

CHAPTER 5 DISCUSSION

5.1 Genetic characteristics of Thai people

5.1.1 Thai people are highly admixed

ADMIXTURE analyses indicate that Thai individuals have admixed genetic pattern, composing of shared genetic components with Indian, Chinese and Iban individuals. This reflects the role of the area that is now Thailand as a genetic crossroad. As the area spans as a geographical link between South Asia, East Asia and the Peninsula and Island Southeast Asia, it hosted multiple migration waves and settlements from the neighbor areas. This contributes to the highly admixed genetic patterns in Thai populations.

The findings of admixed genetic patterns in Thai populations coincide with the history and linguistic evidences of past migrations and indigenous inhabitants of mainland Southeast Asia. Historical and ethnolinguistic studies all point to co-existence of people who speak Tai and Austroasiatic languages in this area. Genetic (Peng *et al.*, 2010; Hill *et al.*, 2006) and ethnolinguistic (Enfield, 2003; Wyatt, 1986) evidences suggest that Austroasiatic speakers, specifically the Mon-Khmer, are among the first aboriginal populations of mainland Southeast Asia. Another significant historical migrations in this area was the southwest fanning migration of Tai speakers from southwest China (Wyatt, 1986; London, 2008), as a result of Han Chinese expansion from Northern China around 2,000 years ago (Li *et al.*, 2007). This gradual demic expansion of the Tai happened over centuries, assimilating Austroasiatic speakers who previously inhabited the area (Stuart-Fox, 1998; Wyatt, 1986; Enfield, 2003).

Additionally, during the 1st BC, inhabitants of the mainland Southeast Asia were in contact with traders from China and India. Chinese influences dominated Vietnam while religious influences from India spread to the South and the West of Southeast Asia region (Enfield, 2003). This explains the minor presence of Indian genetic component in Thai and Khmer individuals and its absence in Vietnamese ones.

5.1.2 There are discreet genetic substructures in Thai people

PCA clinal patterns (Figure 4.10A), discernible subgroup patterns concordant with Thailand geography (Figure 4.10B) and different degree of shared ancestry in population admixture (Figure 4.15) indicate that there are discreet substructures in Thai populations. The clinal or continuous substructure patterns suggest gene flow among larger, less mobile populations, which might be associated with agricultural practices. This explains the more prominent continuous pattern in populations in the Central region, where there is a long history of paddy rice field agricultural practices, and less continuous, more secluded genetic substructure in the South where some of the populations also depend on fisheries.

5.1.3 There is a high degree of variation in Thai people genetics

PCA and ADMIXTURE results indicate that there is a high degree of heterogeneity in Thai populations. Thai people's high degree of genetic variation is shown as highly variable PC scattered plots in national (Figure 4.10, Figure 4.11), regional (Figure 4.7, Figure 4.8A) and even global (Figure 4.1) scales. Such high variation pattern could be compared to that of Nepalese and Indian, as in contrary to the highly homogeneous patterns of many islanders, *e.g.* Japanese and Polynesians. Similar scattered PCA

patterns of Nepalese has been suggested as an indication of high level of genetic diversity (Xing *et al.*, 2010). Similarly, subsampling within Indian populations produces more uniformed clusters among the subpopulations (Reich *et al.* 2009; Majumder 2010). High heterogeneity among the Thai individuals is also evident in ADMIXTURE analyses in all levels. This underlines the need for additional genetic studies in mainland Southeast Asia with careful subsampling.

5.2 Comparative population genetic structures: the Thai with the others

5.2.1 Thai and Khmer people: implication to admixture patterns in mainland Southeast Asia

Individual ancestry estimations show that Thai individuals have shared ancestries with the dominant component in Indian, Chinese and Iban individuals. This specific pattern of three-component admixture is also evident in Khmer individuals. This indicates a close genetic history between Thai and Khmer populations while underpins the anthropological and linguistic findings/hypotheses on the close relationship between the Tai and the Khmer in mainland Southeast Asia.

Linguists and anthropologists have long suggested the cultural and linguistic assimilation of the indigenous Khmer into the newly arrived Tai populations. As mentioned in 5.1.1 in anthropological aspects, the Tai's gradual, and sometimes rapid, lowland migration from the South of China was encountered by Austroasiatic-speaking communities who could respond by either receding to higher land or becoming assimilated into Tai, linguistically and culturally (Enfield, 2003; Enfield, 2005). During the 2nd BC at the time when notable Khmer empires, Angkor and Champa, flourished, Tai-speaking settlements increased and found their way to becoming 'the dominant groups of central mainland Southeast Asia' (Stuart-Fox, 1998). In 11th-12th centuries, the ethnic composition of the marginal areas of Angkorian empire had become substantially Tai (Wyatt, 1986).

Linguistic evidences further suggest extensive structural parallelism between Khmer and Tai languages (Enfield, 2003; Enfield, 2005). This is proposed as the result of contact-induced permeability of language structure. It is also mentioned that the process could render genealogical tree model of language relatedness inappropriate.

This study not only confirms genetic inter-mixing between early Tai and Khmer settlers but also raises awareness in using conventional relatedness methods, *e.g.* phylogenetic calculation and summary statistics, with populations in geographical areas where multilateral demic diffusion exists and the genetic profiles of two populations show a significant set of common patterns.

5.2.2 Thai Ayutthaya, Mon and Iban: close genetic relationship among three seemingly unrelated Southeast Asian ethnic populations

Ayutthaya was one of the first most powerful Tai kingdoms. It lasted for more than 400 years and was also referred to as Siam (London, 2008), an old name general Thai people understand as being equivalent to Thailand. Therefore, Ayutthaya individuals were initially chosen in this study as representatives of a more recent Thai (synonymous with Siamese) population. However, ancestry estimation and phylogenetic analyses all indicate that Ayutthaya people have their own distinct genetic profiles, comparing to other recent samples from four regions of Thailand. Phylogenetic findings (Figure 4.14)

indicate that Ayutthaya individuals form an excluded genetic placement on the tree and are older than the general recent populations in Thailand. Also, the patterns of Thai Ayutthaya's ancestry components are different from those of recent Central Thai population, which is almost identical to those of Bangkok population (Figure 4.17, Figure 4.15).

Furthermore, ADMIXTURE analyses also indicate that the genetic profile of Thai Ayutthaya population is closely resemble to that of Mon population in Thailand. As Mon and Khmer are both in Austroasiatic language family, this finding could be associated with the political processes in Ayutthaya establishment in 15th century when the new Tai movement overthrew the previously dominating Khmer empire (Wyatt, 1986; London, 2008). It also confirms the linguistic theories on assimilation of Mon-Khmer-speaking populations into early Tai settlers in lowland Southeast Asia (Enfield, 2003; Enfield, 2005).

Recent study (Simonson *et al.*, 2011) on autosomal genetic data indicate that the Iban have higher genetic similarity to certain Indonesian and mainland Southeast Asian populations than the neighboring Malay populations. The findings in this study confirm as well as provide more clues to this confounding issue. Ayutthaya population could represent, according to historical and linguistic evidences (see explanation in 5.1.1 and 5.2.1) and genetic evidences in this study (see also discussed results in 5.1.1 and 5.2.1), a mixture of early Austroasiatic and Tai speaking populations while the Malays are Austronesian-speaking populations (Peng *et al.*, 2010; Melton *et al.*, 1998; Bellwood 1995).

5.2.3 The Southern Thai: Austronesian, Tai or Austroasiatic?

The Southern Thai populations, both the one collected in this study and the mixed group of Thai Phuket and Moken from Xing *et al.* (2010), have high genetic variation as indicated in PCA results (Figure 4.7B, Figure 4.8, Figure 4.11). Also, when more ancestral clusters are allowed in the analytic models, Southern Thai individuals form distinct genetic clusters, which are apparent even among the datasets that only include Thai samples (Figure 4.16). Phylogenetic trees (Figure 4.13) place them in a separate older clade. Finally, haplotype analyses (Table 4.5 and Table 4.6) confirm that the recent Southern Thai populations are older than the rest of the recent Thai populations. All of these evidences suggest a distinct genetic structure of the Southern Thai population.

Phylogenetic trees reconstructed from recent Thai populations and neighboring populations (Figure 4.14) place recent Southern Thai individuals in the same clade as a number of Central Thai individuals. However, the clade is separated from Iban/Polynesian clade. It has been discussed that Iban share genetic characteristics that are not of Austronesian (Simonson *et al.*, 2011 and 5.2.2, in this thesis) while Polynesians (Kayser, 2010, Wollstein, 2010) and Moken (Peng *et al.*, 2010) are of Austronesian origin.

The origin of Austronesian-speaking descendants is also much debated from a linguistic perspective (Bellwood, 1995). It is suggested that they share ancestry with Tai-Kadai and probably Austroasiatic speaking populations. A recent genetic study into the Cham (Peng *et al.*, 2010), an indigenous population in mainland Southeast Asia, whose ancestors were regarded as the Austronesian immigrants from Island Southeast Asia,

suggests that this population has high genetic affinity with the Mon-Khmer. The origin of the Cham, therefore, is likely a process of assimilation of local Mon-Khmer populations, accompanied with a language shift. With such strong genetic evidences from the previous study and those presented in this study, the genetics of the Southern Thai is worth further investigation with more thorough subsampling.

5.2.4 Northern hill tribal populations:

Ancestry estimation analyses (Figure 4.18, Table 4.3) indicate that the genetic structures of both highland Tai populations and lowland Thai Ayutthaya population are different from hill tribal populations who speak Sino-Tibetan language family, *e.g.* Hmong and Yao. However, Thai Ayutthaya and Mon populations, who are lowlanders, still share a certain degree of genetic structures to some of highland ethnic Tai, *e.g.* Tai Yuan, Tai Kern and Tai Yong. This could provide insights into migration patterns of the early Tai settlers into the lowland and thus worth further investigation.

5.3 Comparison to mtDNA and NRY studies

Table 5.1 shows the genetics studies in Thailand and Table 5.4 shows some of the studies that are relevant to Thai peopling. All of the studies in Thailand focused on indigenous populations, particularly the Northern tribal populations and the Thai-Khmer in Northeastern Thailand. Most of the studies pertained to small sample sizes and small number of markers, mostly mtDNA and Y-chromosomes. Even though the more recent works have begun to use autosomal SNPs, either the number of markers are still limited or the sample populations are still restricted to certain parts of Thailand, emphasizing mostly on indigenous populations. Despite their contribution to the genetic history of the studied populations, these studies still provide little information on the demographic history of the Thai people who live in the other regions of Thailand.

One complication in the studies of two parental lineages is that the information from Y chromosome and mtDNA genealogy do not coincide in every detail. mtDNA mutates faster than Y-chromosome, leading to inconsistency in their phylogenetic branch points (Pakendorf and Stoneking, 2005). More work is still needed to get mtDNA and Y-chromosome trees into correct alignment. Meanwhile, we can only make limited inferences on population history from the two different genetic loci (Stoneking and Delfin, 2010).

In addition, mtDNA and Y-chromosomes are also subject to varying degrees of cultural practices in marriage (Vikrant Kumar *et al.*, 2006). Although patrilocality and matrilocality may not be important on the continental and global scales (Wilder *et al.*, 2004), they influence the patterns of genetic structures on the local scale, in which most of the studies on mtDNA and Y-chromosomes in Thailand were conducted. Comparisons of hill tribal populations that practice patrilocality and matrilocality in Thailand show different levels of mtDNA and Y chromosome differentiations in the two types of societies. Matrilocal societies have enhanced mtDNA differentiation and reduced Y chromosome differentiation (Oota *et al.*, 2001). It is therefore important to investigate autosomal SNPs in addition to existing mtDNA and Y chromosomal DNA studies. While both the social structures and the genetics (discussed in 5.2.4) of Thai hill tribal people are different from the Thai people in other regions of Thailand, care should be taken when using such data to represent or study Thai people at large. Genetic information from autosomal SNPs is a vital complement to mitochondrial and Y-chromosome researches.

Table 5.1 Genetic studies on the populations in Thailand

| Reference | Populations (individuals) | Locations | Genetic markers | Major results/arguments | Note |
|--|---|----------------------------|--------------------------------|--|---------------|
| (Listman <i>et al.</i> , 2011) | Hmong (25), Karen (22), Lahu (24), Lisu (24), Akha (23), Bangkok Thai (34), Bangkok Chinese (4), Chiang Mai Thai (52) | Northern Thailand, Bangkok | autosomal SNPs (2,445 markers) | The inferred level of shared ancestry among the hill tribes corresponds to their linguistic relationships. | Primary data |
| (Kutanana <i>et al.</i> , 2011) | Kon Mueang (436), Lawa (97), Mon (36), Shan (44), Tai Yuan (135), Khuen (48), Lue (92) and Yong (55) | Northern Thailand | Autosomal STRs | <ul style="list-style-type: none"> Kon Mueang exhibit close relationships with and admixed pattern from neighboring populations, including Mon-Khmer and Tai speaking groups. Higher contribution is from the latter. Variability in the Kon Mueang admixture patterns in the same geographical region might reflect roles of other factors, e.g. variation in admixture timing or parental contribution. | Primary data |
| (Xu <i>et al.</i> , 2010) | Mlabri | Northern Thailand | Autosomal SNPs | <ul style="list-style-type: none"> The Mlabri have higher LD and lower haplotype diversity when compared with its neighboring populations. There is a close relationship between Mlabri and Htin. | Use HUGO data |
| (Siriboonpiputtana <i>et al.</i> , 2010) | Central population (501) | Central Thailand | Y-STRs | <i>(This is an announcement of data, not an in-depth population genetic study)</i> | Primary data |
| (The HUGO Pan-Asian SNP Consortium <i>et al.</i> , 2009) | Tai Lue (20), Tai Kern (18), Tai Yuan (20), Mlabri (18), H ^o Tin (18), Plang (18), Lawa (19), Karren (20), Palong (18), Mon (19) | Northern Thailand | Autosomal SNPs | <p>Genetic ancestry is strongly correlated with linguistic affiliations and geography.</p> <p>More than 90% of East Asian haplotypes are found in Southeast Asian or Central-South Asian.</p> <p>There is a clinal structure with haplotype diversity decreasing from south to north.</p> <p>50% of East Asian haplotypes were found only in Southeast Asian while 5% were found only in Central South Asian.</p> <p>The above indicates that South East Asia is a major geographic source of East Asian population.</p> | Primary data |

Table 5.2 Genetic studies on the populations in Thailand (continue)

| Reference | Populations (individuals) | Locations | Genetic markers | Major results/arguments | Note |
|-----------------------------------|---|------------------------------------|------------------------|--|---|
| (Zimmermann <i>et al.</i> , 2009) | Chiang Mai (190) | Chiang Mai | mtDNA | <ul style="list-style-type: none"> The Northern Thai individuals are different in terms of haplotype composition and genetic structure. The Northern Thai are different from Oota <i>et al.</i>, 2001's samples (hill tribes) but more similar to Fuchareon <i>et al.</i>, 2001's (ethnic Thai from various regions). The most common haplotype is F1a. The most common macrohaplogroups are R, M and N. | Primary data compared with Oota <i>et al.</i> , 2001 and Fuchareon <i>et al.</i> , 2001 |
| Jin <i>et al.</i> , 2009 | not clearly specified | not clearly specified | mtDNA and Y-chromosome | <p>(most of their main findings are directed to Korean peopling and thus not described here)</p> <p>Observation: Dominant haplogroups in Thai samples are M, M7b and F1b. Two samples have R haplotype that is absent in all Korean, Han, Mongolian and Vietnamese.</p> <ul style="list-style-type: none"> Tai-Kadai groups formed close affinities among themselves, being more closely related to other Southeast Asian populations than to those from further north. There is evident mtDNA exchange between Tai-Kadai and Khmer. | Primary data (collected as a part of Korean peopling study) |
| (Lertrit <i>et al.</i> , 2008) | Chao-Bon (20), Thai Korat (32) | Nakhon Ratchasima | mtDNA | <ul style="list-style-type: none"> Tai-Kadai groups formed close affinities among themselves, being more closely related to other Southeast Asian populations than to those from further north. There is evident mtDNA exchange between Tai-Kadai and Khmer. | Primary data and data from Fuchareon <i>et al.</i> , 2001 |
| (Listman <i>et al.</i> , 2007) | Hmong (70), Thai Bangkok (45), Thai Chinese in Bangkok | Northern Thailand, Bangkok | Autosomal STRs | <ul style="list-style-type: none"> Hmong genetic pattern is different from Thai and Chinese. Thai and Chinese genetic patterns are indistinguishable. The Thai and Chinese populations have been large and at equilibrium for a long period (no bottleneck). | Primary data |
| (Besaggio <i>et al.</i> , 2007) | Karen (20), Lisu (14), Akha (14), Lu-mien (13), Hmong (7) | Chiang Rai, Chiang Mai, Maehongson | mtDNA and Y STRs | Genetic diversity of northern Thai hill tribes is affected by sex-specific migration rates and cultural factors | Primary data and data from Oota <i>et al.</i> , 2001 and Kwak <i>et al.</i> , 2005 |
| (Joyjinda, 2007) | Thai | unspecified | mtDNA | Haplogroups B, B*, M, F and C were identified. There were difference in Leber hereditary optic neuropathy (LHON) patients and healthy individuals in haplogroups B* and F. | Data collected for association study |

Table 5.3 Genetic studies on the populations in Thailand (continue)

| Reference | Populations (individuals) | Locations | Genetic markers | Major results/arguments | Note |
|----------------------------------|--|--|-----------------------------|--|---|
| (Tharaphan <i>et al.</i> , 2006) | Thai, Chinese-Thai | Siriraj hospital in Bangkok | mtDNA | Haplogroup F, which is the second most abundant haplogroup after haplogroup M, was absent in the patients with LHON. | Data collected for association study |
| (Oota <i>et al.</i> , 2005) | Mlabri (91) | Northern Thailand | mtDNA, Y-STR, autosomal STR | <ul style="list-style-type: none"> The Mlabri have no mtDNA diversity and reduced diversity of Y-chromosomes and autosomes. They appear to have originated as agricultural groups and adopted hunting-gathering lifestyle later. | Primary data and data from Oota <i>et al.</i> , 2001 |
| (Allard <i>et al.</i> , 2004) | Thai (52) | unspecified | mtDNA | <ul style="list-style-type: none"> Haplogroups A, B4a, D4, and F1a were the most commonly observed clusters in the Chinese data set. The next most common haplogroups in the Chinese data set include the clusters C, M7b1, and N9a. | Data from Fuchareon <i>et al.</i> , 2001 |
| (Tajima <i>et al.</i> , 2002) | Thai-Khmer (65) | Thailand, region not specified | Y-chromosome | <ul style="list-style-type: none"> Three major lineages account for 99% of the North, East and Southeast Asians, suggesting 3 parallel (not hierarchical) parental contribution to these Asian populations. Distribution patterns of Y-haplotypes and mutational ages for the key makers suggest that the 3 groups separately migrated to prehistoric East and Southeast Asia. | Primary data and data from Fuchareon <i>et al.</i> , 2001, Khon Kaen (19) and Chiang Mai (15) |
| (Oota <i>et al.</i> , 2001) | Akha, Lahu, Karen, Lisu (number of individuals not specified) | Northern Thailand | mtDNA and Y-STRs | <ul style="list-style-type: none"> Y-chromosomal diversity is higher in populations who practice matrilocality (males move to the wife's residence) while mtDNA diversity is higher in the populations with patrilocality practice. | Primary data |
| (Fuchareon <i>et al.</i> , 2001) | Phu-Thai (25), Lao Song (25), Chong (25), Mussur (21), Lisu (25), Thai Kon Kaen (44), Thai Chiang Mai (30), Sakai (20) | Mukdahan, Suphanburi, Chanthaburi, Khon Kaen, Chiangmai, Trang | mtDNA | <ul style="list-style-type: none"> Lao Song (TK-Central region) and Chong (AA-the East) are closely related. The hill tribes (ST), Phuthai (TK-Northeast) and native Khon Kaen (Northeast) are closely related. Sakai has a unique genetic position, distant from others. | Primary data |
| (Ke <i>et al.</i> , 2001) | Hill tribal people (exact populations not specified) | Northern Thailand, Northeastern Thailand | Y-chromosome | <ul style="list-style-type: none"> Asian individuals have M89T Y-chromosome which branches from M168. Asian individuals descend from out-of-Africa migration. No in situ homimid contribution. | Primary data (studied as a part of Asian populations) |

Note: Populations other than Thai, which present in these studies, are not listed in this table.



Table 5.4 Relevant genetic studies of the peopling of Thailand

| Reference | Referred Thai populations | Targeted populations/ areas of studies | Genetic markers | Major results/arguments | Note |
|---------------------------------|--|--|-------------------------------|---|---|
| (Simonson <i>et al.</i> , 2011) | Xing <i>et al.</i> , 2009, HUGO, 2009 | Iban | Autosomal SNPs, mtDNA and NRY | <ul style="list-style-type: none"> Autosomal data indicate that the Iban have genetic similarity to Indonesia and mainland Southeast Asian populations (Khmer, Thai, Vietnamese). | Iban have B4a*, M*, M7b, N*, Y2, Z and F1a1 mtDNA haplogroups and O, K, C and F NRY sub- haplogroups. The diversity of mtDNA is higher than that of Y-chromosome. |
| (Peng <i>et al.</i> , 2010) | Fucharoen, Fucharoen, Horai, 2001; Oota <i>et al.</i> , 2001; Yao <i>et al.</i> , 2002; Luangtrakool, 2006; Joyjinda, 2007; Lertrit <i>et al.</i> , 2008; Jin, Tyler-Smith, Kim, 2009; Zimmermann <i>et al.</i> , 2009 | East Asia | mtDNA | <ul style="list-style-type: none"> Southern China and/or Southeast Asia likely served as the source of some post-Last Glacial Maximum dispersal(s). Thai Chiang Mai has M9a | -- |



Table 5.5 Relevant genetic studies of the peopling of Thailand (continue)

| Reference | Referred Thai populations | Targeted populations/ areas of studies | Genetic markers | Major results/arguments | Note |
|-------------------------------|--|--|-----------------|---|--|
| (Bodner <i>et al.</i> , 2011) | Zimmermann <i>et al.</i> , 2009 | Laos | mtDNA | <ul style="list-style-type: none"> The genetics of Laotian populations are very diverse, possibly reflecting contribution from several ethnolinguistic groups and a complex migration history. Haplotype structures, diversity indices and haplotype sharing of Laotian population are similar to other Southeast Asia but different from East Asian, indicating a mixed Southeast Asian origin with limited Northern East Asian contributions. Laotian people are most similar to Austroasiatic and Southern Sino-Tibetan populations. | The most prevalent haplogroups are B5a, F1a1a, C7 and M7b1 (macrohaplogroup N and M). |
| (Kong <i>et al.</i> , 2010) | Fuchareon <i>et al.</i> , 2001 and Zimmermann <i>et al.</i> , 2009 | East Asia | mtDNA | <ul style="list-style-type: none"> Modern human initially populated East Asia did not have archaic gene flow from neighbor regions. Highest genetic diversity in Southern China and differentiated basal mtDNA lineages (from Eurasian founder nodes, M and N) support the Southern Route peopling hypothesis of East Asians. Newly identified haplogroups M76, N10 and M71, which have the highest diversity in southern and southwest China while absent in other areas of East Asia, are also found in Southeast Asians in lower frequencies. The above also suggests that Southern China is the genetic reservoir of modern human after they entered East Asia. | M71 is found in Fuchareon in populations other than the northern. M76, N10 and M74 are found in Zimmermann which studied only northern people. These lineages are also found in Vietnamese, Malay and Borneo, not just the northern region of mainland Southeast Asia. |

Table 5.6 Relevant genetic studies of the peopling of Thailand (continue)

| Reference | Referred Thai populations | Targeted populations/ areas of studies | Genetic markers | Major results/arguments | Note |
|--------------------------------|--|--|-----------------|---|--|
| (Chaubey <i>et al.</i> , 2010) | -- | Indian Austroasiatic group | Y-chromosome | <ul style="list-style-type: none"> The higher diversity of haplogroup O2a (found in high frequencies in Indian AA) and coalescent time in SEA indicate that the origin of Austroasiatic speakers is in SEA. The more complex PCA and structure genetic analyses are proposed to be the result of extensive admixture with local populations. | mtDNA haplogroups of Indian and SEA AA are different while both groups have Y chromosome haplogroup O2a |
| (Peng <i>et al.</i> , 2010) | Fucharoen <i>et al.</i> , 2001; Yao <i>et al.</i> , 2002; Lertrit <i>et al.</i> , 2008; Jin <i>et al.</i> , 2009; Zimmermann <i>et al.</i> , 2009; Dancause <i>et al.</i> , 2009 | Cham people (ancient Austronesian) | mtDNA | <ul style="list-style-type: none"> The Cham has closer affinity with the Mon-Khmer populations in mainland Southeast Asia than with the Austronesian populations. The origin of the Cham was likely a process of assimilation of massive local Mon-Khmer populations accompanied with language shift. The above indicates that Austronesian diffusion in mainland Southeast Asia was mediated by cultural diffusion. | Moken and Cham are two Austronesian groups in mainland SEA. |
| (Irwin <i>et al.</i> , 2007) | -- | Northern Vietnamese | mtDNA | -- | Short communication to announce availability of Vietnamese genetic data. |
| (Li <i>et al.</i> , 2007) | -- | Tai populations of Southern China | mtDNA | <ul style="list-style-type: none"> B, M7, F and R are characteristics of Southern Chinese haplogroups. These haplogroups present in higher frequencies in Tai populations than in AA groups even though the latter are regarded as native Southwestern populations. Tai, Hmong and AA are specific to Southeast China neither the North nor the South. | The most common Tai haplogroups are B4a, F1a, M7b1, B5a, M7b*, M*, R9a and R9b. The most common Austroasiatic haplogroups are F1a, M*, D*, F1b, N*, C, M7b*, M7b1 and F1a1a. |
| (Hill <i>et al.</i> , 2006) | -- | Aboriginal people of Southeast Asia | mtDNA | <ul style="list-style-type: none"> The indigenous people of Peninsular Malaysia (Orang Asli-- original people) could be the descendants of early Austroasiatic speaking agriculturists who came into the area ~40,000 years ago. The group of people is different from present Malay who speaks Austroasiatic languages. | Aboriginal groups of peninsula Southeast Asia have F1a1a mtDNA haplogroup. |