.90 % compared with the polysaccharides content means that Palm is safe for patients with high blood pressure because it contains very low levels of cholesterol and fatty acids in the blood level El-Sohaimy & Hafez, (2010).

High performance liquid chromatography (HPLC) analysis has shown that the carbohydrate content in date consists of a large amount of glucose, fructose and sucrose. The palm contains high level of aspartic acid, proline, glycine, amino acid, valine, lysine and arginine, but low concentrations of a threonine, serine, methionine, alaizuliossin, tyrosine, phenylalanine and lysine (Fadel et al., 2001). The date palm is uniquely useful and food properties of its kind has numerous health benefits beyond its nutritional value. Food associated with consuming the palm is used to enrich the food properties for various types of food.

2.4.1 Sugars

The sugars contribute more widespread in the old countries and the in the production stage, single component was used more as a source of the date sugar fruit. All the component of sugars in the dates are from a mixtures of sucrose ($C_{12}H_{22}O_{11}$), glucose ($C_6H_{12}O_6$) and fructose ($C_6H_{12}O_6$). Glucose and fructose are the derivatives of sucrose after crystallisation process (Salman et al., 2013). Date is determined by relative

quantities of the sucrose, glucose, fructose, mainly through the characteristics of varieties but we can say that most of the dates are invert sugar kind, at any phase that are eaten, more so, when the dates fruits are inverted not all sucrose are converted into glucose and fructose by invertase enzyme.

Chemical analysis was conducted on the fruits on twelve varieties of palm consumed on a large scale in the United Arab Emirates. It showed that glucose and fructose increase rapidly with increasing maturity in the Kimri through Khalal and Rutab where it is ready to be eaten. The total sugar could represent more than 50 % of the fresh weight in the date palm, the sugar values, along with the contents of the low moisture promotes resistance to the fungal damage after harvest (Mohamed et al., 2014). Metals accumulated in the date fruits could be an important source of potassium in regular consumers. The sugars (55-82 % dry) and materials, with glucose and fructose as predominant sugars (Mohamed et al., 2014).

2.4.2 Crude Fibres

It is usually prescribed that crude fiber with insoluble segment of the flesh date which is mainly consist of cellulose, hemicellulose, lignins as well as proteins and insoluble crude fibre. Throughout the maturation stages, it split these materials progressively to the bottom of enzymes for more mildly-soluble compounds to make the fruit more gentle and soft. Mature commercial dates have crude fiber between 2-6 %, but in low-quality dates for manufacturing use this proportion will reach 10 % (Elleuch et al., 2011).

Two classes of date palm Deglet Noor and Allig of Tunisia were chosen by Elleuch et al. (2011) to analyse the chemical component and dietary fiber (DF) respectively. Table 3 shown the results in dry matter basis (DM) and dietary fiber (DF). Ahmed et al. (2013) represented in their studied that insoluble DF, and a large part of the total dietary fiber (DF), is a 9.19-11.7 % Dry Matter for Deglet Noor and Allig. The chemical composition of these concentrates showed great contents of total DF (88 % - 92.4 % DM) and protein and low ash (8.98 - 9.12 % DM) and (2.0 -2.1 % DM, respectively). DF concentrates showed a high capacity to retain water (15.5 g water / g sample) and oil retaining capacity (9.7 g oil / g sample).

Table 3: Dry Matter Basis (DM) and Total Dietary Fiber of Date Palm Deglet Noor

Chemical Component	Dry Matter Basis (DM)	Dietary Fiber (DF)	(DF)
sucrose	13.9 %	ash	2.5 % - 2.52 %
glucose	13.7 %	protein	2.1 % - 3 %
fructose	12.6 %	Total Dietary Fiber	14.4 % - 18.4 %

2.4.3 Vitamins and Minerals

The dates have many vitamins which comprise are A, C, B1 and B2 as shown in **Table 4** and minerals are niacin, sodium, potassium, magnesium, calcium and iron. Besides, chlorine, copper, sulfur and phosphorus also exist in the fruit.

Table 4: Vitamin Content of Date

Vitamin	% 100 / gm
Vitamin A	4.8 – 6
Vitamin C	0.77 - 2.7
Vitamin B1	0.07 - 0.1
Vitamin B2	0.03 - 0.05
Vitamin B3	0.33 - 2.2

Salman et al. (2013). had studied the metallic elements in dates such as sodium, magnesium, potassium and calcium using inductively coupled plasma (ICP) atomic spectroscopy in 34 major palm varieties that had been collected from all areas of cultivation in Iran. The result showed that the concentrations (mg / 100 g, dry weight basis) of sodium, magnesium, potassium and calcium in the range of 4.46 - 47.74, 18.44 - 79.35, 203.61 - 982.97 mg and from 23.24 - 73.85, respectively. Sodium content is less than the other results reported in Pakistan, the United Arab Emirates and the United States, China, Oman and Tunisia. Differences might arise from agricultural, climatic as well as environmental conditions.

Date palm fruits are made up of a variety of vitamins A, complex B, B1, B2 and nicotinic acid. Date palm extracts also contain moisture of about 13.80 % and a total solid of 86.50 %, 2.13 % ash, 5.20 % fiber contents, 3.00 % protein, 73.00 % carbohydrate and 2.90 % fat content. It has a low fat content of about 2.90 % as equated to the content of its sugar (Salman et al. 2013). This shows that date fruit is harmless for patients with blood and heart diseases as a result of it has very low levels of fatty acid and cholesterol. The fruits of the palm have a unique functional and nutritional values of its kind. Many health benefits have been associated with even further value

with nutritional consumption of palm fruits to have high nutritional values for various types of food.

2.4.4 Enzymes

Enzymes in crucial role in the transformation processes that takes place during creation and development of the date fruit, and the activities of four of them are of particular interest to the final product quality:

- i. Invertase: accountable for the inversion of sucrose into glucose and fructose and connected to texture and pliability.
- ii. Polygalacturonase and pectinesterase: both convert insoluble pectic materials into more soluble pectins, which contributes to softness of the fruit.
- iii. Cellulase: breakdown cellulose into smaller chain materials with increasing solubility and finally leading to glucose, thus reducing fibre content.
- iv. Polyphenol oxidase is accountable for biochemical changes of polyphenols to which the tannins fit in; they are vital in non-oxidative browning reactions of the date.

Understanding the roles and action of enzymes are significant for proper operation of heat and humidity. The enzyme action can be stimulated or depressed with regards to the anticipated result. Enzyme activity usually occur in solution or humid environment; the optimal temperature condition typically falls between 30 °C and 40 °C, over and lower which the action will reduce (for instance invertase at 50 °C, loses 50 % activity and 90 % at 65 °C after 10 minutes) (Behija et al., 2011).

Extended storage of dates under low temperatures is guided chiefly on the reducing actions of enzyme. Besbes et al. (2004) studied the consequence of heat treatment (55 °C / 20 min) on polyphenol oxidase (PPO) and peroxidase (POD) actions and total phenolic compounds in Algerian dates (Deglet Nour variety) at Tamar (fully ripe) stage and the dates were kept for 5 months at ambient temperature and in cold storage (10 °C). There was a reduction in Deglet Nour dates for both POD and PPO activities throughout storage for either heat treated and non-treated dates samples.

The second leading industry that relies on using microbes in fermentation production as second to pharmaceuticals is the enzyme industry. The production industrial enzymes such as proteases, amylases, cellulases, lipases, xylanases, invertases, pectinases etc. that are commercially important depends mainly on economical enzyme production made out of simple and low-cost substrates for the economic manufacture of enzymes. Whereas substrates like soya bean, bean meal, sugarcane molasses and other agro manufacturing wastes are investigated, the wastes from date fruit are yet to be exploited though limited enzymes have been experimented. The by-products from date fruit and wastes are also investigated as complementary substrates in the enzyme production medium for few industrial enzymes by investigators. From the literature available it is inferred that date fruits could be successfully used in the production medium as substrate for the production of pectinases, endopectinases and alpha amylase by few microorganisms (Chandrasekaran & Bahkali, 2012).

2.4.5 Antioxidants, Phenolics and Flavonoids

The activities of total phenols and antioxidant in both aqueous and alcoholic extracts of three best quality date varieties (Khalas, Sukkari and Ajwa) from Saudi Arabia were studied by Abdullah et al. (2011). Generally, the water extract has revealed significantly higher contents of total phenols than alcoholic. The phenolic profile indicated that Sukkari has the most rutin concentration (8.10 mg / kg), but, catechin was nearly the same in Sukkari and Ajwa (7.50 and 7.30 mg / kg, respectively). The caffeic acid of 7.40 mg / kg was the highest in Khalas variety.

Sukkari and Ajwa had oxidation value 0.63, 0.70 mg / mL compared with 1.60, 1.43 mg / mL in alcoholic extract. Additionally the linear correlation between total phenols in water was high and positive ($r = 0.96$) and alcohol ($r = 0.85$) extracts and inhibition of lipid oxidation activity. There was also high positive correlation between catechin ($r = 0.96$), and rutin ($r = 0.74$) in water extract, but this correlation reduced in alcoholic extract ($r = 0.66$) for catechin and very weak ($r = 0.38$) for rutin. There was no correlation between caffeic acid and lipid peroxidation in the water and alcoholic extracts.

The free radical scavenging activity of ethanol extract using dual-mode 1,1-diphenylpicrylhydrazyl (DPPH) had been identified (Abdel-Zaher, 2010). Methanol extract (0.1 mL) was added to 0.9 mL of fresh solution DPPH methanol (0.1 mm) and an equal amount of methanol was used as a control. After incubation for 30 minutes at room temperature in the dark, measurement of the optical density (OD) was analyzed

at 517 nm using a spectrophotometer. The account of scavenging (%) was calculated using the following formula:

$$\text{DPPH radical scavenging \%} = [(OD \text{ sample} - OD) / \text{Control OD}] \times 100 \text{ Eq (1)}$$

DPPH test also show no significant correlation, except with Sukkari where there was no difference between aqueous and alcoholic extracts (4.30, 4.10 mg / ml respectively) found by Abdullah et al. (2011).

Returns extraction analysis showed that the active principle associated with DPPH radical-scavenging effect was most. It was determined to focus on the total flavonoids using the modified colorimetric method described by (Zhishen et al., 1999) where catechin was used as the standard.

The use of chemical procedures to detect the presence of total phenols, while using the optical spectrum and chromatographic techniques to identify and quantify individual phenolic compounds. This deals with the application of different methodologies used in the analysis of phenolic compounds in plant-based products, including new technical developments in the quantification of the phenolic acid phenol and phytochemicals are found in all known plants (Khoddami et al., 2013).

The date consists of simple phenols, benzoic and cinnamic acid, coumarin, tannins, lignins, lignans and flavonoids. Organic solvent extraction is the main method used to extract the phenol. The use of chemical procedures to detect the presence of total phenols, while using the optical spectrum and chromatographic techniques to identify and quantify individual phenolic compounds (Khoddami et al., 2013).

2.5 Extraction by Mixture Solvents

The polarity of different solvents is shown in **Table 5** (Organic Chemistry, Wiley-VCH Publishers, 3rd ed., 2013).

Solvent	formula	polarity
Acetic acid	C ₂ H ₄ O ₂	0.648
Acetone	C ₃ H ₆ O	0.355
Acetonitrile	C ₂ H ₃ N	0.460
Acetyl acetone	C ₅ H ₈ O ₂	0.571
2-aminoethanol	C ₂ H ₇ NO	0.651
Aniline	C ₆ H ₇ N	0.420
Anisole	C ₇ H ₈ O	0.198
Benzene	C ₆ H ₆	0.111
Benzonitrile	C ₇ H ₅ N	0.333
Benzyl alcohol	C ₇ H ₈ O	0.608
Chloroform	CHCl ₃	0.259
Dimethylsulfoxide (DMSO)	C ₂ H ₆ OS	0.444
Dioxane	C ₄ H ₈ O ₂	0.164
Ethanol	C ₂ H ₆ O	0.654
Ether	C ₄ H ₁₀ O	0.117
Ethyl acetate	C ₄ H ₈ O ₂	0.228
Hexane	C ₆ H ₁₄	0.009
Methanol	CH ₄ O	0.762
Water	H ₂ O	1.000
Water, heavy	D ₂ O	0.991
<i>p</i> -xylene	C ₈ H ₁₀	0.074

Al-Farsi et al. (2007) conducted the extraction of phenolics and dietary fibre from date seeds and in their research, acetone 50 % and butanone were the most active solvents for extraction. The phenolic contents of seed concentrate to 18.10% and 36.26%, respectively. The total dietary fibre of seeds (57.87 g /100 g) are larger after water and acetone extractions to 83.50 and 82.17 g/100 g, respectively. Nine phenolic

acids (free and liberated) were detected in seeds with *p*-hydroxybenzoic (9.89 mg/100 g), protocatechuic (8.84 mg/100 g), and *m*-coumaric (8.42 mg/100 g) acids found to be among the highest, the total phenolic acid content improved considerably after extraction and purification from 48.64 to 193.83 mg/100 g. The major phenolic acids found in the concentrates were protocatechuic, caffeic and ferulic acids.

In the research of Haider et al. (2013), second-grade dates with hard texture, from six Tunisian cultivars were (Bejo, Baydh, El Hamam, Alpaju, Heish and Alkhseba) analysed for their composition and antioxidant activities. The highest level of phenolic compounds for all dates varieties was detected in the 70 % acetone extract. The level varied from 199.43 to 576.48 mg of / 100 g fresh weight. However, extraction into 50 % methanol produced the maximum antioxidant activity from 89.55 to 109.67 mg equivalents of ascorbic acid/g fresh weight. For both solvents, Bejo showed the highest phenolic content and total antioxidant activity, whereas Baydh and El Hamam showed the lowest. A positive linear correlation between total antioxidant activity and phenolic contents was observed ($R^2 = 0.83$; $R^2 = 0.74$ for acetone / H₂O and methanol / H₂O, respectively).

The efficiencies of 80 °C water, 40 % ethanol and 80 % ethanol in the extraction of phenolics from black currant leaves, as well as the antioxidant capacity of the obtained extracts, were investigated by Noura et al. (2014). Aqueous ethanol (40 %) was found the most effective in the extraction of phenolics followed by 80 °C water while the antioxidant capacity of the investigated extracts correlated with their phenolic content. Seven phenolic acids (gallic, chlorogenic, caffeic, *p*-coumaric,

ferulic, sinapic and salicylic) and three flavonoids (rutin, myricetin and quercetin) were identified and quantified using HPLC with PDA detector. Noura et al. (2014).

In a research conducted by Yen Yuan and Chun Chou (2012), a solid state fermentation of black soybeans with *Bacillus subtilis* BCRC14715 was executed. The consequence of fermentation on the changes in flavonoid content, total phenolic and antioxidant activities including DPPH radical- scavenging effect and Fe^{2+} chelating ability exerted by various solvent (water, 80 % methanol, 80 % ethanol, 80 % acetone) extracts of black soybeans was studied. It was established that fermentation improved the flavonoid content and total phenolic along with antioxidant activity of the black soybean extract. Out of the different extracts studied, the acetone extract of fermented black soybeans displayed the highest flavonoid content and total phenolic.

According to Yen Juan and Chun Chou (2012), the extraction of phenolic compounds after sample preparation is a critical step where several factors may impact the yield of phenolics such as temperature, extraction time, the number of retelling extractions of the sample, solvent type as well as solvent-to-sample ratio. Additionally, the optimal recapture of phenolics is different from one sample to the other and depend on the kind of plant and its active compounds. The selection of solvents extraction such as acetone, water, alcohols (methanol, ethanol and propanol), ethyl acetate and their mixtures will impact the yields of phenolics extracted. For example, a high yield of phenolics can be extracted from sorghum leaf using water whereas extraction of phenolics from wheat bran needs 80 % aqueous ethanol. In another example, an inquiry into the influence of different solvents on extraction of phenolics from aerial parts of *Potentilla atrosanguinea* indicated that 50 % aqueous

ethanol was more effective than pure or 50 % aqueous forms of methanol, and acetone.

According to Khoddami et al. (2013), flavonoids are extremely bioactive compounds found in both edible and non-edible plants. They are regularly extracted with ethanol, methanol, acetone, water or mixtures of these solvents using heated reflux extraction methods. Succeeding extraction, the flavonoid glycosides are often hydrolyzed into the glycone forms by applying HCl under N₂. Extracted flavonoids from diverse types of herbal plant materials with 50 % methanol acidified by 1.2 ml HCl. Ascorbic acid was added to prevent oxidation of the mixture. The hydrolysis of the flavonoid glycosides was carried out for 2 h at 80 °C.

Kchaou et al. (2012), focused on optimization of enzymatic extraction of flavonoids from celery stalks. The pulpy aqueous homogenate was mixed with 1 N HCl or NaOH to adjust the pH and the mixture incubated at the desired temperature. A complex mixture of enzymes was added to the sample under stirring at 150 rpm. The enzymes were then inactivated by heating at 90 °C for 10 min and the supernatant of the centrifuged mixture was collected for total flavonoid determination.

Biesaga extracted flavonoids in maize samples using heated reflux, microwave-assisted extraction (MAE), ultrasonic-assisted extraction (UAE) and maceration and compared the stability of the extracted compounds. The highest stability of the extracted flavonoids in methanol-water (60:40 v/v) was for compounds extracted with traditional heated reflux in a water bath and MAE within 1 min under 160 W (Khoddami et al., 2013).

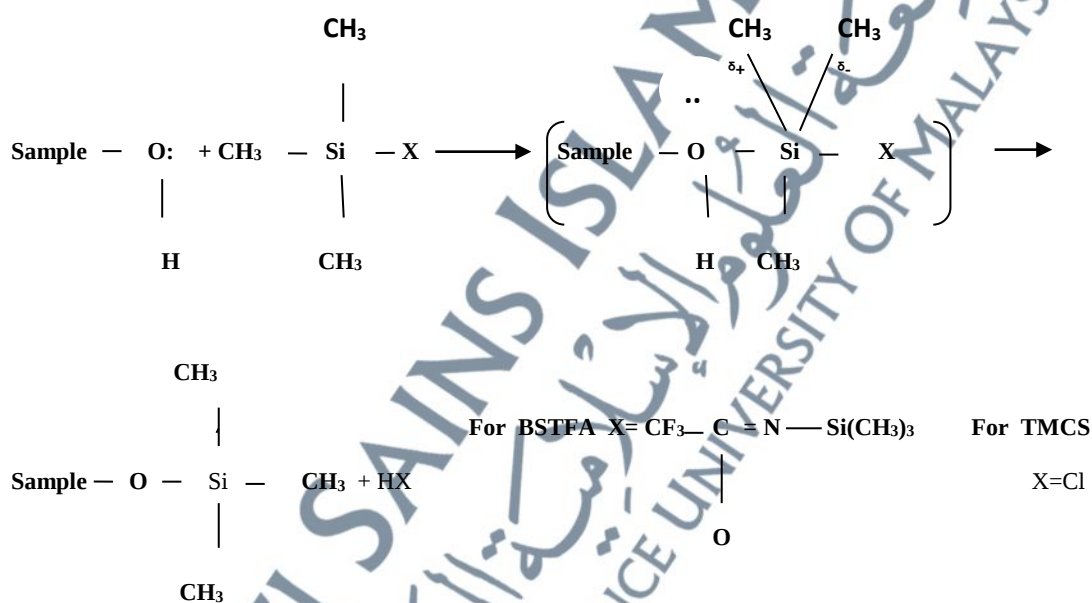
Extraction of phenolic compounds with sugars was carried out by introduced 25 grams of ground peas meal into a 100 ml dark glass bottle and suspended in 200 ml of methanol-water, ethanol-water or acetone-water (80:20, v / v). Tightly capped bottles placed in water bath at 80 °C. After 15 min during which the content was shaken twice, the extract was cooled and filtered under partial vacuum. The material left on the filter paper was transferred back to dark glass bottles for further extraction with 200 ml of the same extraction solution. This procedure was repeated three times over 30, 60 and 90 min of extraction, each time collecting the solution for analysis. Supernatants were combined and evaporated using rotary vacuum evaporator to remove any remaining solvent; the water was then removed by lyophilization (Chavan & Amarowicz, 2013).

2.6 Derivatization Using BSTFA

The *N*, *O* bis (trimethylsilyl) trifluoroacetamide (BSTFA) is very versatile, reacting with a range of polar organic compounds and replacing active hydrogens with a – Si(CH₃)₃ (trimethylsilyl) group. It reacts rapidly and more completely than *N*,*O*-bis(trimethylsilyl)acetamide (BSA). The trimethylsilane (TMS) derivatives are thermally stable but more susceptible to hydrolysis than their parent compounds. BSTFA and its by-products (trimethylsilyltrifluoroacetamide and trifluoroacetamide) are more volatile than many other silylating reagents, causing less chromatographic interference.

Hydrogen fluoride, a by-product of silylation with BSTFA (see Mechanism), reduces detector (FID) fouling. The trimethylchlorosilane (TMCS) increases the

reactivity of BSTFA (or other silylation reagents). Amides and many secondary amines and hindered hydroxyls, incompletely derivatized by BSTFA alone, can be derivatized by adding 1-20 % TMCS to BSTFA. A mixture of BSTFA and TMCS has good solvent properties and can function as a silylation reagent without additional solvents. Alternatively, the mixture is very soluble in most commonly used silylation solvents.



The BSTFA joined with TMCS is ideal initially when engaged to support trimethylsilylation of alcohols, amines, carboxylic acids, among others. Being a substitute for the analysis of low volatile compounds by gas chromatography, the combination of these compounds favors the replacement of the amine and

phosphonate groups which may be found in the structures of glycine (GLY) and α -amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid (AMPA).

The derivatization step of BSTFA was performed whereby the sample solution (200 μ L) was placed in a vial (1 mL) and evaporated to dryness under nitrogen at 60 °C. Silylating agent (BSTFA containing 1 % TMCS; 100 μ L) was added to the residue and the vial was vortex mixed and heated at 80 °C for 30 min. After cooling, the derivatized solution was evaporated to dryness and the residue was redissolved in 100 μ L chloroform. This solution (1 μ L) was analyzed by GC–MS (Szyrwińska et al., 2007).

TMCS acts as a catalyst, increasing the power of the donor silyl (BSTFA) and assuring greater competence for the reaction. However, few reports have been found in the literature regarding the derivatization of glycine (GLY) and α -amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid (AMPA) using this combination of reagents (Farag et al., 2014). The study on the use of BSTFA and N-(tert-butyltrimethylsilyl)-N-methyltrifluoroacetamide (MTBSTFA) to derivatise the natural hormone estrone (E1) and the synthetic estrogen 17-ethinylestradiol (EE2). The resulting trimethylsilyl (TMS) and *t*-butyldimethylsilyl (TBS) derivatives of EE2 were incompletely transformed to their respective E1 derivatives. Therefore, these reagents may not be appropriate for instantaneous determination of estrogens in environmental samples, which raises questions about the reliability of results from some earlier studies (Shareef et al., 2004).

2.7 Evaporation Techniques

In extraction process, the solvent must be evaporated to concentrate the extract. The evaporation technique is performed in different ways; such techniques include stationary evaporation and rotational type evaporation. The rotational type evaporation is the most popular process utilized to evaporate a solvent from an extract (Nkambule, 2008).

2.7.1 Rotary Evaporator

Rotary evaporator is an important evaporation unit combined with a rotating evaporation flask that is developed by Buchi. Vacuum evaporation is mostly used due to its effectiveness in removing the solvent. In addition, it is faster than the stationary evaporation flasks and also can produce a very effective heat transfer, at the same time, confirming fully mixing and avoiding overheating of the extract (Nkambule, 2008). Rotation movement transfers a liquid sample as a thin film to the entire flask surface, noticeably increasing evaporation speed and helping transfer of heat from a heating source. Moreover, sealing system, which is rotating flask and vapor vessel are combined by, assists the operation to be under vacuum. Additional increase in the speed of the evaporation process due to a decrease in boiling point of the solvent and effective removal of the vapor phase. Also, vacuum operation allows heat-labile constituents to be effectively concentrated without degradation (Buchi, 2008). Zhang et al. (2014) employed quadratic rotary evaporator to determine optimal extraction process in Areca seed.

According to Mohd et al., (2011), fifty gram of citrifolia leaf powder was accurately weighed and placed in extraction vessel. Carbon dioxide and ethanol were pumped into the extraction vessel at a flow rate of 15 and 2 g / min, respectively. The critical temperature and pressure employed were 50 °C for more than 3 hrs with the collector set at 30 °C. The solvent was removed under vacuum using rotary evaporator at 40 °C. The results from this study provided useful data regarding the effect of different extraction techniques on the recovery of phenolics and antioxidant activity of extractable components from citrifolia leaf. It can be suggested that temperature plays a significant role in preserving the antioxidant property of citrifolia leaf extract during extraction.

According to Hossain & Nagooru (2011), studied Neem leaf powders (127 g) were extracted with methanol solvent (500 mL, 72 h) by using Soxhlet extractor. After extraction, it was filtered and the methanol solvent was evaporated completely by using rotary evaporator (Yamato Rotary Evaporator, Model RE 801). The solvent free methanol crude extract (9.6 g) was suspended in distilled water (100 mL). The suspension was transferred into a separatory funnel. Then it was extracted successively with different organic solvent with increasing polarity such as hexane, chloroform, ethyl acetate and butanol resulting: hexane (1.327 g), ethyl acetate (4.425 g), chloroform (1.00 g), and butanol (1.20 g) and residual methanol fractions (2.09 g), respectively. All the crude extracts were filtered using filter paper (Whatman No. 41) to obtain particle-free crude extract. The defatted extraction procedure was repeated twice for all solvent for complete extraction and then filtered. The combined extracts were concentrated and evaporated by using rotary evaporator and dried under vacuum.

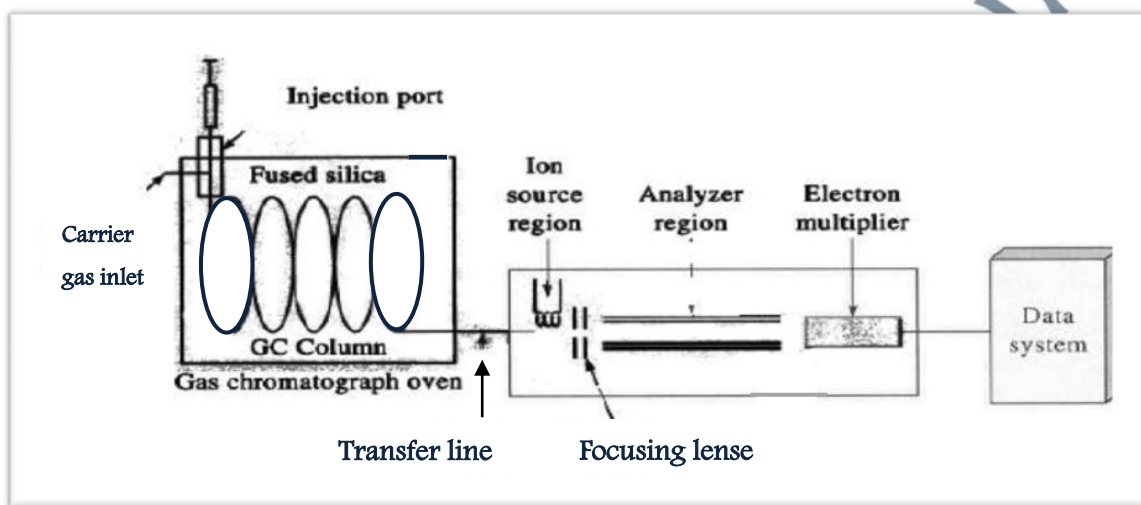
2.8 GC-MS Technique

Chromatography technology is widely utilized for separation, isolation and purification. This technique was established in 1906 by a Russian botanist, Mikhail Tswett, who separated various color components from extract of crushed leaf by passing it through a 13 column consisting alumina, sucrose and calcium carbonate. Consequently, it was called chromagraphein, which is a Greek word, where chroma means “color” and graphein means “to write” (Pedersen & Myers, 2010; Wixom, 2001). It is a physical process of separation where, the separated components are distributed between two phases. Those two phases are mobile phase and stationary phase (Ahuja, 2003). Silica, alumina or C-18 powders are the most used stationary phases, whereas the mobile phase is gas for gas Chromatography.

GC-MS instruments separate chemical mixtures and identify the compounds at a molecular level. In GC, a mixture would be separated into different substances once heated. The heated gases are passed within a column by an inactive gas like helium. The substances flow into the MS, when they are separated from the column. Mass spectrometry detects compounds based on the molecule mass of the analyses as an electrical signal. That signal will be sent to a data system, which creates an image termed a chromatogram presenting the peaks of analyses (McNair & Miller, 2011).

Fig. 3 shows schematic of a typical capillary GC-MS.

Figure 3: Schematic of a typical capillary GC-MS (McNair & Miller, 2011).



The mobile phase carries the analyses through the stationary phase. The analyzer ideally equilibrate or differentially partition among the two phases, leading to a diverse migration rate through the system. Two type of chromatography are liquid chromatography and gas chromatography (Villas-Boas et al., 2007). The phytoconstituent measurement is typically attained by gas (GC) and liquid chromatography (LC) linked with specific detection schemes (Eisenhauer et al., 2009).

The principle of gas chromatography is to separate compounds according to their polarity difference or boiling points in moving gas. That means a mixture of compounds is separated based on their affinity of interaction with the stationary phase filled into the capillary column of gas chromatography. Gas chromatography is very sensitive and offers quantitative and qualitative results in a single process (Harborne, 1998).

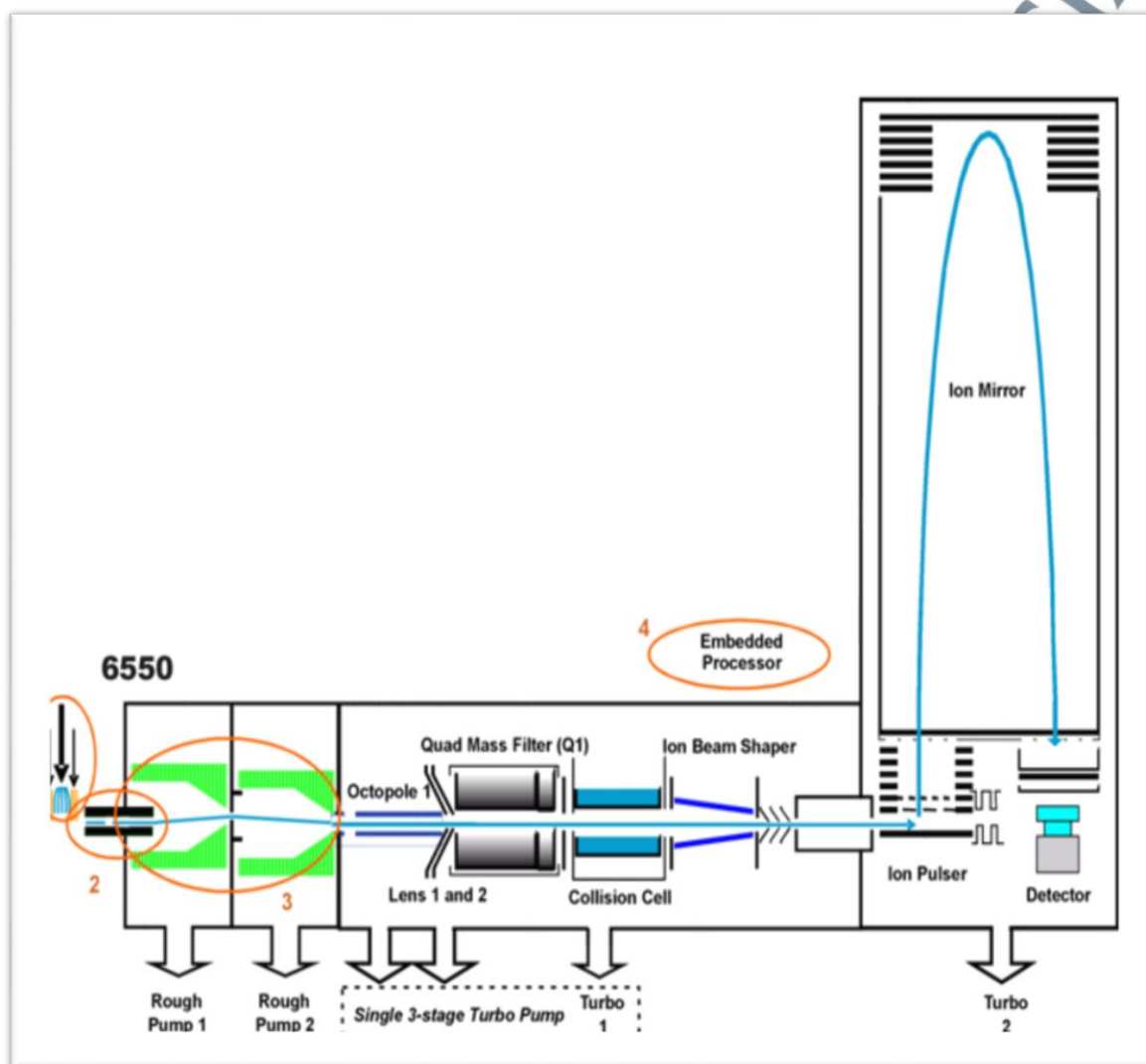
Mass spectroscopy (MS) provides important information about the molecular weight and mass fragmentation pattern in explaining a structure of a compound. It can be used to analyze different types of components and provides information regarding the fundamental composition of a sample, inorganic and organic structure, biological molecules, quantitative and qualitative composition of composite structure, solid surface composition and atom isotopic ratio in sample (Selvamangai & Bhaskar, 2012). In many researches, (GC-MS) has been proven as an important scientific metabolic profiling in plant and non-plant species (Goncalves et al., 2006).

2.9 LC-TOF-MS

LC-TOF-MS is the abbreviation for liquid chromatography Time of Flight Mass Spectrometry. Charged ions of various sizes are generated on the sample slide, as shown in the diagram. A potential difference V_0 between the sample slide and ground attracts the ions in the direction shown in the diagram. The velocity of the attracted ions v is determined by the law of conservation of energy. As the potential difference V_0 is constant with respect to all ions, ions with smaller m/z value (lighter ions) and more highly charged ions move faster through the drift space until they reach the detector. Consequently, the time of ion flight differs according to the mass-to-charge ratio (m/z) value of the ion. The method of mass spectrometry that exploits this phenomenon is called Time of Flight Mass Spectrometry (Zhang et al., 2014).

The main parts in LC-TOF-MS are ion source, ion transport, analyzer, detector, data acquisition system and data system (PC) **Fig. 4** shows Schematic of Agilent 6550 TOF LC / MS with major improvements circled.

Figure 4: Schematic of Agilent 6550 TOF LC / MS



Recent technological advances in high-resolution and high-throughput methods generally termed as metabolomics, allow for large scale GC-MS profiling regarding the high number of measured metabolites and experiments carried out (Lisec et al., 2006). Gas chromatography time-of-flight mass spectrometry (GC-TOF-MS) provides fast scanning, high sensitivity and mass accuracy compared to common quadrupole (GC-QMS) or ion trap instrumentation (GC-ITMS), and is considered the standard GC

platform in many metabolomics labs. Sample processing for GC-MS-based metabolite profiling include solvent extraction, concentration to dryness and consecutive derivatization, often carried out in a two-step procedure.

In the first step, methoximation (also called methoxyamination) is achieved by a reaction of sample components with e.g., O-methoxylamine hydrochloride diluted in pyridine to stabilize thermolabile enolic aldehydes and ketones and to convert them into oximes or alkyl oximes. In the second step, extracted metabolites are derivatized with silylating reagents (Lisec et al., 2006). The latter step is crucial for the adequate derivatization of non-volatile compounds, in order to capture a huge variety of metabolites with polar characteristics and high boiling points on a GC-MS system. Detectable compounds comprise sugars (mono-, di- and trisaccharides), sugar alcohols / acids, amino and fatty acids, phosphorylated intermediates and many plant secondary metabolites such as phenolics, terpenoids, steroids and alkaloids.

In the study of Ferrer and Thurman (2007), they focused do on the importance of unequivocally detecting emerging contaminants, as well as establishing their presence in the environment by accurate mass spectrometric measurement techniques. A novel approach to studying the production of secondary metabolites by fungi using LC / TOF-MS has been developed. Fungi grown on culture media are solvent-extracted and directly analyzed by LC / TOF-MS. Searching against a database of 465 secondary metabolites, mycotoxins and other compounds of interest can be readily identified.

The methodology was validated by spiking culture media with 20 mycotoxin standards and identifying these toxins in the crude solvent extracts. Subsequently, using seven different fungi from culture collections, after culturing for 7 to 14 days in three different media, anticipated metabolites were readily identified. The review

explained the unique features of these instruments and gives examples of their applications (Ferrer & Thurman, 2010). The use of liquid chromatography / mass spectrometry (LC-MS) methods for protein analysis is well acknowledged and has had a significant impact on understanding the complexities of biological systems. Less well publicized is the role that mass spectrometry has had in the characterization of expressed proteins in molecular biology laboratories.

LC-MS is one of the most current technologies for protein identification in both simple and complex samples. This application note will demonstrate the use of liquid chromatography and high-mass-accuracy time-of-flight mass spectrometry (LC-TOF-MS) for the characterization of expressed proteins.

The limits of detection using LC-TOF-MS for mycotoxins were less than 10 ngg^{-1} in foodstuff analysed in this work. The LC / TOF-MS was appropriate for the screening of mycotoxins in food stuff quickly with high sensitivity, and its performance was demonstrated for the confirmation for target mycotoxins Xiaolin and Bruce (2010).

2.10 Principal Component Analysis (PCA)

PCA is among the most popular multivariate statistical technique and it is widely used by almost all scientific disciplines. It is also likely to be the oldest multivariate technique, as its origin can be traced back to Pearson, but its modern instantiation was formalized by Hotelling who also coined the term “principal component” (Abdi & Williams, 2010).

Principal component analysis is defined as a method of data reduction to elucidate the relationships between two or more characters and to divide the total variance of the original characters into a limited numbers of uncorrelated new variables. This will appropriate visualization of the differences among the individuals and identify possible groups. The reduction is accomplished by linear transformation of the original variables into a new set of uncorrelated variables known as principal components (PCs).

According to Jolliffe (2002), the central idea of PCA is to reduce the dimensionality of a data set consisting of a large number of interrelated variables, while retaining as much as possible of the variation present in the data set. This is achieved by transforming to a new set of variables, the principal components (PCs) which are uncorrelated, and which are ordered so that the first few components retain most of the variation present in all the original variables.

Abdi (2003), stated that the goal of PCA is to decompose a data table with correlated measurements into a new set of uncorrelated (i.e., orthogonal) variables. These variables are called (depending upon the context) principal components, factors, eigenvectors, singular vectors, or loadings. Each unit is also assigned a set of scores which correspond to its projection on the components. The results of the analysis are often presented with graphs plotting the projections of the units onto the components, and the loadings of the variables. The importance of each component is expressed by the variance (i.e., eigenvalue) of its projections or by the proportion of the variance explained. The most interesting information gained from PCA is the scores and

loadings. As already stated, the distance of samples (points) from a particular PC is called the scores on the respective PC.

The scores therefore, hold information on the samples. Samples with similar scores are close to each other in the row space and are therefore somehow similar. The goal of PCA as known is to summarize and visualize multivariate data and this visualization is given in the form scores plots. The loadings on the other hand hold information on the variables (Wise et al., 2006) and are a measure of alignment of the particular PC with the variable axes. Variables with high loadings on a particular PC are well described by that PC. The loadings plots make it possible to find out why some samples are being similar and forming clusters in the scores plots, and this is a great advantage of PCA over clustering.

PCA is also a way to sum up the original data with a large amount of variables into simpler data with only a few variables (Wise et al., 2006). Similar to clustering, PCA is often used to find patterns in data to determine if some samples differ from others, observed as clusters of samples located at a distance from other samples in plots called scores plots. PCA finds the directions with the largest variance, meaning the largest spread of the points, and fits a line into that direction in such a way that the minimum squared distance from the line to each point is minimized (Beebe et al., 1998). The fitted lines are known as the principal component (PC).

The first step in PCA is to calculate eigenvalues, which define the amount of total variation that is displayed on the PC axes. The proportion of variation accounted for

by each PC is expressed as eigenvalue divided by the sum of the eigenvalues. The eigenvector defines the relation of the PC axes to the original data axes.

The first PC summarizes most of the variability present in the original data relative to all remaining PCs. The second PC explains most of the variability not summed up by the first PC and uncorrelated with the first, and so on (Jolliffe, 1986). As PCs are orthogonal and independent of each other, each PC unveils different properties of the original data and may be interpreted independently. In this way, the total variation in the original data set may be broken down into components that are cumulative (Thompson et al., 1998).

PCA aims to describe as much of the data as possible with as little principal components as possible, to extract most of the interesting data from the vast amount of initial data that is mostly irrelevant. In other words, a PCA model that consists of only a few principal components is created to approximate and simplify the data. The PCs are fitted so that the first principal component is aligned with the direction with the most variance (the direction where the spread of the points is largest). The second PC is fitted so, that it goes to that direction orthogonal to PC1 that has the next largest variance in the points. The third PC and all the rest of the PCs are always orthogonal to all of the preceding PCs and similarly go to the direction of the largest remaining variance. The PCs are in order of significance: PC1 describes the most variance and is therefore most important. Because there is random noise distributed evenly to the whole row space and the amount of variance explained by the first PC is largest, the signal-to-noise ratio of PC1 is highest of all PCs, decreasing for each following PC. Eventually, the later PCs have signal-to-noise so low that they mostly describe just

noise and are therefore useless (Abdi & Williams, 2010; Jolliffe, 2002; Beebe et al., 1998).

According to Farag et al. (2014) Principal component analysis (PCA) was used to define both similarities and differences among the three artichoke leaf cultivars. In addition, batches from seven commercially available artichoke market products were analysed and variable quality, particularly in caffeic acid derivatives, flavonoid and fatty acid contents.

2.10.1 PCA Prerequisite Notions and Notations

Matrices are denoted in upper case bold, vectors are denoted in lower case bold, and elements are denoted in lower case italic. Matrices, vectors and elements from the same matrix all use the same letter (e.g., $\mathbf{A}$, $\mathbf{a}$, a). The transpose operation is denoted by the superscript T . The identity matrix is denoted $\mathbf{I}$. The data table to be analyzed by PCA comprises I observations by J variables and it is represented by the $I \times J$ matrix $\mathbf{X}$, whose generic element is $x_{i,j}$. The matrix $\mathbf{X}$ has rank L where $L \leq \min \{I, J\}$. In general the data table will be preprocessed before the analysis (Wise et al., 2006). The columns of $\mathbf{X}$ will be centered so that the mean of each column is equal to 0 (i.e., $\mathbf{X}^T \mathbf{1} = \mathbf{0}$, where $\mathbf{0}$ is a J by 1 vector of zeros and $\mathbf{1}$ is an I by vector of ones). As shown in Equation 2, the inertia of a column is defined as the sum of the squared elements of this column as is computed as:

$$\gamma_i^2 = \sum_j x_{i,j}^2. \quad \text{Eq (2)}$$

The sum of all the γ_i^2 is denoted Γ and it is called the *inertia* of the data table or *total inertia*. The center of gravity of the rows also called centroid, denoted g , is the vector of the means of each column of $\mathbf{X}$. When $\mathbf{X}$ is centered, its center of gravity is equal to the $I \times J$ row of vector 0^T . The distance of the i -th observation to g is Shown in Equation 3:

$$d_{i,g}^2 = \sum_j (x_{i,j} - g_j)^2 \quad \text{Eq (3)}$$

When the data are centered, Equation 3 is reduced to:

$$d_{i,g}^2 = \sum_j x_{i,j}^2 \quad \text{Eq (4)}$$

2.10.2 Eigenvalues

After PCA, the size of each component can be measured and often called an eigenvalue. The earlier (and more significant) the components are, the larger their size and the higher the degree of correlation among the variables in the data, the fewer components to capture common information. Eigenvalue of a PC is defined as the sum of squares of the scores as Shown in Equation 5;

$$g_a = \sum_{i=1}^I t_{ia}^2 \quad \text{Eq (5)}$$

where g_a is the a^{th} eigenvalue (Brereton, 2003). Eigenvalues are presented in percentages. Successive eigenvalues correspond to smaller percentage. The cumulative percentage eigenvalue is often used to determine (approximately) what

proportion of the data has been modeled using PCA and is given by $\sum_a^A = 1g_a$. The closer the value to 100 %, the more faithful is the PC model (Brereton, 2003).

2.10.3 Scores and Loadings

PCA results in a transformation of the original data matrix into scores and loadings. The scores have as many rows as the original data matrix and the loadings have as many columns as the original data matrix. Each scores matrix consists of a series of column vectors, and each loadings matrix a series of row vectors. The scores matrices and loadings matrices are composed of several such vectors, one for each principal component. Correlation of principal components and original variables is given by loadings and loadings reflect the relative importance of the variable components and not the importance of the variable itself. Variable loadings can be classified as "strong", "moderate", and "weak", corresponding to absolute loading values of greater than 0.75, 0.75-0.50, and 0.50-0.30, respectively (Jolliffe, 2002).