



## CHAPTER II

### LITERATURE REVIEW

#### 1. The genus *Cinnamomum*

The genus *Cinnamomum* Schaeffer, belonging to the family Lauraceae, comprises about 350 species (Mabberly, 1997) occurring in the tropical and subtropical regions. Its species distribute in continental Asia, East and Southeast Asia, Australia, Pacific islands and a few species in Central and South America.

##### 1.1 Botanical aspect

Description of plants in the genus *Cinnamomum* Schaeffer is as follows.

Schaeffer, Bot. Exped. 74. Oct-Dec. 1760, nom. Conserve.

Mostly evergreen trees or shrubs. Bark, branchlets and leaves very scented. Buds naked or peculate, in the latter case scales distinct or indistinct, imbricate. Leaves alternate, subopposite or opposite, sometimes aggregate on the top of branchlet, leathery, trinerved or triplinerved, also penninervate. Flowers small to middle-sized, yellow or white, bisexual, rarely polygamous, in a panicle, panicle axillary, subterminal or terminal, composed of (1-)3-flowered cymes. Perianth tube short, cupuliform or campanulate, lobes 6, subequal, entirely deciduous or their upper halves deciduous but rarely entirely persistent after anthesis. Fertile stamens 9, rarely less or more, in 3 whorls, the filaments of 1st and 2nd whorls eglandular, but those of 3rd whorl each with a pair of stalked or stalkless glands near at the base, the anthers 4-celled, rarely those of 3rd whorl 2-celled, cells introrse (of 1st and 2nd whorls) or extrorse (of 3rd whorl). Staminodes 3, of the innermost whorl, cordate or sagitate, stipitate. Ovary always as long as style, style slender, stigma capitate or discoid, sometimes 3-lobed. Fruit fleshy, subtended by a perianth-cup, perianth-cup cupuliform, campanulate or conical, truncate or undulate, irregularly denticulate or sometimes with 6 truncate lobe bases at the apex (Li *et al.*, 2005).



Figure 2-1 Morphology of (1-4) *Cinnamomum verum* and (5-6) *Cinnamomum camphora*

### 1.2 *Cinnamomum* of commerce

Many species of *Cinnamomum* have medicinal and spice values and are of great demand commercially. The major use of several *Cinnamomum* species is as a spice, for example, cinnamon and cassia, which are among the earliest known spices used by humankind. Cinnamon is the dried inner bark of *C. verum* (syn. *C. zeylanicum*) native to Sri Lanka. It is known in trade as Ceylon cinnamon or Sri Lankan cinnamon. Cassia or cassia cinnamon comes from different sources, e.g. *C. cassia* (Chinese cassia), *C. burmannii* (Indonesian cassia or Padang cassia), *C. loureirii* (Vietnamese cassia or Saigon cassia), and *C. tamala* (Indian cassia). Cinnamon and cassia, both in whole and ground form, are widely used for culinary purposes and for flavouring processed foods (bakery products, sauces, pickles, beverages), in perfumes, pharmaceutical products and incense. The bark and leaves can be further used for distillation of essential oils which are also used in flavouring and perfumery.

Although the camphor tree (*C. camphora*) belongs to the genus *Cinnamomum*, it is not used as a spice. The leaves, wood and roots of *C. camphora* is an important

source of natural camphor. The production of camphor was originated in China, and then has spreaded to Japan and Taiwan since World War II.

The use of *C. camphora* as a source of Ho leaf oil, on the other hand, has expanded in recent years, and it is now an important source of natural linalool. Chinese Ho oil has largely replaced the use of rosewood as a source of natural linalool (Coppen, 1995).

In addition to its use as chemical feedstock, the wood of *C. camphora* and other *Cinnamomum* species is also traded as timber for decorative work, furniture, and cabinet making. Many species have fragrant wood which is resistant to insect attack and suitable for making moth-proof chests (Lemmens et al., 1995). In southern part of Thailand, the wood sculpture made from Thep-thaa-ro (*C. porrectum*) has become one of handicraft products promoted in the One Tambon One Product (OTOP) project.

### 1.3 *Cinnamomum* species in Thailand

The occurrence of the genus *Cinnamomum* in Thailand has been reported by Tem Smitinand (2001). This genus consists of 19 species as shown below:

#### *Cinnamomum bejolghota* (Buch.-Ham.) Sweet

Vernacular name: ขนนมะแวง Khanun ma waeng, เขียวใหญ่ Chiak yai (Trang); จวงตง Chuang dong (Nong Khai); เขียด Chait, บริแวง Bori waeng (Ranong); ฝนแสนห่า Fon saen ha, สมุลแว้ง Samum lawaeng (Nakhon Si Thammarat); พะแว Pha wae, โมงหอม Mong hom, ระแวง Ra waeng (Chon Buri); มหาปราบ Maha prap (Trat); มหาปราบตัวผู้ Maha prap tua phu (Chanthaburi); แลงแวง Laeng waeng (Pattani); อบเชย Op choei (Bangkok, Uttaradit).

Small to big tree, 5-25 m tall. Bark green, scented. Branches always opposite, robust; branchlets terete or obtusely tetragonous, red-brown when dry, puberulent initially soon glabrescent. Buds small, ovoid; scales densely sericeous. Leaves subopposite, elliptic-oblong, 12-30 by 4-9 cm, thick leathery, green and nitid above, greenish or yellow-green and ± glaucous beneath, glabrous on both surfaces, trinerved or triplinerved, the apex obtuse, acute or acuminate, the base subrounded or attenuate,

the margin entire, the basal lateral nerves arising at 0.5-1.5 cm above the leaf base, oblique, as straight costa attaining to the leaf apex, as costa also slightly impressed or elevated above and conspicuously elevated beneath, the transversal veins and veinlets inconspicuous above but slightly conspicuous beneath, reticulate; petioles robust, 1-1.5 cm long, plane-convex. Panicle axillary on the upper part of branchlet, 12-116 cm long, densely many-flowered, much-branched, the branches about 3 cm long; peduncle 7-11 cm long, as rachis sparsely gray-pubescent. Flowers yellow, up to 6 mm long; pedicels 4-6 mm long, gray-pubescent. Perianth tube short, obconical, about 1 mm long; the lobes 6, ovate-oblong, 5 by 2.5 mm, acute, gray-pubescent except the apex subglabrous on both surfaces. Fertile stamens 9, about 3.5 mm long (of 1st and 2nd whorls) or 3.7 mm long (of 3rd whorl); the filaments complanate, those of 3rd whorl each with a pair of long-stalked orbicular-reniform glands, others glandless; the anthers of 1st and 2nd whorls ovate-oblong, almost as long as filaments and with introrse cells, but those of 3rd whorl narrower, oblong, about 1.7 mm long, and with extrorse cells. Staminodes 3, of the innermost whorl, conspicuous, 3 mm long, sagittate-deltoid, long-stalked. Ovary oblong, 1.5 mm long, style slender, up to 3 mm long, stigma discoid. Fruit ellipsoid, 1.3 by 0.8 cm, green when fresh; perianth-cup in fruit yellow but with purple-red tinged, somewhat dilated, obconical, the apex up to 7 mm wide, dentate, the teeth truncate; fruit stalk purple, somewhat dilated (Li *et al.*, 2005).

*Cinnamomum burmannii* (Nees) Blume

Vernacular name: **อบเชยชวา** Op choei chawa (Bangkok).

Evergreen shrub or small tree, up to 15 m tall. Leaves subopposite; petiole 0.5-1 cm long; blade oblong-elliptical to lanceolate, 4-14 cm by 1.5-6 cm, pale red and finely hairy when young, when older glabrous and glossy green above, glaucous pruinose below. Inflorescence a short axillary raceme; pedicel 4-12 mm long; perianth 4-5 mm long, after anthesis the lobes tearing off transversely about halfway; stamens about 4 mm long, staminodes 2 mm. Berry ovoid, about 1 cm long (Guzman, and Siemonsma, 1999).



Figure 2-2 *Cinnamomum burmannii* (1) flowering branch, (2) flower, (3) stamen of 1<sup>st</sup> and 2<sup>nd</sup> whorl, (4) stamen of 3<sup>rd</sup> whorl with glands, (5) staminode of 4<sup>th</sup> whorl, (6) fruits

*Cinnamomum camphora* (L.) J. Presl

Vernacular name: พรมเส็ง Phrom-seng (Shan-Northern); การบูร Karabun, อบเชยญวน Op choei yuan (Central).

Large, evergreen, fragrant tree, 15-30 m tall; root system extensive and shallow; trunk short, stout; bark deeply furrowed; crown spreading, up to 30 m wide; twigs brown, yellowish or pinkish when young, glabrous; buds stout, ovoid, pubescent, with many imbricate scales. Leaves alternate, aromatic; petiole slender, 1.5-3 cm long; blade broadly ovate-elliptical to oblong-lanceolate, 5-12 cm by 2-7 cm, base obtuse, margin slightly undulate, apex acute or acuminate, chartaceous, deep green, shiny, glabrous above, glabrous or sparsely hairy beneath, with 3 main veins and 2 conspicuous, impressed glands in vein axils, major veins prominent on both sides. Inflorescence an axillary, many-flowered panicle, up to 7 cm long; pedicel 1-1.5 mm long, glabrous; flowers bisexual, small; perianth tubular, 6-lobed, membranaceous, partly persistent in fruit; lobes ovate, 2.5-3 mm by 1 mm, obtuse, yellowish-green glabrous outside, pubescent inside, transversely tearing off near the base; fertile stamens 9, in 3 whorls,

pubescent; 1<sup>st</sup> and 2<sup>nd</sup> whorls eglandular, anthers oblong, 0.5 mm long, introrse; 3<sup>rd</sup> whorl with 2 subsessile, ovate glands at the base and extrorse anthers; 4<sup>th</sup>, innermost whorl consisting of 3 eglandular staminodes, ovoid, with short filaments; anthers open upwards by flaps; ovary superior, ovoid, subsessile, glabrous; style up to 2 mm long. Fruit a compressed-globose berry, 7-10 mm in diameter, violet-black when ripe, one-seeded. Seed 6-7 mm in diameter (Oyen, and Xuan Dung, 1999).



Figure 2-3 *Cinnamomum camphora* (1) flowering branch, (2) flower, (3) longitudinal section through flower; (4) stamen with 2 basal glands, (5) fruits

*Cinnamomum cassia* Presl (syn. *C. aromaticum* Nees)

Vernacular name: **อบเชยจีน** Op choei chin (Bangkok).

Evergreen tree up to 18 m tall, strongly aromatic in all its parts; bole up to 70 cm in diameter; bark thick, smooth in young trees, rough in mature trees, grey; twigs brown-hairy. Leaves alternate to nearly opposite; blade oblong-elliptical, 8-20 cm by 4-7.5 cm, dark shiny green. Inflorescence 7.5-18 cm long; pedicel 2-3 mm long; flowers small, about 3 mm long, pubescent, white or whitish yellow; perianth lobes after anthesis tearing off transversely at base. Fruit ovoid to ellipsoidal, 1-1.5 cm long, black to blackish-purple. Seed ovoid, 1 cm long, dark brown with paler stripes (Guzman, and Siemonsma, 1999).



Figure 2-4 *Cinnamomum cassia* (1) flowering branch, (2) fruit and tepal, (3) fruit, (4) dissected flower, (5) outer stamen and inner stamen with basal gland, (6) staminode (Ravindran, Nirmal-Babu and Shyraj, 2003)

*Cinnamomum crenulicupulum* Kosterm.

Vernacular name: ฮางแกง Hang kaeng (Chiang Mai)

*Cinnamomum deschampsii* Gamble

Vernacular name: เขียดตัวเมีย Chiat tua mia (Narathiwat); เตยอ Tae-yo (Malay-Narathiwat).

Bushy tree, stem about 1.5 m tall, and 30 cm in diameter. Leaves coriaceous; petiole 0.5 cm long; blade oblong to elliptical-ovate, 7.5-15 cm x 5-7 cm. Inflorescence a lax spreading panicle with silky flowers. Fruit an ellipsoidal one-seeded berry, about 1 cm long.

The bark is thick, very aromatic and with a very pleasant flavour, resembling Chinese cassia (Guzman, and Siemonsma, 1999).

*Cinnamomum glaucescens* (Nees) Drury

Vernacular name: กะเพราตัน Kaphrao ton (Nakhon Ratchasima)

Tree; branches stout, smooth black when dry. Leaves very variable in size, thickly coriaceous, often glaucous beneath, brown when dry; nerves erecto-patent. Petioles 1.2-2.5 cm long, slender. Panicles crowded, densely tomentose. Flowers densely tomentose on both sides. Fruit oblong, 3 cm long (Ravindran *et al.*, 2003).



Figure 2-5 *Cinnamomum glaucescens* (ราชบัณฑิตยสถาน, 2547)

*Cinnamomum illicoides* Chev.

Vernacular name: ขาดัน Kha ton, ตะไคร้ตัน Takhrai ton (Central); พลูตัน Phlu ton (Chaing Mai)

Tree, 5-18 m tall, up to 90 cm in diam. Corona globose. Bark brown, longitudinally deep-fissured. Old branchlets terete, black-gray; young ones greenish. Leaves alternate, ovate or narrowly ovate-elliptic, 6-11 by 2-16 cm, subleathery, greenish and nitid above, brownish and opaque beneath, pinninervate, the apex acute or short-acuminate, the base broadly cuneate to subrounded, the lateral nerves 2-15-paired, ascendant but curved near the leaf margin, as costa elevated on both surfaces, the axils of lateral nerves always conspicuously dome-shaped beneath, the veins and veinlets reticulate, inconspicuously foveolate on both surfaces; petioles 1.2-12 cm long.



Flowers unknown. Inflorescence paniculate, axillary or subterminal, 6.5-7 cm long, pedunculate, peduncle robust, about 2.5 cm long, as rachis yellow brown-villous. Fruit obovoid, about 2 cm long, purple-black; perianth-cup in fruit campanulate, green, 1.2-1.8 cm long and broad (Guzman, and Siemonsma, 1999).

*Cinnamomum iners* Reinw. ex Blume

Vernacular name: กระแจ่มอง Krachae mong, กระเขียด Kachiat, กระทั่งนั้น Kathang nan (Yala); กระดังงา Kradang nga (Kanchanaburi); กระพังหัน Kaphang han, โกล่ K ole, เนอม่า Noe-ma (Karen-Kanchanaburi); เขียด, เคียด Khiat, เขียด Chiat, ชะนุดัน Chanu ton (Peninsular); เขียด Chait, มหาปราบตัวผู้ Maha prap tua phu, อบเชย Opchoei, อบเชยตัน Opchoei ton (Central); ดีกสี่สอ Dik-si-so (Karen-Chiang Mai); บอกดอก Bok kok (Lampang); ฝักดาบ Fak dap (Phitsanulok); พญาปราบ Phaya prap (Nakhon Ratchasima); สะวง Sawong (Prachin Buri).



Figure 2-6 *Cinnamomum iners* (1) tree habit, (2) flowering twig, (3) flower; (4) stamen, (5) fruits

A medium-sized tree up to 24 m tall. Bole up to 60 cm in diameter, bark surface smooth, lenticellate, greyish-brown, inner bark pinkish; leaves opposite or subopposite, (5-)7.5-30 cm by 2-13 cm, base cuneate, rarely rounded, apex blunt to acute, often glaucous below, 3-veined, main veins prominent above, tertiary venation scalariform to scalariform-reticulate, faint to distinct below, petiole 1-2 cm long; inflorescence an axillary or terminal panicle, up to 18 cm long; flowers sometimes partly unisexual, densely silky hairy; fruit oblong to narrowly ovoid, c. 1.5 by 1 cm, seated on a perianth cup with persistent perianth lobes (Lemmens, Soerianegara, and Wong, 1995).

*Cinnamomum kerii* Kosterm.

Vernacular name: ละมุนละมั่ง La mun la maeng (Loei).

Tree 15 m high; branchlets stiff, smooth, glabrous; terminal buds desely grey sericeous. Leaves opposite, suboblanceolate, 5.5-9.5 by 2-3 cm, apex very shortly acuminate, base gradually tapering into petiole, rigidly coriaceous, glabrous on both surfaces, upper surface with filiform, prominulous main nerves, usually slightly sunk in a groove, lower surface with prominulous midrib, the 2 lateral nerves slender, reaching the leaf apex, almost flush with the surface. Petioles 7-10 mm long, slightly concave, dark (in sicco). Inflorescences panicles resembling racemes, axillary (sometimes more than one in each axil), rarely extra-axillary, 2-4 cm long, consisting of a long slender, glabrous peduncle, the few branches slender, up to 5 mm long. Pedicel 2-3 mm, slender, dubglabrous. Flower tube very short, tepals elongate, ovate, 2-14 mm long, acute, slightly sericeous, inside densely grey-sericeous; stamens 2.5 mm long, outer stamens filament slender, sericeous, anther large, as long as filament, 4-large-celled, introrse; inner stamens slightly smaller, anthers narrower, the upper pair-celled very small, extrorse, basal glands small, sessile; staminodes small, sagittate, stipitate; style robust, as long as the ovary, stigma small, subpeltate.

*Cinnamomum mollissimum* Blume

Vernacular name: เขียดใบใหญ่ Chait bai yai (Yala)

A small tree up to 15 m tall, bole up to 20 cm in diameter, with small buttresses, bark surface smooth, greyish, inner bark brown; leaves opposite, 10–20 cm × 3.5–7.5 cm, base cuneate, apex pointed, woolly hairy, 3-veined, main veins sunken above, tertiary venation scalariform-reticulate, prominent below, petiole 0.5–2 cm long; inflorescence an axillary or terminal panicle, c. 10 cm long; flowers woolly hairy; fruit ovoid, c. 0.7 cm long, seated on a cup-shaped, wavy-margined perianth (Lemmens *et al.*, 1995).

*Cinnamomum porrectum* (Roxb.) Kosterm. (syn. *C. parthenoxylon* (Jack) Nees)

Vernacular name: จวง Chuang, จวงหอม Chuang hom (Peninsular); จะไคตัน Cha khai ton, จะไคหอม Cha khai hom (Northern); เทพธำโร Thep tharo (Central, Chanthaburi, Surat Thani); พลูตันขาว Phlu ton khao (Chiang Mai); มือแตกะมาจิง Mue-dae-ka-ma-ning (Malay-Peninsular); การบูร Karabun (Nong Khai).



Figure 2-7 *Cinnamomum porrectum* (1) tree habit, (2) leaf, (3) flowering twig, (4) flower, (5) fruits

A medium-sized to large, more or less deciduous tree up to 45 m tall, bole straight cylindrical, up to 105 cm in diameter, sometimes buttressed, bark surface deeply irregularly fissured or cracked, dark grey or greyish-brown, inner bark reddish-brown, laminated; leaves subopposite to spiral, 5-15 cm by 2.5-8 cm, base cuneate to rounded, apex blunt to acuminate, glabrous with 2-18 pairs of lateral veins, main veins prominent above, tertiary venation reticulate, faint on both surfaces, petiole 1.2-3 cm long; inflorescence an axillary or pseudo-terminal panicle, 2.5-15 cm long; flowers glabrous or sparingly hairy; fruit globose to slightly depressed globose, 0.8-1 cm across, seated on a funnel-shaped perianth cup with an entire margin (Lemmens *et al.*, 1995).

*Cinnamomum puberulum* Ridl.

Vernacular name: เขียดตัวผู้ Chiat tua phu (Narathiwat); เตยยอแต่ Tae-yo-ya-tae, ยอแต่ Yo-ya-tae (Malay-Peninsular).

Tree, up to 15 m tall; bole 40 cm in diameter. Young twigs pale yellow hairy, older twigs glabrous, blackish. Leaves alternate, tri-veined; petiole up to 1.5 cm long; blade elliptical to oblong, 6.5-13 cm by 2.5-3.5 cm, leathery, lower surface adpressed hairy and slightly glaucous. Inflorescence an axillary or terminal racemose panicle, 6 cm long; flowers yellow haired. Fruit a cylindrical one-seeded berry, 1 cm by 0.6 cm (Ravindran *et al.*, 2003).

*Cinnamomum rhynchophyllum* Miq.

Vernacular name: เตยยอ Tae-yo (Malay-Narathiwat)

Tree up to 20 m tall; bole 30 cm in diameter. Leaves subopposite, tri-veined; petiole up to 1.5 cm long; blade elliptical to oblong, 7-23 cm x 2-8 cm, apex with 1-2 cm long acumen, leathery, hairy and slightly glaucous below, secondary veins with few lateral veins running towards margin and joining to form looped intramarginal vein. Inflorescence a terminal or axillary hairy panicle, up to 15 cm long, with yellow flowers.

Fruit an ovoid-truncate, one-seeded berry, about 1 cm in diameter (Lemmens *et al*, 1995).

*Cinnamomum sintoc* Blume (syn. *C. cinereum* Gamble.)

Vernacular name: ลูกข่า Luk kha (Chon Buri).

A medium-sized to fairly large tree up to 39 m tall, bole up to 70 cm in diameter, with buttresses up to 2 m high, bark surface smooth to shallowly fissured, lenticellate, grey-brown, inner bark red with white striations; leaves opposite or subopposite, 7-22.5 cm by 2.5-8.5 cm, base narrowly to broadly cuneate, apex blunt to acuminate, hirsute below but glabrescent, 3-veined, main veins prominent above, tertiary venation reticulate, very faint on both surfaces, petiole 0.8-1.8 cm long; inflorescence an axillary or pseudo-terminal panicle, 10-15 cm long; flowers grey tomentose; fruit oblong to ellipsoid, c. 1.8 cm 0.8 cm, seated on a cup-shaped perianth with an entire margin (Lemmens *et al*, 1995).



Figure 2-8 *Cinnamomum sintoc* (1) fruiting twig, (2) flower

*Cinnamomum subavenium* Miq.

Vernacular name: ชะเอม Cha em, ชะเอมเคี้ยว Cha em khrua (Loei); สุรามะริด Sura marit (Nakhon Ratchasima); เส่กอเล Se-ko-le (Karen-Chiang Mai).

A medium-sized tree up to 30 m tall, bole up to 65 cm in diameter, with buttresses up to 2 m high, bark surface smooth, pustular, grey-brown to reddish-brown, inner bark pale brown, soon becoming red on exposure; leaves opposite to alternate, 5.5-15 cm by 2-6 cm, base cuneate, apex pointed, drying reddish on both surfaces, glabrescent below, 3-veined, main veins slightly raised above, tertiary venation scalariform, faint below, petiole 0.6-1 cm long; inflorescence an axillary panicle, 10 cm long; flowers with the perianth densely hairy inside; fruit ovoid, 1-1.2 cm by 0.6-0.7 cm, seated on a cup-shaped perianth with an entire or slightly wavy margin (Lemmens *et al.*, 1995).

*Cinnamomum tamala* (Hamilton) Nees & Eberm.

Vernacular name: แกลง Kaeng (Chiang Mai).



Figure 2-9 *Cinnamomum tamala* (1) flowering twig, (2) fruit. (ราชบัณฑิตยสถาน, 2547)

Tree, up to 20 m tall and 20 cm in diameter. Bark gray-brown, scented. Branchlets terete, tea-brown, glabrous, young ones  $\pm$  angulate, sparsely gray-puberulent initially soon glabrescent. Leaves alternate or those on the young branchlets sometimes subopposite, ovate, oblong or lanceolate, 7.5-15 by (2.5)2-15.5 cm, thin leathery, green and nitid above, green-white and opaque beneath, glabrous on both surfaces, triplinerved, the apex long-acuminate, the base acute or broadly cuneate, the margin entire, the costa straight up to the leaf apex, the basal lateral nerves slightly elevated above but every elevated beneath, the transversal veins undulate, the veinlets reticulate,  $\pm$  conspicuous on both surfaces; petioles 0.5-1.3 cm long, slightly sulcate on the ventral side, glabrous. Panicle axillary or terminal, 5-10 cm long, many-flowered, branched, the end of branch being a 2-15-flowered cyme; peduncle 1-4 cm long, as rachis sparsely fine-gray-puberulent. Flowers white-green, up to 6 mm long; pedicels 4-6 mm long, slender fine-gray-puberulent. Perianth sparsely gray-puberulent outside but densely so inside; the tube obconical, short, less 2 mm long; the lobes obovate-oblong, about 4 by 1.5 mm, obtuse. Fertile stamens 9, 3.8 mm long (of 1st and 2nd whorls) or 4 mm long (of 3rd whorl); the filaments gray-villous, about 2.5 mm long, those of 3rd whorl each with a pair of finely stalked ovate-cordate glands at the lower 1/3 part, others glandless; the anthers of 1st and 2nd whorls ovate-oblong, 1.3 mm long, with introrse cells, those of 3rd whorl oblong, 1.5 mm long, with extrorse cells. Staminodes 3, of the innermost whorl, villous, 1.7 mm long, long-stalked, the apex triangular-sagittate. Ovary ovoid, 1.2 mm long, villous, style slender, 3.6 mm long, stigma small, inconspicuous (Li *et al.*, 2005).

*Cinnamomum verum* J. Presl (syn. *C. zeylanicum* Bl.)

Vernacular name: อบเชยเทศ Op choei thet (Bangkok).

Evergreen tree up to 18 m tall; bole low-branching, up to 60 cm in diameter; buttresses 60 cm tall, 70 cm deep, thin, light pinkish-brown; bark about 10 mm thick, strongly aromatic; the bark on young shoots is smooth and pale brown, on mature branches and stems rough, dark brown or brownish-grey; oil cells are located in the phloem, and are oval or round in cross-section; wood of mature trees varies from light

brownish-grey to grey or yellowish-brown, without marking, more or less lustrous and faintly scented. Leaves opposite, somewhat variable in form and size, strongly aromatic; petiole 1-2 cm long, grooved on upper surface; blade ovate to elliptical, 5-25 cm by 2-110 cm, conspicuously 3 veined, or 5 veined, base rounded, apex acuminate, glabrous, cricaceous, shiny dark green. Inflorescence consisting of lax axillary or terminal panicles up to 10 cm long or longer; peduncle creamy white, softly hairy, 5-7 cm long; flowers small, 3 mm in diameter, with foetid smell, pale yellow, subtended by small ovate hairy bract; perianth 8 mm long, silky hairy, with short campanulate tube and 6 persistent tepals about 3 mm long; fertile stamens 9, in 3 whorls, with 2 small glands at the base of the stamen of the 3<sup>rd</sup> whorl; a fourth innermost whorl consists of 3 staminodes; filaments hairy, stout; anthers 4- or 2-celled; ovary superior, 1-celled, with a single ovule, style short. Fruit a 1-seeded berry, ellipsoidal to ovoid, 1-2 cm long, black when ripe, surrounded by the enlarged perianth at the base (Guzman, and Siemonsma, 1999).

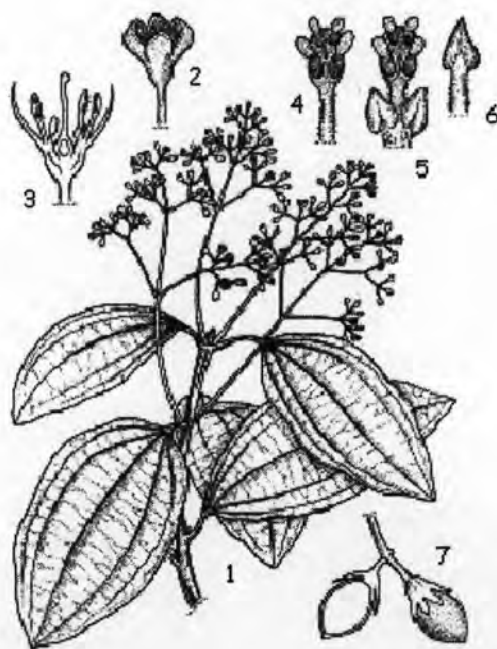


Figure 2-10 *Cinnamomum verum* (1) flowering branch, (2) flower, (3) schematic longitudinal section through flower, (4) stamen of 1<sup>st</sup> and 2<sup>nd</sup> whorl, (5) stamen of 3<sup>rd</sup> whorl with glands, (6) staminode of 4<sup>th</sup> whorl, (7) fruits and longitudinal section through fruit



## 2. Chemical constituents of the leaf essential oil in *Cinnamomum* spp.

Essential oils are very common in the genus *Cinnamomum*. The leaves of many species contain essential oils made up from phenylpropanoids and terpenoids. The following reviews are focussed only on the species of *Cinnamomum* reported in Thailand, of which the chemical compositions of essential oil have been previously investigated. Only the major components of these leaf oils are shown below.

Table 2-1 Major components of some *Cinnamomum* species

| Species   | Major components (%)   | References                     |
|---|--|--------------------------------|
| <i>C. bejolghota</i>                            | linalool (57.4)<br>1-8-cineol (10.2)   | Baruah <i>et al.</i> , 1997    |
| <i>C. burmannii</i>                             | 1,8-cineol (28.5)<br>borneol (16.5)<br>$\alpha$ -terpineol (6.4)<br>p-cymene (6.1) | Ji <i>et al.</i> , 1991        |
|   | linalool (54.9)  | Ding, <i>et al.</i> , 1994     |
| <i>C. camphora</i><br>camphor-type              | camphor (83.9)   | Shi <i>et al.</i> , 1989       |
| linalool-type                                   | linalool (90.6)  |                                |
| cineol-type                                     | 1,8-cineol (50.0)  |                                |
| borneol-type                                    | borneol (81.8)   |                                |
| iso-nerolidol-type                              | iso-nerolidol (57.67)  |                                |
|   | linalool (94.9)  | Jantan, and Goh, 1992          |
|   | camphor (84.1)   | Pelissier <i>et al.</i> , 1995 |
| camphor-type                                    | camphor (65.8)   | Stubbs <i>et al.</i> , 2004    |
| cineol-type                                     | cineol (49.8)<br>sabinene (16.6)   |                                |
| <i>C. camphora</i> var.<br><i>linaloolifera</i> | linalool (91.1)  | Dung <i>et al.</i> , 1993      |

Table 2-1 (Continued)

| Species   | Major components (%)  | References  |
|---|---|---|
| <i>C. cassia</i>                                      | cinnamaldehyde (77.2)<br>coumarin (15.3)                          | Senanayake, 1977: cited in<br>Guzman and Siemonsma,<br>1999         |
|   | benzyl benzoate (92.2)  | Lockwood, 1979  |
|   | cinnamaldehyde (74.1)<br>2-methoxy-<br>cinnamaldehyde (10.5)      | Zhu <i>et al.</i> , 1993: cited in<br>Guzman and Siemonsma,<br>1999 |
| <i>C. glaucescens</i>                                 | elemicin (92.9)   | Baruah and Nath, 2006   |
| <i>C. iners</i>                                       | benzylbenzoate (32.7)<br>geraniol (16.2)                          | Jantan and Goh, 1992  |
|   | ( <i>E</i> )-caryophyllene (20.4)<br>spathulenol (19.8)           | Ubonnuch, 1998  |
| <i>C. porrectum</i><br>(syn <i>C. parthenoxylon</i> ) | geraniol (35.8)<br>neral (28.3)                                   | Zhu, Lu, and Li, 1984   |
|   | 1,8-cineol (20.1)<br>$\alpha$ -terpineol (16.3)<br>safrole (15.4) | Jantan and Goh, 1992  |
|   | ( <i>Z</i> )-isosafrole (97.3)                                    | Ubonnuch, 1998  |
|   | safrole (99.8)  | Palanuvej, Verawatganone,<br>and Ruangrunsi, 2005                   |
| <i>C. rhynchophyllum</i>                              | benzyl benzoate (77)  | Jantan, <i>et al.</i> , 2004  |
| <i>C. sintoc</i>                                      | safrole (23.4)<br>$\gamma$ -muurolene (13.5)                      | Jantan, Yalvema <i>et al.</i> ,<br>2005                             |
| <i>C. subavenium</i>                                  | patchouli alcohol (27.7)<br>benzyl benzoate (19.6)                | Jantan, Muhammad and<br>Nee, 2005                                   |

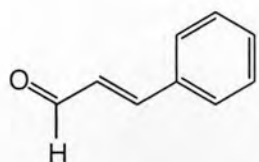
Table 2-1 (Continued)

| Species          | Major components (%)   | References                          |
|------------------|--|-------------------------------------|
| <i>C. tamala</i> | linalool (60.7)<br>$\alpha$ -pinene (10.5)<br>$\beta$ -pinene (10.4) | Nath, Hazarika and Singh,<br>1994   |
|                  | $\beta$ -caryophyllene (25.3)<br>linalool (13.4)                     | Ahmed <i>et al.</i> , 2000          |
|                  | trans-sabinene hydrate<br>(29.8)<br>(Z)- $\beta$ -ocimene (17.9)     | Mir, Ali, and Kapoor, 2002          |
| <i>C. verum</i>  | eugenol (70.1)   | Senanayake, Lee, and<br>Wills, 1978 |
|                  | eugenol (84.1)   | Mallavarapu <i>et al.</i> , 1995    |
|                  | benzyl benzoate (65.4)<br>linalool (10.8)                            | Nath, Pathak and Baruah,<br>1996    |
|                  | linalool (85.7)  | Jirovetz, <i>et al.</i> 2001        |
|                  | eugenol (76.6)<br>linalool (8.5)                                     | Raina, 2001                         |

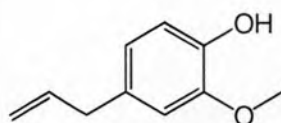
## 2.1 Phenylpropanoids

Phenylpropanoids are naturally occurring phenolic compounds based on a phenylpropane skeleton. As the name implies, the phenylpropanoids contain a three-carbon side chain attached to a phenol. They are biosynthesized from the aromatic protein amino acid phenylalanine and they may contain one or more C<sub>6</sub>-C<sub>3</sub> residues. Common examples include hydroxycoumarins, phenylpropenes, and lignans. The phenylpropenes are important components of many essential oils and include cinnamaldehyde, the major principle of bark oil of cinnamon. Eugenol is also a phenylpropene present in the leaf oil of cinnamon. The phenylpropenes are usually isolated in the essential oil fraction of plant tissues, together with the volatile terpenes. They are lipid soluble, as distinct from most other phenolic compounds. Pairs of allyl and

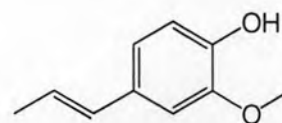
propenyl isomers (e.g. eugenol and isoeugenol) are known, sometimes occurring together in the same plant (Harborne, 1998).



*trans*-Cinnamaldehyde



Eugenol



Isoeugenol

## 2.2 Terpenoids

Compounds with basic skeleton derived from mevalonic acid, or a closely related precursor, are termed terpenoids. Terpenoids are all based on the isoprene molecule (2-methylbutadiene) and their carbon skeletons are built up from the union of two or more of these  $C_5$  units. The classification of terpenoids is presented in Table 2-2 according to the number of isoprene units as building blocks.

Table 2-2 Classification of Terpenoids

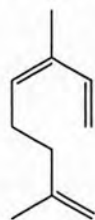
| Class                                    | Number of carbon atoms | Number of isoprenes | Sources   |
|--|------------------------|---------------------|---|
| Hemiterpenoids ( $C_5H_8$ )              | 5                      | 1                   | Volatile oils, esters                                 |
| Monoterpenoids ( $C_{10}H_{16}$ )        | 10                     | 2                   | Volatile oils, glycosides, mixed terpenoids           |
| Sesquiterpenoids ( $C_{15}H_{24}$ )      | 15                     | 3                   | Volatile oils, bitter principles                      |
| Diterpenoids ( $C_{20}H_{32}$ )          | 20                     | 4                   | Resins, chlorophyll                                   |
| Sesterterpenoids ( $C_{25}H_{40}$ )      | 25                     | 5                   | Rare (mostly in animals)                              |
| Triterpenoids ( $C_{30}H_{48}$ )         | 30                     | 6                   | Resins, waxes, steroids, saponins, cardiac glycosides |
| Tetraterpenoids ( $C_{40}H_{64}$ )       | 40                     | 8                   | Carotenoids   |
| Polyterpenoids ( $C_5H_8$ ) <sub>n</sub> | $\infty$               | n                   | Rubber and gutta                                      |

As shown in Table 2-2, terpenoids range from the volatile oil components ( $C_5$  -  $C_{15}$ ) through the less volatile diterpenes ( $C_{20}$ ) to the involatile triterpenoids ( $C_{30}$ ) and carotenoids pigments ( $C_{40}$ ). The volatile terpenes are the components of essential oils and often associated in odour fractions with aromatic compounds such as the phenylpropanoids. Chemically, the terpene essential oils can be divided into two classes: mono- and sesquiterpenes.

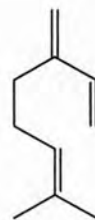
### 2.2.1. Monoterpenoids

Monoterpenoids are colourless, water insoluble liquids with a characteristic aroma, with boiling points ranging from  $140^{\circ}\text{C}$  to  $180^{\circ}\text{C}$ . These  $C_{10}$  compounds are formed by the head-to-tail, head-to-head or tail-to-tail condensation of two isoprene residues. These substances can be further divided into three groups depending on whether they are acyclic (e.g. geraniol), monocyclic (e.g. limonene) or bicyclic (e.g.  $\alpha$ - and  $\beta$ -pinene). Within each group, the monoterpenoids may be simple unsaturated hydrocarbons (e.g. limonene) or may have functional groups and be alcohols (e.g. menthol), aldehydes or ketones (e.g. menthone, carvone).

#### 2.2.1.1. Acyclic monoterpenoids

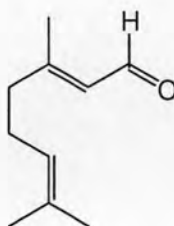


Ocimene

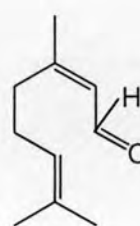


Myrcene

#### Aldehydes

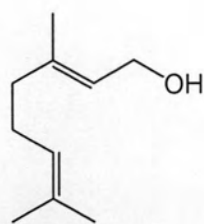


Geranial

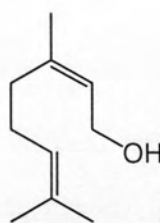


Neral

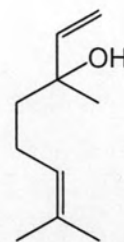
## Alcohols



Geraniol

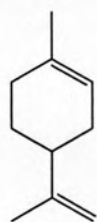
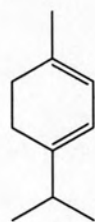
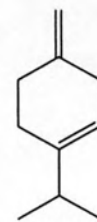
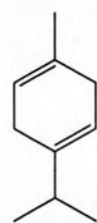
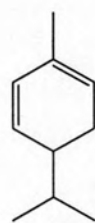
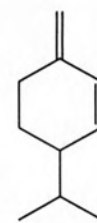


Nerol

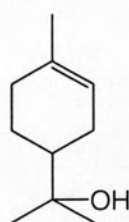
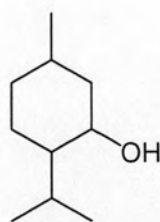


Linalool

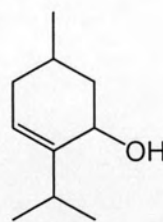
## 2.2.1.2. Monocyclic monoterpenoids

*d*-Limonene $\alpha$ -Terpinene $\beta$ -Terpinene $\gamma$ -Terpinene $\alpha$ -Phellandrene $\beta$ -Phellandrene

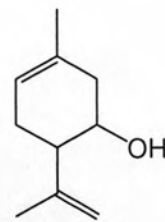
## Alcohols

 $\alpha$ -Terpineol

Menthol

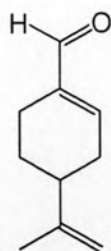


Piperitol

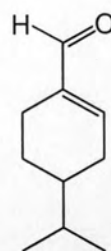


Carveol

## Aldehydes

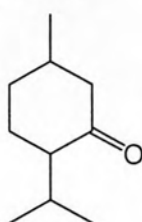


Pirillaldehyde

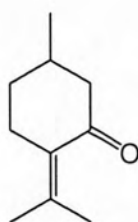


Phellandral

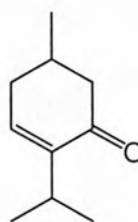
## Ketones



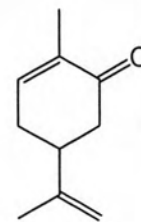
Menthone



Pulegone

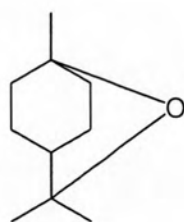


Peperitone

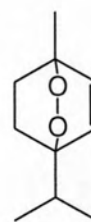


Carvone

## Oxides



1,8-Cineol

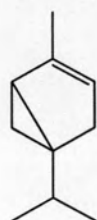


Ascaridole

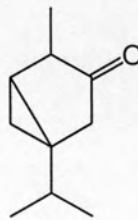
## 2.2.1.3. Bicyclic monoterpenoids

The bicyclic monoterpenoids may be divided into five groups.

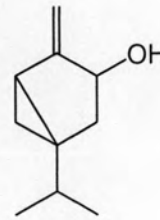
## Thujane group

 $\alpha$ -Thujene

Sabonene

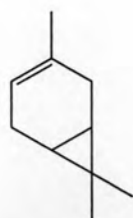


Thujone

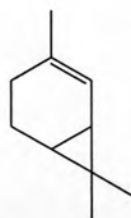


Sabinol

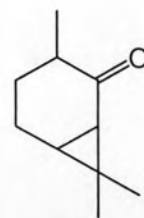
## Carane group



Car-3-rene

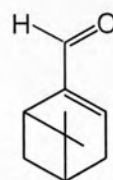


Car-4-rene

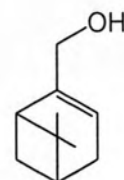


Carone

## Pinane group

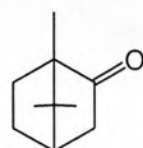
 $\alpha$ -Pinene $\beta$ -Pinene

Myrenal

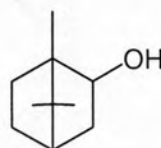


Myrtenol

## Camphane group



Camphor

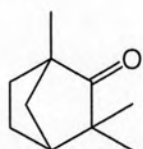


Borneol

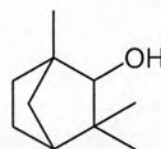


Camphene

## Fenchane group



Fenchone



Fenchyl alcohol

 $\alpha$ -Fenchene

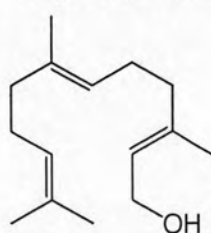
Simple monoterpenoids are widespread and tend to occur as components of the majority of essential oils. Some compounds are regularly found together in leaf oils, especially  $\alpha$ - and  $\beta$ -pinene, limonene, 3-carene and  $\alpha$ -phellandrene.



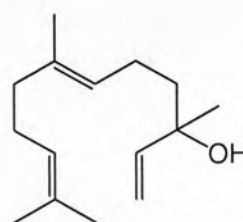
### 2.2.2. Sesquiterpenoids

Sesquiterpenoids are composed of three isoprene units, which have boiling points range over 200°C. Like the monoterpenes, the sesquiterpenes fall chemically into groups according to the basic carbon skeleton. The common ones are either acyclic, monocyclic or bicyclic. However, within each group there are many different compounds known, according to a recent estimate, there are several thousand sesquiterpenoids with well-defined structures, belonging to some 200 skeletal types.

#### 2.2.2.1. Acyclic sesquiterpenoids

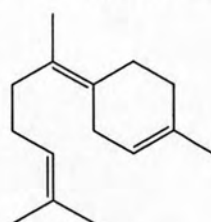
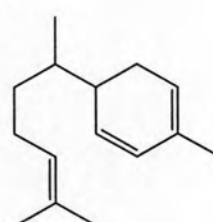


Farnesol



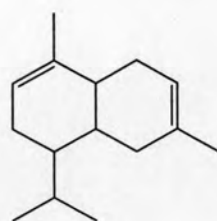
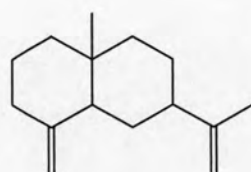
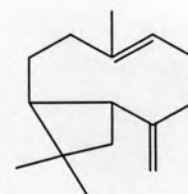
Nerolidol

#### 2.2.2.2. Monocyclic sesquiterpenoids

 $\gamma$ -Bisabolene

Zingiberene

#### 2.2.2.3. Bicyclic sesquiterpenoids

 $\alpha$ -Cadinene $\beta$ -Selinene

Caryophyllene

### 3. Chemometrics

#### 3.1 Introduction to chemometrics

Previously, the main problem confronting analytical scientists was how to obtain data. At that time measurements were labor-intensive, time-consuming, with low sensitivity and manual operating. Currently, the field of chemistry is facing major change, many modern analytical instruments are equipped with advance optical, mechanical and electronic components with high sensitivity and automated operating to produce high-quality data. Moreover, a modern technology called "hyphenated instrumentation" using two or more devices such as gas chromatography coupled with mass spectrometer (GC-MS) has been developed. Tremendous amount of data are generated from these pieces of equipment. For example, hundreds of volatile compounds can be investigated easily by a typical run of essential oils on automated GC-MS. To mine valuable information from these data, different mathematical techniques have been used (Chau, 2004).

Chemometrics, defined as "The chemical discipline that uses mathematical and statistical methods to design or select optimal procedures and experiments, and to provide maximum chemical information by analyzing chemical data" (Massart *et al.*, 1986), has been considered a branch of analytical chemistry since the middle of 1970's. Chemometrics provides powerful methods to reduce the large amount of data and used to analyze and interpret a cluster of raw data into knowledgeable information (Figure 2-11).

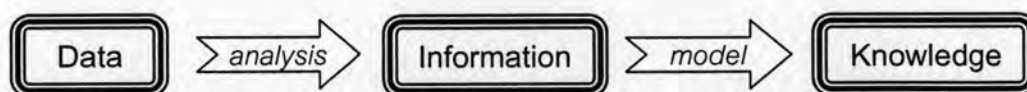


Figure 2-11 Transformation order of data to knowledge chemometrically

Most analytical procedures attempt to make a problem univariate, such as look at only a single unknown compound. But some problems are complex by nature. Typical for problems of modern analytical chemistry is the use of many variables, or multivariate,

to describe a system. Examples of such widespread problems are: recognition of the chemical structure from spectral data (spectral classification), quantitative analyses of substances in complex mixtures (multivariate calibration), prediction of properties or activities of chemical compounds or technological materials (quantitative structure-activity or structure-property relationships) and, as in this research, determination of the origin of samples (cluster analysis and classification) (Varmuza, 2001).

A wide discipline of chemometrics is pattern recognition, which involves the classification and identification of samples. Its purpose is to develop a semi-quantitative model that can be applied to the identification of unknown sample patterns (Aboul-Enein *et al.*, 2000)

### 3.2 Pattern recognition

Human are very good at perceiving similarities and differences between objects of different shapes. For example, discriminating a square from a circle is learned by very young children through games and toys. The goal of pattern recognition in analytical chemistry is very similar: finding similarities and differences between chemical samples based on measurements made on the samples (Beebe, Pell, and Seasholtz, 1998). Pattern recognition is one of the first and most publicized success stories in chemometrics. Much chemistry involves using data to determine patterns. For examples, is there a pattern in the spectra allowing physical information to be related to chemical knowledge? Can a chromatogram be used to decide on the origin of a wine and, if so, what main features in the chromatogram distinguish different wines? (Brereton, 2003)

Given a pattern, its recognition or classification may consist of one of the following two tasks: 1) supervised classification (e.g., discriminant analysis) in which the input pattern is identified as a member of a predefined class, 2) unsupervised classification (e.g., clustering) in which the pattern is assigned to a hitherto unknown class (Jain, Duin and Mao, 2000). Unsupervised pattern classification or clustering may be viewed as exploratory data analysis and it may provide a successful conclusion to a study. On the other hand, it may be a mean of preprocessing the data for a supervised

classification procedure (Webb, 1999). Many methods have their origins in numerical taxonomy (Brereton, 2003) as useful tools for presenting the natural clustering of samples within the data set. The two methods of unsupervised pattern recognition are reviewed below:

### 3.2.1. Hierarchical cluster analysis (HCA)

Hierarchical cluster analysis (HCA) is an unsupervised technique that examines the interpoint distances between all of the samples and represents that information in the form of a two-dimensional plot called a dendrogram (Beebe *et al.*, 1998). The vector of data points for a given sample may be considered as a point in an  $n$ -dimensional hyperspace, where  $n$  is the number of variables (Hibbert, 1997). The similarity between two points closest together is assessed by measuring the distances between them in the hyperspace. Samples that are similar will lie close to one another, whereas dissimilar samples are distant from each other.

The starting point of hierarchical clustering experiment is the similarity matrix which is formed by first computing the distances between all pairs of points in the data set. The choice of the distance metric to express similarity depends on the type of measurement variables used. Typically, three types of variables – categorical, ordinal, and continuous – are used to characterize chemical samples. Measurement variables are usually continuous. For continuous variables, the Euclidean distance is the best choice, because interpoint distances between samples can be computed directly by

$$d_{kl} = \sqrt{\sum_{j=1}^n (x_{kj} - x_{lj})^2}$$

where there are  $n$  measurements and  $x_{kj}$  is the  $j$ th measurement on sample  $k$ ,  $d_{kl}$  is the Euclidean distance between sample  $k$  and  $l$  (Figure 2-12). The smaller this value is, the more similar the samples are (Lavine, 2000; Brereton, 2003).

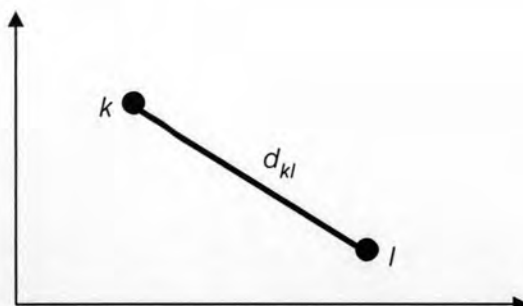


Figure 2-12 Euclidean distance between two data points in a two dimensional space

Then each distance is converted into a similarity value by the following equation:

$$s_{kl} = 1 - \frac{d_{kl}}{d_{max}}$$

where  $s_{kl}$  (which varies from 0-1) is the similarity between samples  $k$  and  $l$ ,  $d_{kl}$  is the Euclidean distance between sample  $l$  and  $k$ , and  $d_{max}$  is the distance between the two most dissimilar samples (i.e. the largest distance) in the data set. The similarity values are organized in the form of a table or matrix.

However, the Euclidean distance is only one of many distance measures that can be used and the second point to note is that none of these distance measures necessarily have any physical meaning. They are simply measure of proximity that are interpreted as similarity. Therefore, in all of the pattern recognition techniques it is the relative distance between points (i.e. is point A closer to B than it is to C) that is important (Beebe *et al.*, 1998).

Examples of the other ways of determining how similar samples are to each other are as follows

- Square Euclidean Distance

$$\text{Distance (X, Y)} = \sum (X_i - Y_i)^2$$

- Pearson Correlation

$$\text{Similarity (X, Y)} = \frac{\sum Z_{xi}Z_{yi}}{N-1}$$

$Z_{xi}$  is Z score of  $X_i$

- Manhattan Distance

$$\text{Distance}(X, Y) = \sum |X_i - Y_i|$$

The next step is to link the samples. The similarity matrix is then scanned for the largest value, which corresponds to the most similar point pair. The two samples constituting the point pair are combined to form a new point. The rows and columns corresponding to the old data points are then removed from the matrix. The similarity matrix for the data set is then recomputed. In other words, the matrix is updated to include information about the similarity between the new nearest point and every other point in the data set. The new nearest point pair is identified, and combined to form a single point. This process is repeated until all points have been linked (Lavine, 2000).

There are a variety of ways to compute the distances between data points and clusters in hierarchical clustering (Figure 2-13). The single-linkage method assesses similarity between a point and a cluster of points by measuring the distance to the closest point in the cluster. The complete linkage method assesses similarity by measuring the distance to the farthest point in the cluster. Average linkage assesses the similarity by computing the distances between all point pairs where a member of each pair belongs to the cluster. The average of these distances is used to compute the similarity between the data point and the cluster.

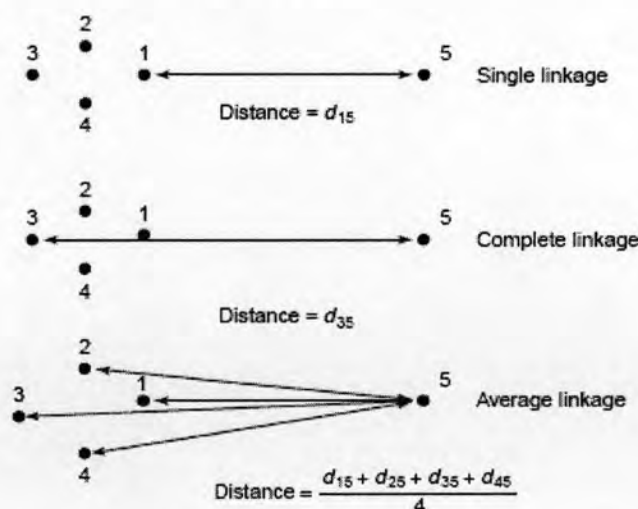
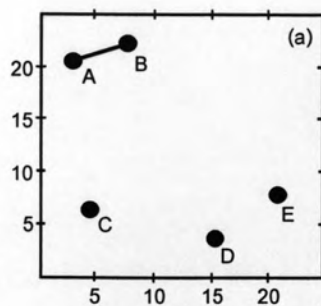


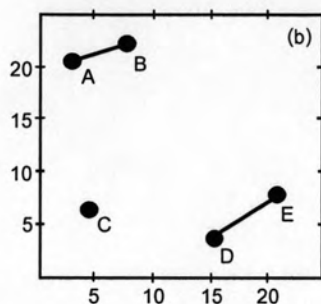
Figure 2-13 The distance between a data cluster and a point using single linkage, complete linkage, and average linkage.

As illustrated in Figure 2-14, in step (a) the samples are initially in five individual points and their similarity matrix were computed. The similarity matrix is then scanned for the largest value, which corresponds to the two samples that are most similar. Sample A and B are the closest two points with a score of 0.77 so they are joined as indicated by the solid line. Once samples A and B are joined into one cluster, the procedure is repeated with the remaining four clusters. In the next step, sample D and E are the closest to each other, and are linked. Then, sample C is linked with the DE cluster. Finally, the two remaining clusters are joined, resulting in a single cluster.

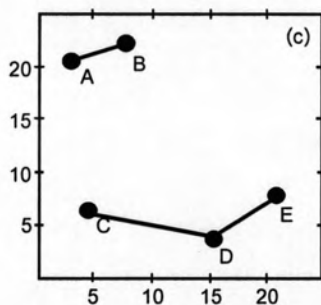




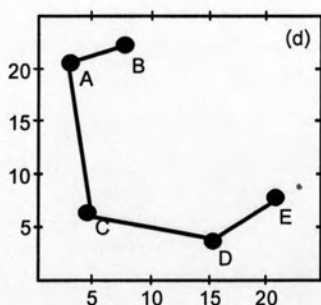
| Samples | A    | B    | C    | D    | E    |
|---------|------|------|------|------|------|
| A       | 1.00 |      |      |      |      |
| B       | 0.77 | 1.00 |      |      |      |
| C       | 0.36 | 0.28 | 1.00 |      |      |
| D       | 0.12 | 0.15 | 0.57 | 1.00 |      |
| E       | 0.00 | 0.10 | 0.30 | 0.71 | 1.00 |



| Samples | A,B  | C    | D    | E    |
|---------|------|------|------|------|
| A, B    | 1.00 |      |      |      |
| C       | 0.36 | 1.00 |      |      |
| D       | 0.12 | 0.57 | 1.00 |      |
| E       | 0.00 | 0.30 | 0.71 | 1.00 |



| Samples | A,B  | C    | D, E |
|---------|------|------|------|
| A, B    | 1.00 |      |      |
| C       | 0.36 | 1.00 |      |
| D, E    | 0.12 | 0.57 | 1.00 |



| Samples | A,B  | C, D, E |
|---------|------|---------|
| A, B    | 1.00 |         |
| C, D, E | 0.36 | 1.00    |

Figure 2-14 Single link clustering for a five-sample, two-measurement data set (left) and its similarity matrix (right) (Adapted from Beebe *et al.*, 1998).



The clustering information can be represented in a two-dimensional dendrogram regardless of the number of measurement variables. The dendrogram for the data in Figure 2-14 is shown in Figure 2-15. The samples are listed on the left-hand side of the graph, and the vertical lines indicate which samples are in a cluster. The x axis is a measure of similarity between clusters which ranging from 0 to 1, and the locations of the vertical lines correspond to the similarity between two jointed clusters.

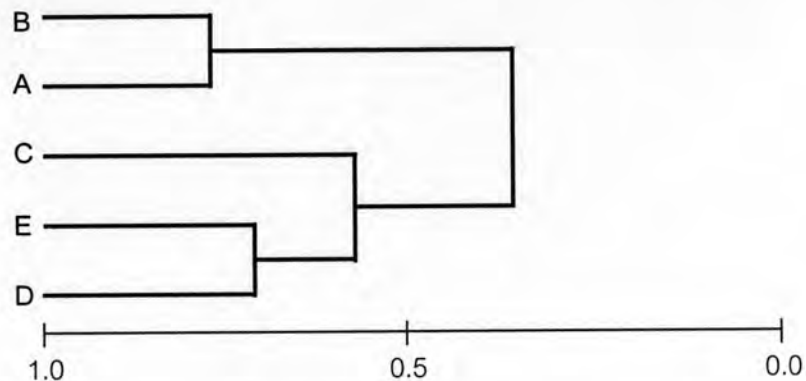


Figure 2-15 Dendrogram of the data in Figure 2-14

A major problem in hierarchical clustering is defining a cluster. There is no cluster validity measure that can serve as an indicator of the quality of a proposed partitioning of the data. Hence, clusters are defined intuitively, depending on the context of the problem, not mathematically, which limits the utility of this technique. Clearly, prior knowledge about the problem is essential when using this method. The criterion for determining the threshold value for similarity is often subjective and depends on the nature of problem investigated, for examples, the goal of the study, previous experience, and common sense.

### 3.2.2. Principal component analysis (PCA)

Principal component analysis (PCA) is probably the oldest and best known of the techniques used for multivariate analysis. The overall goal of PCA is to reduce the dimensionality of a data set while retaining the information present in the data. A new measurement space is constructed in which to plot the samples by redefining the axes

using "factors" rather than the original measurement variables. These new axes, referred to as factors or principal components (PCs), allow the analyst to probe matrices with many variables and view the true multivariate nature of the data in a relatively small number of dimensions.

A two-dimensional plot of an example data is shown in Figure 2-16. The data matrix consists of two columns, representing the two variables, and 40 rows, representing the samples. Each row of the matrix is represented as a point (O) on the graph.

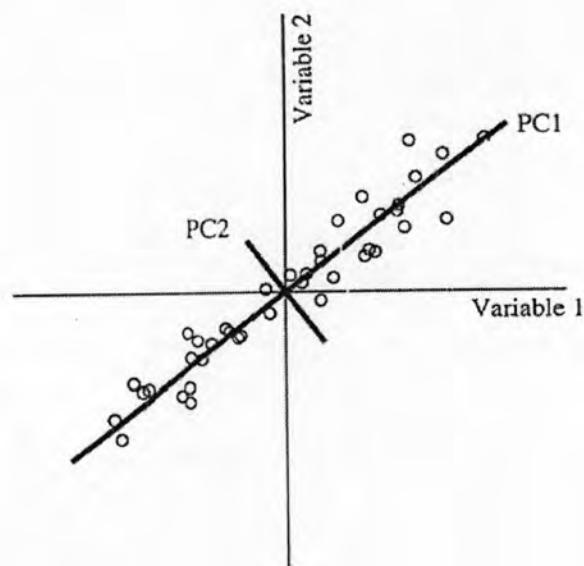


Figure 2-16 Principal component axes (drawn in bold) developed from the two-measurement system.

It is of interest to study the relationship between the samples. The distances between samples are used to define similarities and differences. In mathematical terms, the goal of PCA is to describe the interpoint distances (spread or variation) using as few axes or dimensions as possible. This is accomplished by constructing PC axes that align with the data. The largest or first principal component is formed by determining the direction of largest variation in the original measurement space and modeling it with a line fitted by linear least squares that pass through the center of the data. The second largest principal component lies in the direction of next largest variation - it passes

through the center of the data and is orthogonal to the first PC, and so forth. The maximum number of PCs that can be calculated is the smaller of the number of samples or variables.

The first principal component explains the maximum amount of variation possible in the data set in one direction. Furthermore, the percent of the total variation in the data set described by principal component can be precisely calculated. In this example, the first PC describes 98% of the variation and the remaining variation, 2%, are described by the second PC. Knowing the percent variation described is very important when interpreting the plots. For example, if close to 100% of the variation is described using two PCs, a two-dimensional plot can effectively be used to study the variation in the data set. However, a two-dimensional plot will not be adequate if only 20% of the variation is described. Therefore, conclusions based on examination of PC plots should be tempered by how much of the variation is captured.

Coordination of the sample with the new PC axes can be found by drawing a perpendicular line from the sample to the PC axes. In Figure 2-17, the dashed line shows the coordinates of one sample relative to the first and second principal component. The coordinates of the samples relative to the principal component axes are typically termed "scores".

Excluding non-significant principal components can be used to filter noise from a data set. Each principal component describes some amount of signal and some amount of noise in the data because of accidental correlation between signal and noise. The larger principal components primarily describe signal variation, whereas the smaller principal components essentially describe noise. When smaller principal components are deleted, noise is being discarded from the data. Plotting the scores defined by the two or three largest principal components often provides enough information about the overall structure of the data.

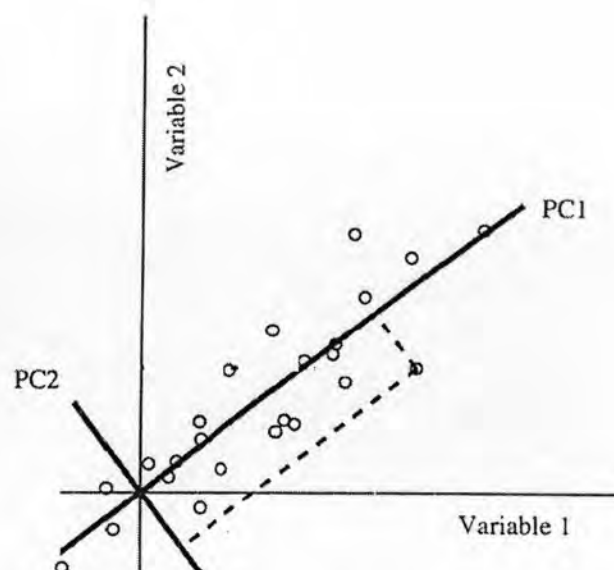


Figure 2-17 The first quadrant of the data in Figure 2-16 showing the coordinates of one point relative to the PC axes (the dash line)

### 3.3 Chemometric analysis of data from essential oil

In 1965, McKern stated that although many promising leads had appeared, it was considered that both chemical and botanical data were still too insufficient to enable the chemical compositions of the essential leaf oils of plants to be used for taxonomic purposes. However, since 1964 enormous advances have been made in analytical instruments. The instruments available today are extremely sensitive and in some cases so stable that accurate analyses can be duplicated for very long periods. The quality of the chemical data now allows us to use chemometric methods to make taxonomic judgements with some confidence (Dunlop, Bignell, and Hibbert, 1997). Essential oils have been used to investigate taxonomic problems in many researches. For example, Hsiao and Lin (1995) have used the cluster analysis on gas chromatogram data from the leaf essential oils of genus *Clerodendrum* (Verbenaceae) to study the relationship among taxa. Without identifying any components, cluster analysis indicated congruence between morphological and chemical relationship at the intraspecific level. Roussis *et al.* (2000) also used cluster analysis to investigate data of the essential oils from the aerial part of four *Helichrysum* species. Although their results show discrepancy with the published taxonomic relationship, interestingly, all chemical

profiles coming from the same taxon are homocategorial. Demetzos, Anastasaki, and Perdetzoglou, (2002) studied the interpopulation variability of the volatile compounds of *Cistus creticus* (Cistaceae) to find out the natural chemotypes of the taxon. Based on their results, cluster analysis of labdane type diterpene patterns could differentiate the groups of individuals by their collection sites.

There are several published data from many countries on the study of the variability and chemical composition of the essential oil in species of *Cinnamomum* (Lauraceae). Among those works, some of them conducted comparative studies to use as chemotaxonomic guide for species identification (Brophy, Goldsack and Forster, 2001; Jantan *et al.*, 2003) and some studied the chemical varieties or chemotypes (Shi *et al.*, 1989; Stubbs and Brushett, 2001; Stubbs *et al.*, 2004). Mostly, these studies did not involve the application of chemometric analysis in chemical classification.

However, there was a chemometric study of *Cinnamomum* has been done almost 20 years ago. In 1988, Guang-fu and Yang used chemometrics combined with morphometrics to study the relationship among *Cinnamomum* species in Hubei province, China (Figure 2-18). Based on the cluster analysis of the chemical constituents and morphological characters, they conducted the phylogenetic system of the genus.

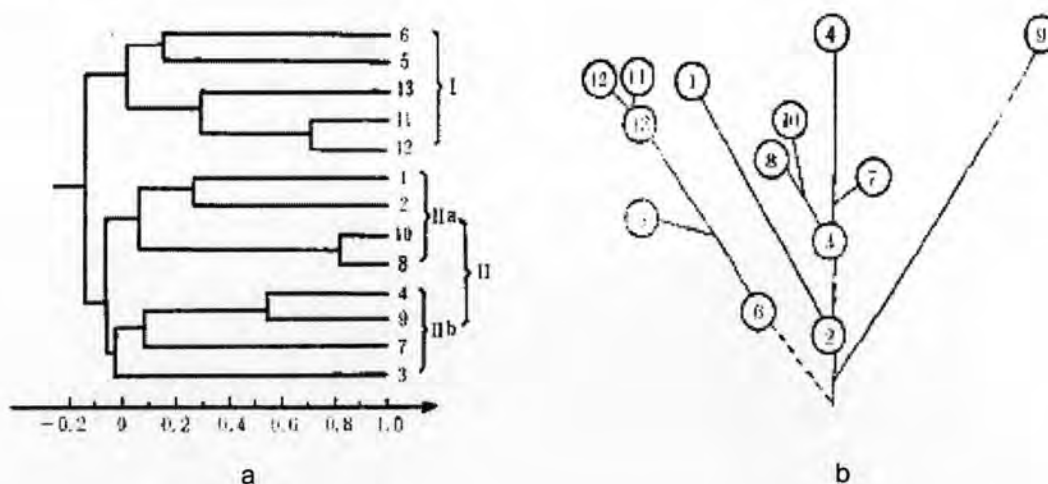


Figure 2-18 a) Cluster analysis shown by correlation coefficient (UPGMA) and b) Evolutionary relationship of *Cinnamomum* in Hubei province (see Table 2-3 for number identification)

Table 2-3 The species and localities of *Cinnamomum* in Hubei province

| Number | Scientific name                                 | Major components       | Locality  |
|--------|---|------------------------|-----------|
| 1      | <i>C. bodinieri</i> var.<br><i>hupehanum</i>    | Citral                 | Changyang |
| 2      | <i>C. bodinieri</i> var.<br><i>hupehanum</i>    | Camphor                | Changyang |
| 3      | <i>C. camphora</i>                              | Cineol                 | Lichuan   |
| 4      | <i>C. camphora</i> var.<br><i>linaloolifera</i> | Linalool               | Lichuan   |
| 5      | <i>C. pauciflorum</i>                           | $\alpha$ -Pinene       | Lichuan   |
| 6      | <i>C. appelianum</i>                            | Cineol                 | Xianfeng  |
| 7      | <i>C. longepaniculatum</i>                      | Bulnesol               | Changyang |
| 8      | <i>C. platyphyllum</i>                          | Trans-methylisoeugenol | Lichuan   |
| 9      | <i>C. parthenoxylon</i>                         | Linalool               | Xianfeng  |
| 10     | <i>C. septentrionale</i>                        | Trans-methylisoeugenol | Lichuan   |
| 11     | <i>C. wilsonii</i>                              | Citral                 | Changyang |
| 12     | <i>C. wilsonii</i>                              | Citral                 | Xianfeng  |
| 13     | <i>C. wilsonii</i>                              | Cinnamic acetate       | Lichuan   |

Recently, Cheng et al. (2006) used cluster analysis on percentage of leaf oil composition of *Cinnamomum osmophloeum* from nine provenances in Taiwan to classify this plant into six chemotypes: cinnamaldehyde type, cinnamaldehyde/cinnamyl acetate type, cinnamyl acetate type, linalool type, camphor type and mixed type.