Forensic efficiency and genetic polymorphisms of 12 X-chromosomal STR loci in Northeastern Thai populations

Suparat Srithawong[1] · Kanha Muisuk[2] · Nonglak Prakhun[1] · Nisarat Tungpairojwong[1] · Wibhu Kutanan[1,3]

Abstract

Northeastern Thailand comprises one-third of the country and is home to various populations, with Lao Isan constituting the majority, while others are considered minority groups. Previous studies on forensic short tandem repeats (STRs) in Thailand predominantly focused on autosomal STRs but there was a paucity of X-STRs, exclusively reported from the North and Central regions of the country. In this study, we have newly established a 12 X-STRs from a total of 896 samples from Northeastern Thailand, encompassing Lao Isan as the major group in the region, alongside nine minor populations (Khmer, Mon, Nyahkur, Bru, Kuy, Phutai, Kalueang, Nyaw, and Saek). Across all ten populations, the combined powers of discrimination in both genders were high and the combined mean exclusion chance (MEC) indices calculated for deficiency, normal trio and duo cases were also high (> 0.99999). DXS10148 emerged as the most informative marker, while DXS7423 was identified as the least informative. Genetic comparison based on X-STRs frequency supported genetic distinction of cerain minor groups such as Kuy, Saek and Nyahkur from other northeastern Thai groups as well as genetic differences according to the geographic region of Thai groups (Northeast, North and Central). In sum, the overall results on population genetics are in agreement with earlier reports on other genetic systems, indicating the informativeness of X-STRs for use in anthropological genetics studies. From a forensic perspective, despite the limitations of small sample sizes for minority groups, the present results contribute to filling the gap in the reference X-STRs database of the major group Lao Isan, providing valuable frequency data for forensic applications in Thailand and neighboring countries.

Keywords Forensic microsatellites · X-STRs · X-chromosomal microsatellites · Northeastern Thailand · Allelic frequency

Introduction

Thailand is located in the center of mainland Southeast Asia bordered by several countries: Myanmar to the North and West, Laos to the North and Northeast, Cambodia to the Northeast and East, and Malaysia to the South. With a population size ~69 million, the major ethnic groups are Khonmueang, Lao Isan, central Thais and southern Thais who live respectively in four main regions; North, Northeast, Central and South (Myers 2005; Eberhard et al. 2020). The Northeast is the largest region of Thailand, comprising nearly one-third of the total area with a population of ~15 million (Eberhard et al. 2020). The major group Lao Isan historically relocated from Laos during the fourteenth–eighteenth century AD (Myers 2005; Mishra 2010). At that time, besides the Lao people, other ethnic groups from both Laos and Vietnam were also relocated to the area of present-day Northeastern Thailand as minor groups, including several Tai-Kadai (TK) speaking-Phutai, Seak, Nyaw and Kalueang. However, prior to Lao Isan settlement, the indigenous people in this region were of the Austroasiatic (AA) speaking groups, e.g. Khmer and Mon (Myers 2005; Mishra 2010; Higham 2014).

Previous genetic studies utilizing mitochondrial DNA and the Y chromosome indicated contrasting paternal and

✉ Nisarat Tungpairojwong
   knisar@kku.ac.th

✉ Wibhu Kutanan
   wibhuk@nu.ac.th

[1] Department of Biology, Faculty of Science, Khon Kaen University, Khon Kaen, Thailand

[2] Department of Forensic Medicine, Faculty of Medicine, Khon Kaen University, Khon Kaen, Thailand

[3] Department of Biology, Faculty of Science, Naresuan University, Pitsanulok, Thailand

maternal genetic ancestries of Lao Isan (Kutanan et al. 2017; 2019). However, analysis of autosomal DNA markers revealed that the Lao Isan people exhibited a genetic mixture with neighboring AA-speaking groups and displayed regional genetic differences among major groups in Thailand (Srithawong et al. 2020, 2021; Kutanan et al. 2021). While there have been numerous genetic investigations in northeastern Thai populations, with a particular focus on uni-parental and bi-parental genetic features (Kutanan et al. 2014, 2017, 2019, 2021; Srithawong et al. 2015, 2020, 2021; Than et al. 2022), the genetic characteristics of the X chromosome have not yet been reported.

According to gender, X-chromosome exhibits a specific inheritance pattern; father transmits it to daughter as an unchanged block while mother transmits one recombined X-chromosome to her offspring in the same way as autosome (Gomes et al. 2020). These unique properties make X-chromosome a complementary informative marker to autosomes and Y-chromosomes in forensic and population genetics (Gomes et al. 2020; Schaffner 2004). X-chromosomal short tandem repeats (X-STRs) are utilized in forensic investigations as additional evidence for several complicated scenarios of personal identification, e.g. identification cases of missing persons and mass disaster victims, and complex kinship analyses, e.g. deficiency paternity cases and incest cases (Asamura et al. 2006; Liu et al. 2007). In addition, some properties of X-STRs, e.g. high polymorphic, simple analyses and easy genotyping make X-STRs suitable for population-wide genetic study. However, the usefulness of X-STRs in both forensic and population genetic studies requires precise knowledge of not only allele and haplotype frequencies but also the genetic linkage and linkage disequilibrium (LD) status among markers (Inturri et al. 2011).

Within a decade, X-STRs have been continuously developed for commercial kits, e.g. the Investigator Argus X-12 kit (Qiagen GmbH, Hilden, Germany) with high standard production based on numerous makers, methodologies and a large database (Garcia et al. 2022) [6]. X-STRs data have been reported worldwide (Asamura et al. 2006; Li et al. 2011; Shin et al. 2005; Tetzlaff et al. 2012; Pepinski et al. 2007; Illescas et al. 2012; Baeta et al. 2013) but there is a paucity of X-STRs data in Thailand with data available from the northern and central regions of Thailand (Vongpaisarnsin et al. 2016; Khacha-ananda et al. (2020) but not in the Northeast and the South. A previous study suggested that variations in allele frequency of X-STRs might differ among individuals living in different geographic locations (Liu et al. 2007). Furthermore, our earlier study, based on autosomal STRs, also indicated regional genetic differences among Thai populations (Srithawong et al. 2020). Therefore, regional heterogeneity in Thailand may also hold true for X-STRs. To address this question, X-STRs data from northeastern Thailand must be developed. Here, we aim to characterize the genetic variations and forensic parameters of X-STR markers using the Investigator® Argus X-12 kit (Qiagen, Germany) on 10 northeastern Thai populations, encompassing both major and minor groups. Additionally, we compare genetic relatedness based on X-STRs between northeastern Thais and other Asian populations.

Materials and methods

Materials

Population samples

A total of 896 samples belonging to 10 populations included Lao Isan ($n = 498$), Khmer ($n = 40$), Mon ($n = 82$), Nyahkur ($n = 38$), Bru ($n = 36$), Kuy ($n = 56$), Phutai ($n = 57$), Kalueang ($n = 41$), Nyaw ($n = 36$) and Seak ($n = 12$). Almost all of the genomic DNA samples were obtained from our previous studies (Kutanan et al. 2017, 2018), except for 402 newly collected individuals of Lao Isan, using the 1.2-mm-diameter size blood spot on the FTA card. Sample donors were unrelated for at least two generations and provided samples with written informed consent. Ethical approval for this study was provided by Khon Kaen University (HE642161).

Methods

PCR amplification and STR typing

Genomic DNA was extracted using the Chelex-100 method. According to the manufacturer's instructions, each sample was amplified for 12 X-STR loci using reaction mixtures included in the Investigator® Argus X-12 QS kit (Qiagen, Hilden, Germany). The PCR amplicons were separated by capillary electrophoresis on Genetic Analyzer ABI 3500 (Applied Biosystems, USA). The GeneMapper® ID-X software v1.4 (Thermo Fisher Scientific, Inc., Waltham, MA, USA) was used for determining STR allele calling. Raw genotypes of 12 X-STRs of females and males are shown in Online Resource 1, (ESM_1).

Statistical analyses

Allele frequencies for each locus and haplotype frequency of the four linkage groups (linkage group I: DXS10135-DXS10148-DXS8378; linkage group II: DXS7132-DXS10079-DXS10074; linkage group III: DXS10103-HPRTB-DXS10101; and linkage group IV: DXS10146-DXS10134-DXS7423) as well as the gene diversity (GD) and haplotype diversity (HD) were estimated

using StatsX v2.0 software (Lang et al. 2019). StatsX v2.0 software was also used to calculate polymorphism information content (PIC) (Botstein et al. 1980), power of discrimination (PD) for males (PDm) and females (PDf) (Desmarais et al. 1998) and mean exclusion chances (MEC) for deficiency cases (MEC_Kruger) (Krüger et al. 1968), for normal trios (MEC_Kishida) (Kishida et al. 1997), and for duo cases (MEC_Desmarais duo) (Desmarais et al. 1998). Expected heterozygosity (H_E) (Nei 1974), observed heterozygosity (H_O), Hardy–Weinberg equilibrium (HWE) for females, and the pairwise linkage disequilibrium (LD) in the male and female using the exact P test for the 12 loci were performed by Arlequin version 3.5 software (Excoffier and Lischer 2010). Arlequin was also used to perform analysis of molecular variance (AMOVA) by grouping populations according to linguistic categories; i.e. Tai-Kadai and Austroasiatic languages.

A genetic distance matrix based on the number of different allele (F_{st}) was also created by Arlequin, then a matrix was plotted in two dimensions by means of multidimensional scaling (MDS) using Statistica v. 10 demo (StatSoft, Inc, United states). Furthermore, the genetic relationship between 10 studied populations and 17 availably published populations were illustrated by a neighbor-joining tree based on Nei's D_A distances using POPTREE2 software (Takezaki et al. 2014) and two dimensions multidimensional scaling (MDS) using Statistica.

Results and discussions

We genotyped 12 X-STRs using the Investigator® Argus X-12 kit on a total of 896 samples from major northeastern Thai (Lao Isan) and minor groups (AA-speaking-Khmer, Mon, Nyahkur, Kuy and Bru and TK-speaking-Phutai, Kalueang, Nyaw and Seak) from northeastern Thailand. The comprehensive results, encompassing allelic frequency, haplotype frequency, and forensic parameters of X-STRs for both Lao Isan and other minor groups in northeastern Thailand, are presented in Online Resource 2–6; ESM_2–6. For the Lao Isan, all loci are highly polymorphic and informative, indicating that this marker set is a valuable tool for forensic investigations. The newly generated reference X-STRs database of Lao Isan holds significance for forensic applications not only in Thailand but also in Laos. The X-STRs data also support regional genetic differences among major Thai groups in each region: northeastern Thai, Northern Thai, and Central Thai, as well as genetic distinctions of some minor groups, such as Kuy, aligning with findings from other genetic marker systems in our previous studies (Srithawong et al. 2020; Kutanan et al. 2021). This marker set could serve as an optional tool in forensic practice and population

genetic studies in various populations. For more details, we present and discuss results in three parts: genetic and forensic parameters of the major group, genetic and forensic parameters of minor groups, and genetic relatedness among populations in northeastern Thailand and other Asian groups.

Genetic and forensic parameters of major northeast Thai population (Lao Isan)

Forensic parameters

There was a total of 498 genotypes of Lao Isan (NE-TH) samples (males = 228 and females = 270). The HWE tests were performed on female samples and some loci were deviated from the HWE: DXS10148, DXS7132, DXS10074, DXS10079 and DXS7423, after applying Bonferroni correction ($0.05/12 = p < 0.004$) (Online Resource 7, ESM_7) The deviation of some loci could be a result of the effect of population substructure (Tao et al. 2018) which was observed in genetically heterogeneous Lao Isan groups (Srithawong et al. 2020).

We calculated allele frequencies of 12 X-STRs in both genders; there were 189 and 149 alleles with corresponding allelic frequencies ranging from 0.0019 to 0.4963 and 0.0043–0.4868 in females and males, respectively (Online Resource 2–3, ESM_2–3). Based on the exact p values, the frequencies of almost all of allele between genders were not significantly different, with the exception of DXS7132 and DXS10074 loci ($p < 0.0007$) (Table 3). We then calculated the allele frequencies of pooled 12 X-STR data of male and female (Online Resource 2–3, ESM_2–3); a total of 191 alleles with allele frequencies ranging from 0.0013 to 0.4935 were found. With high numbers of alleles (23 alleles) and the highest value for all forensic parameters, DXS10148 is the largest informative marker (PIC = 0.9416), while the least informative marker (PIC = 0.5007) is DXS7423 with 7 alleles and low values of forensic indexes (Online Resource 4, ESM_4). GD and PIC were lowest in DXS7423 (0.5888 and 0.5007, respectively) and highest in DXS10148 (0.9946 and 0.9416, respectively). The PDm and PDf spanned 0.5880 (DXS7423) to 0.9444 (DXS10148) and 0.7430 (DXS7423) to 0.9941 (DXS10148), respectively, while the ranges of MEC_Kruger, MEC_Kishida, MEC_Desmarais Trio and MEC_Desmarais Duo were 0.2999 (DXS7423)—0.8875 (DXS10148), 0.5008 (DXS7423)—0.9415 (DXS10148), 0.5007 (DXS7423)—0.9416 (DXS10148) and 0.3593 (DXS7423)—0.8926 (DXS10148), respectively (Online Resource 4, ESM_4). The combined MEC indices were much greater than 0.99999959 for deficiency CMEC_Kru, CMEC_Kishida, CMEC_Des-trio and duo cases (CMECDes-duo) (Table 1).

Table 1 Genetic variation and forensic efficiency parameters of 12 X-STR loci based on allelic frequencies in the ten studied populations

Ethnicity	Sample size	Total Allele	Gene diversity (GD)	Average PIC	PDm	PDf	MEC_Kruger	MEC_Kishida	MEC_Desmarais	MEC_Desmarais_duo	Average H_o_female	Average H_E_female	Loci departed from HWE[d]_female
Lao Isan	498	191	0.80793	0.7779	1	1	0.99999887	1	1	0.99999959	0.7673	0.81018	DXS10148, DXS7132, DXS10074, DXS10079 and DXS7423
Phutai	57	130	0.8086	0.7686	1	1	0.99999871	1	1	0.999999511	0.7632	0.7978	
Seak	12	76	0.9357	0.8761	0.99999996	1	0.99994335	0.999999819	0.99999982	0.999980705	0.8095	0.7372	
Nyaw	36	106	0.7927	0.7451	1	1	0.99999080	0.999999998	0.99999998	0.999996738	0.7451	0.7665	
Kalueang	41	115	0.7993	0.7551	1	1	0.99999551	0.99999991	0.99999999	0.999998417	0.7851	0.7914	
Bru	36	94	0.7648	0.7178	0.99999998	1	0.99996888	0.99999911	0.99999991	0.999988950	0.7902	0.7575	
Nyahkur	38	101	0.7878	0.7417	0.99999999	1	0.99999013	0.999999978	0.99999998	0.999996570	0.8389	0.7671	
Mon	82	133	0.8076	0.7713	1	1	0.99999805	1	1	0.999999313	0.7632	0.7915	DXS7423
Khmer	40	111	0.8110	0.7662	0.99999999	1	0.99999609	0.999999993	0.99999999	0.999998647	0.7320	0.7952	
Kuy	56	103	0.7819	0.7405	0.99999999	0.9999999	0.99998793	0.99999999	0.99999997	0.999995815	0.7803	0.7562	

H_O observed heterozygosity in females, H_E expected heterozygosity in females, *HWE* Hardy–Weinberg test in female samples, *GD* gene diversity, *PIC* polymorphism information content, *PDf* power of discrimination in females, *PDm* power of discrimination in males, *MEC_Kruger* mean paternity exclusion chance for autosomal STR markers in trios and complex kinship cases, *MEC_Kishida* mean paternity exclusion chance for X-chromosomal markers in trios involving daughters, *MEC_Desmarais* mean paternity exclusion chance for X-chromosomal markers in trios involving daughters (Desmarais version), *MEC_Desmarais duo* mean paternity exclusion chance for X-chromosomal markers in father/daughter duos

Linkage disequilibrium (LD) and four linkage groups (LG1-4)

The LD was estimated for all pairs of loci (Online Resource8, ESM_8) and their *p*-values for the LD exact tests between the male and female are shown Table 3. Technically, linkage expectations were only based on physical distances between loci; however, LD may also result from random genetic drift, founder effects, mutations, selection and population admixture or population stratification (Pereira et al. 2012). In this study, the significant associations were found both within LG and between LGs. After Bonferroni correction calculated from 66 pairwise comparisons, $(0.05/66 = p < 0.0007)$, there were six pairs of linkage in male: DXS8378(LG1)/DXS10135(LG1), DXS7132(LG2)/DXS10074(LG2), DXS10074(LG2)/DXS10079(LG2), DXS10074 (LG2)/DXS10146(LG4), DXS10101(LG3)/DXS10103(LG3) and DXS10103(LG3)/HPRTB(LG3) and five pairs in female: (DXS10074 (LG3)/DXS7132(LG3), DXS10146(LG4)/DXS7132(LG2), DXS10103(LG3)/DXS10101(LG3), HPRTB(LG3)/DXS10101(LG3) and HPRTB(LG3)/DXS10103(LG3) (Online Resource 8, ESM_8). The significant associations of linkage disequilibrium were previously reported in a northwestern Italian population (Robino et al. 2006), four Chinese ethnic groups (Han, Tibet, Uighur and Hui) (Yang et al. 2017) and Liaoning Manchu population from China (Xing et al. 2019). Interestingly, here we also found strong linkage groups between the STR trios within LG3 ($p < 0.05$); DXS10101/DXS10103, DXS10103/HPRTB and HPRTB/DXS10103 in both Lao Isan males and females (Online Resource 8, ESM_8).

Haplotype frequency

Due to linkage in the studied STRs, we also classified the haplotypes into four linkage groups based on their physical localizations (Diegoli 2015): linkage group 1 (LG1): DXS8378-DXS10135-DXS10148, linkage group 2 (LG2): DXS7132-DXS10074-DXS10079, linkage group 3 (LG3); DXS10101-DXS10103-HPRTB and linkage group 4 (LG4): DXS7423-DXS10134-DXS10146. The haplotypes and corresponding frequencies of these four linkage groups (LG1-4) in Lao Isan males along with forensic parameters are shown in Table 2, Online Resource 5–6; ESM_5–6. We found a total of 195, 115, 103 and 119 haplotypes in LG1 to LG4, respectively, with frequencies ranging from 0.057 to 0.0044. The most common haplotypes were observed in LG3 as H 31-19-11 with haplotype frequency at 5.7% (13 males) and H 31-19-12 (5.26% haplotype frequency in 12 males). Haplotype diversity (HD) ranged from 0.9985 (LG1) to 0.9851 (LG3) whereas PIC varied from 0.9941 (LG1) to 0.9804 (LG3) and the PDm and PDf were greater than 0.9 (Online Resource 6, ESM_6). The combined MEC (CMEC) indices based on haplotype frequencies for deficiency cases CMEC_Kru (0.999999792953847), CMEC_Kishida (0.999999992808508), normal trio CMEC_Des-trio (0.999999978393156) and duo cases CMEC_Des-duo (0.9999996836695) were equally high for Lao Isan males (Table 2 and ESM_5).

Genetic and forensic parameters of minor groups in Northeastern Thailand

Forensic parameters

A total of 398 genotypes of 12 X-STRs of nine ethnic groups were tested for HWE (Online Resource 7, ESM_7); there was no deviation from the HWE after Bonferroni correction ($p > 0.05/12 = 0.0041$) in each individual population with the exception of DXS7423 in Mon population (Online Resource 7, ESM_7). Allele frequencies of all loci between genders in all ethnic groups were not significantly different based on the exact *P*-value, excepting DXS7132 and DXS10074

Table 2 Genetic variation and forensic parameters of 12 X-STR loci based on haplotype frequencies

Ethnicity	Sample size	Total Haplotypes	Average HD	Average PIC	PDm	MEC_Kruger	MEC_Kishida	MEC_Desmarais	MEC_Desmarais_duo
Lao Isan	228	532	0.9913	0.9913	0.991317361	0.991317361	0.999999993	0.991317361	0.999999684
Phutai	38	126	0.9851	0.9577	0.99999746	0.999957611	0.99999691	0.99999708	0.999963931
Nyaw	19	59	0.9664	0.9096	0.999950759	0.999190175	0.999934556	0.999935543	0.999369832
Kalueang	22	74	0.9815	0.9335	0.999984329	0.999745096	0.999981403	0.999980583	0.999786397
Nyahkur	23	64	0.9525	0.9038	0.999955552	0.999286477	0.999942596	0.999941419	0.999434198
Mon	44	139	0.9855	0.9618	0.999998288	0.999970809	0.999997899	0.999998046	0.999975202
Khmer	26	89	0.9884	0.9481	0.999993978	0.999900538	0.999993147	0.999992787	0.999913907
Kuy	34	57	0.9484	0.9150	0.999959994	0.999318173	0.999946702	0.999947698	0.999471295

HD haplotype diversity, *PIC* polymorphism information content, *PDm* power of discrimination in males, *MEC_Kruger* mean exclusion chance for deficiency cases, *MEC_Kishida* mean exclusion chance for normal trios, *MEC_Desmarais Duo* mean exclusion chance in duo cases

Table 3 Differentiation test between all pairs of loci of male and female based on exact test

	Lao Isan	Phutai	Nyaw	Kalueang	Seak	Khmer	Mon	Nyahkur	Bru	Kuy
DXS8378	0.9319 (0.0044)	0.9523 (0.0012)	0.6879 (0.0025)	0.6463 (0.0037)	0.3837 (0.0042)	0.8503 (0.0017)	0.60230 (0.0067)	0.9348 (0.0009)	0.8557 (0.0023)	0.5553 (0.0059)
DXS10135	0.7818 (0.0181)	0.8628 (0.0037)	0.8899 (0.0042)	0.3372 (0.0083)	0.7437 (0.0048)	0.7556 (0.0075)	0.8521 (0.0046)	0.9413 (0.0029)	0.3630 (0.0051)	0.3305 (0.0059)
DXS10148	0.9334 (0.0092)	0.7780 (0.0116)	0.8106 (0.0033)	0.5772 (0.0090)	1.0000 (0.0000)	0.8163 (0.0063)	0.4954 (0.0087)	0.3954 (0.0048)	0.8972 (0.0030)	0.6448 (0.0094)
DXS7132	**0.0000 (0.0000)**	**0.0000 (0.0000)**	**0.0000 (0.0000)**	**0.0000 (0.0000)**	0.0051 (0.0007)	**0.0000 (0.0000)**	**0.0000 (0.0000)**	**0.0000 (0.0000)**	**0.0000 (0.0000)**	**0.0000 (0.0000)**
DXS10074	**0.0000 (0.0000)**	**0.0000 (0.0000)**	**0.0000 (0.0000)**	**0.0002 (0.0001)**	**0.0051 (0.0006)**	**0.0000 (0.0000)**	**0.0000 (0.0000)**	**0.0000 (0.0000)**	**0.0000 (0.0000)**	**0.0000 (0.0000)**
DXS10079	0.9734 (0.0039)	0.8176 (0.0036)	1.0000 (0.0000)	0.9960 (0.0006)	0.7902 (0.0040)	0.7505 (0.0048)	0.82120 (0.0060)	0.6765 (0.0023)	0.9417 (0.0014)	0.1745 (0.0053)
DXS10101	0.4693 (0.0169)	0.6987 (0.0061)	0.3933 (0.0079)	0.8555 (0.0032)	0.0211 (0.0019)	0.9683 (0.0026)	0.6500 (0.0047)	0.9457 (0.0024)	0.6654 (0.0119)	0.5355 (0.0044)
DXS10103	0.1311 (0.0128)	0.4752 (0.0059)	0.6675 (0.0036)	0.2752 (0.0067)	0.6445 (0.0026)	0.7086 (0.0020)	0.5705 (0.00630)	0.3516 (0.0068)	0.3885 (0.0060)	0.6585 (0.0049)
HPRTB	0.9629 (0.0026)	0.6150 (0.0062)	0.7105 (0.0036)	0.9909 (0.0007)	0.1050 (0.0039)	0.7740 (0.0042)	0.7843 (0.0062)	0.8431 (0.0026)	0.3328 (0.0054)	0.8764 (0.0013)
DXS7423	0.9721 (0.0024)	0.4564 (0.0042)	0.1167 (0.0022)	0.9303 (0.0013)	1.0000 (0.0000)	0.3695 (0.0000)	0.5230 (0.00690)	0.8907 (0.0030)	0.8726 (0.0028)	0.1549 (0.00470)
DXS10134	0.9237 (0.0050)	0.5552 (0.0073)	0.7377 (0.0075)	0.3151 (0.0061)	0.5609 (0.0044)	0.6603 (0.0056)	0.3815 (0.0093)	0.6476 (0.0063)	0.9009 (0.0043)	**0.0035 (0.0008)**
DXS10146	0.7745 (0.0107)	0.9109 (0.0052)	0.8915 (0.0052)	0.4097 (0.0108)	0.6106 (0.0032)	0.4230 (0.00730)	0.9331 (0.0035)	0.7520 (0.0056)	0.9134 (0.0034)	0.2176 (0.0077)

Values in parentheses were p values and bold letters indicates values significantly different from zero ($p < 0.05$)

(Table 3). We then pooled samples from both genders in each individual population and calculated allele frequencies and forensic parameters (Table 1, Online Resource 2–4, ESM_2–4).

Within these nine groups, the number of observed alleles varied from 2 to 21 across the different loci. The allele frequencies and the forensic parameters, including the PD in females (PDf) and males (PDm), PIC, GD, MEC, the CDP for females (CDPf) and males (CDPm), and the combined MEC (CMEC) indices for the 12 X-STR loci in individual groups are shown in Online Resource 2–4, ESM_2–4. The combined power of discrimination for both males (CPDm > 0.99999) and females (CPDf > 0.99999) as well as the CMEC indices (MEC_Krüger, MEC_Kishida, MEC_Desmarais Trio and MEC_Desmarais Duo) were high for all populations (Table 1 and Online Resource 4, ESM_4). The GD values of all 12 X-STR markers were greater than 0.7 in all populations, indicating variation of forensic markers and allelic diversity in the studied populations (Khacha-ananda et al. 2020) (Table 1). The PIC values of DXS7423 and DXS8378 markers in all groups were less than 0.6, in agreement with previous study of the northern Thai populations (Khacha-ananda et al. 2020). DXS10148 showed the highest value for all forensic parameters, followed by DXS10135 and DXS10101, suggesting that these loci were the most informative markers, while the least informative was DXS7423 (Online Resource 2–4, ESM_2–4). Previous studies also reported the less informative nature of DXS7423 in several Asian populations, i.e. Southern Han, Tibetan, Uighur and Hui (Yang et al. 2017), Guizhou Sui from southwest China (Guo et al. 2019), Central Thai and northern Thai populations (Vongpaisarnsin et al. 2016; Khacha-ananda et al. 2020) and Sri Lankan ethnic groups (Perera et al. 2021).

Linkage disequilibrium (LD) and four linkage groups (LG1-4)

Elevated LD is expected for populations which are either small, reproductively isolated or have low population growth (Ardlie et al. 2002), thus we estimated LD for all pairs of markers in seven populations, excluding Seak and Bru males due to small sample sizes (Online Resource 8, ESM_8). These linkage groups did not form stable haplotypes as indicated by LD of STRs within the groups and significant LD between the groups (Online Resource8, ESM_8). The high LD in populations, particularly Kuy's male is probably driven by either genetic drift or consanguineous groups, which was also supported by the presence of highly shared haplotypes without any unique haplotypes (Online Resource 8, ESM_8). Previous study on autosomal STRs of Kuy indicated their genetic divergence from other northern Thai groups; lesser gene flow (compared with their neighbor the

AA-speaking Khmer) and genetic drift could have promoted Kuy's genetic distinction (Chantakot et al. 2017).

The high values of gene and haplotype diversities (Tables 1 and 2), along with elevated LD values in AA-speaking Khmer and Mon (Online Resource 8, ESM_8), likely contributed to the admixed population structure observed in these groups. This pattern is consistent with findings in several Asian populations, such as the Sinhalese from Sri Lanka (Perera et al. 2021), and Liaoning Manchu from China (Xing et al. 2019). Previous genome-wide studies have provided evidence of South Asian genetic ancestry in several present-day Thai populations (Kutanan et al. 2021; Changmai et al. 2022). South Asian ancestry is identified as one of the parental sources shaping the current admixed genetic structure of Khmer and Mon in Thailand, although in minor proportions (~10–30%) (Kutanan et al. 2021; Changmai et al. 2022).

Haplotype frequencies

The forensic and statistical parameters of the four haplotypes (including PIC, HD, PDm, PDf, MEC_Kruger, MEC_Kishida, MEC_Desmarais and MEC_Desmarais_duo) are shown in Table 2, Online Resource 5–6, ESM_5-6; the PIC, PDm and PDf values for these haplotypes were greater than 0.9, while MEC values ranged above 0.8. Among the four linkage groups, LG1 proved more informative than other groups for almost all populations, mirrored by the haplotype diversity (Online Resource 5–6, ESM_5–6). Nevertheless, all four clusters showed a high haplotype diversity for all seven ethnicities. Approximately 75% to 85% of haplotypes were identified as unique haplotypes; in exception the Kuy did not show any unique haplotypes (Online Resource 5–6, ESM_5–6). The most common haplotypes are different among the studied populations; H 10-26-19 in LG1 is common in Kaleung (18.18%), while H 14-17-19 and H 13-17-20 in LG2 are prevalent haplotypes with a frequency of 26.09% and 21.74%, respectively in Nyahkur. In LG3, the most common haplotypes were observed: H 32-16-13 in Phutai with 15.79%, H31-19-11 in Nyaw with 21.05%, and H33-16-13 in Mon (11.36%), (Online Resource 5–6, ESM_5–6).

Genetic variation and genetic relatedness within northeastern Thai groups and between Asian populations

Results of Analysis of Molecular Variance (AMOVA) indicated that total genetic variation among populations accounts for 3.2%. Interestingly, the TK groups (3.13%, $p < 0.05$) exhibited small greater genetic heterogeneity than AA groups (2.97%, $p < 0.05$) (Table 4). The genetic variation showed non-significant differences among the TK and AA language families, coupled with much lower variation among groups than the variation among populations within each group, indicating that language families do not correspond to genetic structure. In addition, a direct comparison of Lao Isan vs. minor TK-groups, and Lao Isan vs. minor AA-groups, showed no significant differences between groups.

Among 45 pairwise comparisons of genetic distances (F_{st}) calculated from ten populations, there were 41 pairs (91.11%) with statistical differences ($p < 0.05$) and 4 pairs (8.89%) without significant differences ($p > 0.05$). No significant difference was observed between the pair of Phutai-Nyaw ($Fst = 0.00625$), Phutai-Kalueang ($Fst = 0.00682$), Phutai-Khmer ($Fst = 0.00149$) and Khmer-Kalueang ($Fst = 0.00919$). Subsequently, we carried out multidimensional scaling (MDS) plot based on F_{st} matrix to visualize genetic relationship among populations (Fig. 1). Based on MDS dimension 1, Seak, Lao Isan, Kuy and Nyahkur were quite distinct from other populations but Nyaw and Bru were separated from the others based on dimension 2 and 3, respectively. The Mon, Kalueang, Khmer and Phutai were positioned in the center of the plots, mirroring their close genetic relationship (Figs. 1A and B). In general, population clustering is not consistent with language family, in agreement with AMOVA results (Table 4).

Table 4 AMOVA results

| Groups	Number of groups	Number of populations	% of variance		
			Within populations	Among population within groups	Among groups
All sample	1	10	96.80	3.20*	
TK	1	5	96.87	3.13*	
AA	1	5	97.03	2.97*	
TK/AA	2	10	96.42*	3.00*	0.58
Lao Isan/TK	2	6	96.20*	2.78*	1.01
Lao Isan/AA	2	5	96.75*	1.62*	1.63

*Indicates $p < 0.05$
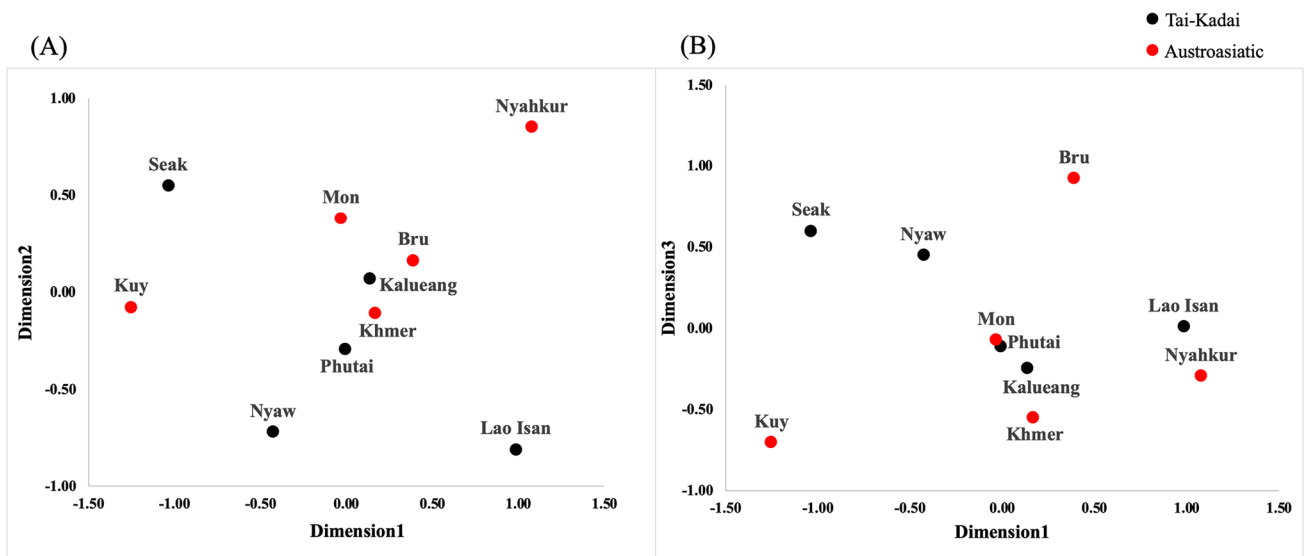
Fig. 1 The three-dimensional MDS plot based on F_{st} distance matrix for 10 populations for **a** dimension 1 vs. 2 and **B** dimension 1 vs.3. Black and red dots represent Tai-Kadai and Austroasiatic-speaking populations, respectively. The stress value is 0.0409

To reveal genetic relationship among populations worldwide, the allele frequencies of 12 X-STRs of ten northeastern Thai groups were compared with available data from 17 populations (Vongpaisarnsin et al 2016; Khacha-ananda et al. 2020; Shrivastava et al. 2015; Uchigasaki et al. 2013; Zeng et al. 2011; Sufian et al. 2017; Elakkary et al. 2014; Almarri and Lootah 2018; Salvador et al. 2018; Messaoudi et al. 2019; Poulsen et al. 2016). We constructed

a neighbor-joining (NJ) tree and the first two dimensions of MDS based on pairwise Nei's Da genetic distance averaged over 12 X-STRs (Fig. 2A and B). In general, the Lao Isan population was clustered with other northeastern Thai minorities with the divergence of Seak (Fig. 2A). Distinction of Seak from the other groups in Thailand was also reflected by linguistic evidence. In contrast to other TK groups in Thailand who speak the Southwestern Tai branch,
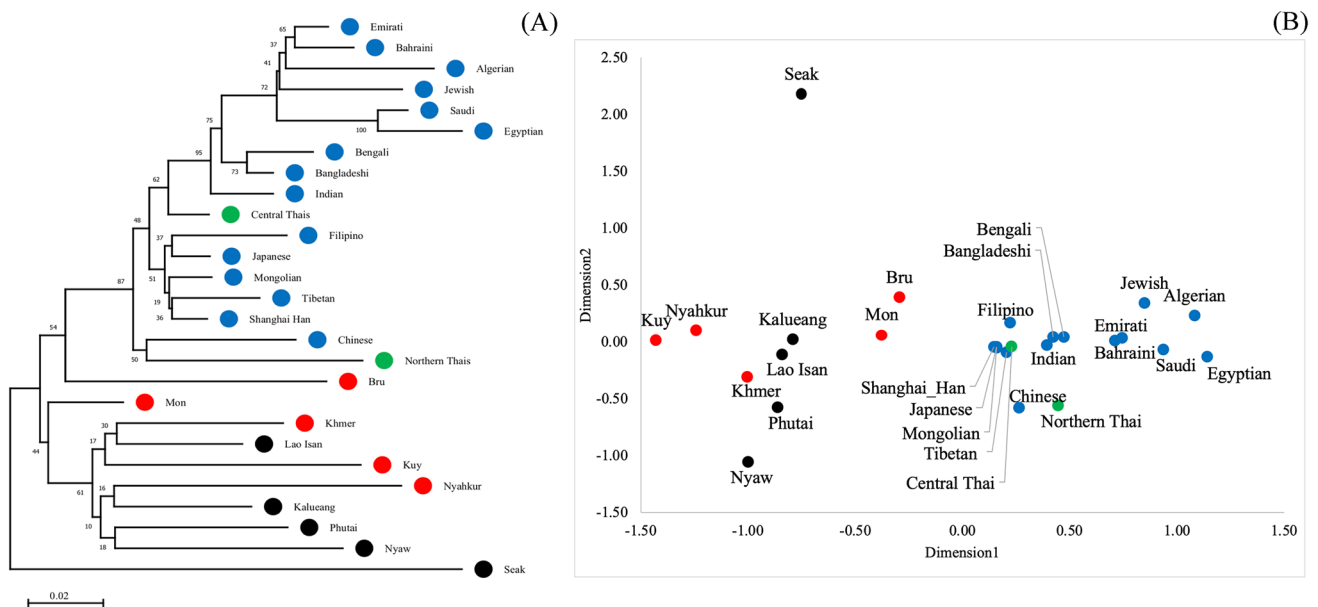
Fig. 2 Genetic relationship among 27 populations based on Nei's genetic distances. **A** the NJ tree and **B** the two dimensional MDS plots with stress value at 0.0973. Black and red dots represent Tai-Kadai and Austroasiatic-speaking populations from northeastern Thailand, respectively, whereas green and blue dots represent Thai populations from other regions and other Asian populations, respectively

the language of Seak is of the Northern Tai branch in TK family that was spoken mainly by the Tai in Gwangsi Province of China (Eberhard et al. 2020). Historically, the Seak originated from the area of Thua Thien-Hue Province in Vietnam and then migrated westward into central Laos and settled scattered there (Schliesinger 2003). In 1851 CE, some groups of Seak were taken to present-day Northeastern Thailand from the Khummuan Province of Laos by the Siamese army as war prisoners (Schliesinger 2000, 2001, 2003). Previous mtDNA studies revealed low genetic diversity and large divergence of Seak that were probably influenced by genetic drift during population movement (Kutanan et al. 2014; 2017). However, the Y chromosome showed an opposing genetic result, indicating the contrasting maternal and paternal genetic history of Seak.

The AA-speaking Mon and Bru are positioned in between northeastern Thai clade and clades of other populations (Fig. 2A); Mon and Bru were laid far away from northeastern Thai group shifting to cloud of other Asian groups (North and South Asians) in dimension 1 of MDS (Fig. 2B), indicating genetic difference of both from other northeastern Thais. In agreement with previous studies, autosomal STRs indicated genetic differentiation of Bru (Inturri et al. 2011) while genome-wide data supported unique genetic structure of Mon that resulted from admixture with South Asian group (Kutanan et al. 2021).

In agreement with autosomal STRs and genome-wide data (Srithawong et al. 2020; Kutanan et al. 2021), genetic difference according to geographic locations of major Thai groups were observed using X-STRs; almost all of the northeastern Thai populations were clustered together and separated from the northern Thai, which were genetically closer to Chinese, and central Thai who were more related to South Asian populations (Indian, Bengali and Bangladeshi) than the other Thais These results are consistent with several lines of investigation that indicated South Asian genetic ancestry in the central Thai group (Srithawong et al. 2020, 2021; Kutanan et al. 2017, 2019).

Conclusion

Previous reports on X-STRs data from Thailand were primarily focused on the North and Central regions. This study expands upon existing research by establishing additional X-STRs data in the northeastern region, utilizing the Investigator® Argus X-12 kit. Forensic parameters within the major group, Lao Isan, demonstrated that all loci are highly polymorphic and informative, rendering this marker set a valuable tool for forensic investigations in northeastern Thailand. For the minor ethnic groups, it is important to note a limitation in this study—the low sample size. Caution is advised in applying these findings in forensic

applications and interpreting them in anthropological studies. Nevertheless, a rigorous recruitment procedure for sample collection, coupled with an overall agreement with previous studies that employed various genetic markers, suggests that the results are representative of these smaller populations. The present results also support regional genetic differences among major groups in each region: northeastern Thai, Northern Thai, and Central Thai. Consequently, different haplotype frequency databases should be considered for forensic purposes. Additionally, more exhaustive sampling in this area would be beneficial for further forensic and anthropological genetic studies. Furthermore, incorporating forensic X-STRs data from southern Thailand could strengthen the forensic X-STR database for the entire country.

Author contributions Suparat Srithawong**,** Kanha Muisuk, Nonglak Prakhun and Wibhu Kutanan collected the samples. Suparat Srithawong**,** Kanha Muisuk and Nonglak Prakhun extracted DNA and performed genotyping. Suparat Srithawong designed the project, analyzed data and drafted the manuscript. Nisarat Tungpairojwong and Wibhu Kutanan designed the project, drafted and edited the manuscript.

Data availability Raw genotypes of 12 X-STRs of females and males are shown in Online Resource 1, (ESM_1).

Declarations

Conflict of interest The authors declare no conflict of interest.

Ethical approval The rights of participants and their identity have been protected during the whole process of this research. All experiments were performed in accordance with relevant guidelines and regulations based on the experimental protocol on human subjects which was approved by the Khon Kaen University Ethic Committee (Protocol No. HE642161).

References

Almarri MA, Lootah RA (2018) Allelic and haplotype diversity of 12 X-STRs in the United Arab Emirates. Forensic Sci Int Genet 33:e4–e6. https://doi.org/10.1016/j.fsigen.2017.12.013

Ardlie KG, Kruglyak L, Seielstad M (2002) Patterns of linkage disequilibrium in the human genome. Nat Rev Genet 3:299–309. https://doi.org/10.1038/nrg777

Asamura H, Sakai H, Kobayashi K, Ota M, Fukushima H (2006) Mini X-STR multiplex system population study in Japan and

application to degraded DNA analysis. Int J Legal Med 120:174–181. https://doi.org/10.1007/s00414-005-0074-6

Baeta M, Núñez C, Aznar JM, Sosa C, Casalod Y, Bolea M, Gonzlez-Andrade F, de Pancorbo MM, Martinez-Jarreta B (2013) Analysis of 10 X-STRs in three population groups from Ecuador. Forensic Sci Int Genet 7:e19-20. https://doi.org/10.1016/j.fsigen.2012.08.004

Botstein D, White RL, Skolnick M, Davis RW (1980) Construction of a genetic linkage map in man using restriction fragment length polymorphisms. Am J Hum Genet 32:314–331

Changmai P, Jaisamut K, Kampuansai J, Kutanan W, Altınışık NE, Flegontova O, Inta A, Yüncü E, Boonthai W, Pamjav H, Reich D, Flegontov P (2022) Indian genetic heritage in Southeast Asian populations. PLoS Genet 18:e1010036. https://doi.org/10.1371/journal.pgen.1010036

Chantakot P, Pittayaporn P, Srithongdaeng K, Srithawong S, Kutanan W (2017) Genetic divergence of Austroasiatic speaking groups in the Northeast of Thailand: a case study on Northern Khmer and Kuy. Chiang Mai J Sci 44:1279–1294

Desmarais D, Zhong Y, Chakraborty R, Perreault C, Busque L (1998) Development of a highly polymorphic STR marker for identity testing purposes at the human androgen receptor gene (HUMARA). J Forensic Sci 43:1046–1049

Diegoli TM (2015) Forensic typing of short tandem repeat markers on the X and Y chromosomes. Forensic Sci Int Genet 18:140–151. https://doi.org/10.1016/j.fsigen.2015.03.013

Eberhard DM, Simons GF, Fennig CD (2020) Ethnologue: languages of the World, 23rd edn. SIL International, Dallas

Elakkary S, Hoffmeister-Ullerich S, Schulze C, Seif E, Sheta A, Hering S, Edelmann J, Augustin C (2014) Genetic polymorphisms of twelve X-STRs of the investigator Argus X-12 kit and additional six X-STR centromere region loci in an Egyptian population sample. Forensic Sci Int Genet 11:26–30. https://doi.org/10.1016/j.fsigen.2014.02.007

Excoffier L, Lischer L (2010) Arlequin software 3.5.: a new series of programs to perform population genetics analyses under Linux and Windows. Mol Ecol Resour 10:564–567. https://doi.org/10.1111/j.1755-0998.2010.02847.x

Garcia FM, Bessa BGO, dos Santos EVW, Pereira JDP, Alves LNR, Vianna LA, Casotti MC, Trabach RSR, Stange VS, Meira DD, Louro ID (2022) Forensic applications of markers present on the X chromosome. Genes 13:1597. https://doi.org/10.3390/genes13091597

Gomes I, Pinto N, Antão-Sousa S, Gomes V, Gusmão L, Amorim A (2020) Twenty years later: a comprehensive review of the X chromosome use in forensic genetics. Front Genet 11:926. https://doi.org/10.3389/fgene.2020.00926

Guo J, Ji J, He G, Ren Z, Zhang H, Wang Q, Zhang H, Wang Q, Yang M, Nabijiang Y, Zhang Z, Zhang J, Huang J, Wang CC (2019) Genetic structure and forensic characterization of 19 X-chromosomal STR loci in Guizhou Sui population. Ann Hum Biol 46:246–253. https://doi.org/10.1080/03014460.2019.1623911

Higham C (2014) Early mainland Southeast Asia: from first humans to Angkor. River Books, Bangkok

Illescas MJ, Pérez A, Aznar JM, Valverde L, Cardoso S, Algorta J, de Pancorbo MM (2012) Population genetic data for 10 X-STR loci in autochthonous Basques from Navarre (Spain). Forensic Sci Int Genet 6(5):e146-148. https://doi.org/10.1016/j.fsigen.2012.02.014

Inturri S, Menegon S, Amoroso A, Torre C, Robino C (2011) Linkage and linkage disequilibrium analysis of X-STRs in Italian families. Forensic Sci Int Genet 5:152–154. https://doi.org/10.1016/j.fsigen.2010.10.012

Khacha-ananda S, Mahawong P (2020) Genetic analysis of 12 X-short tandem repeats loci in a northern Thai population. Med Sci Law 1:34–43. https://doi.org/10.1177/0025802420965000

Kishida T, Wang W, Fukuda M, Tamaki Y (1997) Duplex PCR of the Y-27H39 and HPRT loci with reference to Japanese population data on the HPRT locus. Jpn J Legal Med 51(2):67–69

Krüger J, Fuhrmann W, Lichte KH, Steffens C (1968) On the utilization of erythrocyte acid phosphatase polymorphism in paternity evaluation. Dtsch Z Gesamte Gerichtl Med 64(2):127–146

Kutanan W, Ghirotto S, Bertorelle G, Srithawong S, Srithongdaeng K, Pontham KD (2014) Geography has more influence than language on maternal genetic structure of various northeastern Thai ethnicities. J Hum Genet 59(9):512–520. https://doi.org/10.1038/jhg.2014.64

Kutanan W, Kampuansai J, Srikummool M, Kangwanpong D, Ghirotto S, Brunelli A, Stoneking M (2017) Complete mitochondrial genomes of Thai and Lao populations indicate an ancient origin of Austroasiatic groups and demic diffusion in the spread of Tai-Kadai languages. Hum Genet 136(1):85–98. https://doi.org/10.1007/s00439-016-1742-y

Kutanan W, Kampuansai J, Brunelli A, Ghirotto S, Pittayaporn P, Ruangchai S, Schröder R, Macholdt E, Srikummoo M, Kangwanpong D, Arias L, Stoneking M (2018) New insights from Thailand into the maternal genetic history of Mainland Southeast Asia. Eur J Hum Genet 26(6):898–911. https://doi.org/10.1038/s41431-018-0113-7

Kutanan W, Kampuansai J, Srikummool M, Brunelli A, Ghirotto S, Arias L, Macholdt E, Hübner A, Schröder R, Stoneking M (2019) Contrasting paternal and maternal genetic histories of Thai and Lao populations. Mol Biol Evol 36(7):1490–1506. https://doi.org/10.1093/molbev/msz083

Kutanan W, Liu D, Kampuansai J, Srikummool M, Srithawong S, Shoocongdej R, Sangkhano S, Ruangchai S, Pittayaporn P, Arias L, Stoneking M (2021) Reconstructing the human genetic history of Mainland Southeast Asia: insights from genome-wide data from Thailand and Laos. Mol Biol Evol 38(8):3459–3477. https://doi.org/10.1093/molbev/msab124

Lang Y, Guo F, Niu Q (2019) StatsX v20: the interactive graphical software for population statistics on X-STR. Int J Legal Med 133(1):39–44. https://doi.org/10.1007/s00414-018-1824-6

Li C, Ma T, Zhao S, Zhang S, Xu J, Zhao Z et al (2011) Development of 11 X-STR loci typing system and genetic analysis in Tibetan and Northern Han populations from China. Int J Legal Med 125(5):753–756. https://doi.org/10.1007/s00414-011-0592-3

Liu QL, Lv DJ, Sun HY, Lu HL, Wu XL, Wu XY (2007) Development and forensic application of a pentaplex X-STR loci typing system. Yi Chuan 29:1459–1462 (**PMID: 18065380**)

Messaoudi SA, Babu SR, Alsaleh AB, Albajjah M, Alsnan N (2019) Genetic diversity and haplotypic structure of a Saudi population sample using Investigator Argus X-12 amplification kit. bioRxiv [Preprint]. https://www.biorxiv.org/content/https://doi.org/10.1101/760819v3

Mishra PP (2010) The history of Thailand. ABC-CLIO/Greenwood, Santa Barbara

Myers RL (2005) The Isan saga: the inhabitants of rural northeast Thailand and their struggle for identity, equality and acceptance (1964–2004). Master's thesis, San Diego State University, San Diego

Nei M, Roychoudhury AK (1974) Sampling variances of heterozygosity and genetic distance. Genetics 76(2):379–390

Pepinski W, Niemcunowicz-Janica A, Skawronska M, Koc-Zorawska E, Janica J, Berent J, Soltyszewski I (2007) X-chromosomal polymorphism data for the ethnic minority of Polish Tatars and the religious minority of Old Believers residing in northeastern Poland. Forensic Sci Int Genet 1(2):212–214. https://doi.org/10.1016/j.fsigen.2007.01.007

Pereira V, Gusmão L, Valente C, Pereira R, Carneiro J, Gomes I, Morling N, Amorim A, Prata MJ (2012) Refining the genetic portrait

of Portuguese Roma through X-chromosomal markers. Am J Phys Anthropol 148(3):389–394. https://doi.org/10.1002/ajpa.22061

Perera N, Galhena G, Ranawaka G (2021) X-chromosomal STR based genetic polymorphisms and demographic history of Sri Lankan ethnicities and their relationship with global populations. Sci Rep 11:1. https://doi.org/10.1038/s41598-021-92314-9

Poulsen L, Tomas C, Drobnič K, Ivanova V, Mogensen HS, Kondili A, Miniati P, Bunokiene D, Jankauskiene J, Pereira V, Morling N (2016) NGMSElect™ and Investigator(®) Argus X-12 analysis in population samples from Albania, Iraq, Lithuania, Slovenia, and Turkey. Forensic Sci Int Genet 22:110–112. https://doi.org/10.1016/j.fsigen.2016.02.004

Robino C, Giolitti A, Gino S, Torre C (2006) Development of two multiplex PCR systems for the analysis of 12 X-chromosomal STR loci in a northwestern Italian population sample. Int J Legal Med 120:315–318. https://doi.org/10.1007/s00414-006-0115-9

Salvador JM, Apaga DLT, Delfin FC, Calacal GC, Dennis SE, De Ungria MCA (2018) Filipino DNA variation at 12 X-chromosome short tandem repeat markers. Forensic Sci Int Genet 36:e8–e12. https://doi.org/10.1016/j.fsigen.2018.06.008

Schaffner SF (2004) The X chromosome in population genetics. Nat Rev Genet 5:43–51. https://doi.org/10.1038/nrg1247

Schliesinger J (2000) Ethnic groups of Thailand: non-Tai-speaking peoples. White Lotus Press, Bangkok

Schliesinger J (2001) Tai Group of Thailand: introduction and overview. White Lotus Press, Bangkok

Schliesinger J (2003) Ethnic groups of Laos: profiles of Austro-Thai-speaking peoples. White Lotus Press, Bangkok

Shin SH, Yu JS, Park SW, Min GS, Chung KW (2005) Genetic analysis of 18 X-linked short tandem repeat markers in Korean population. Forensic Sci Int 147:35–41. https://doi.org/10.1016/j.forsciint.2004.04.012

Shrivastava P, Jain T, Gupta U, Trivedi VB (2015) Genetic polymorphism study on 12 X STR loci of investigator Argus X STR kit in Bhil tribal population of Madhya Pradesh, India. Leg Med (tokyo, Japan). 17(3):214–217. https://doi.org/10.1016/j.legalmed.2014.11.004

Srithawong S, Srikummool M, Pittayaporn P, Ghirotto S, Chantawannakul P, Sun J, Eisenberg A, Chakraborty R, Kutanan W (2015) Genetic and linguistic correlation of the Kra-Dai-speaking groups in Thailand. J Hum Genet 60(7):371–380. https://doi.org/10.1038/jhg.2015.32

Srithawong S, Muisuk K, Srikummool M, Mahasirikul N, Triyarach S, Sriprasert K, Kutanan W (2020) Genetic structure of the ethnic Lao groups from mainland Southeast Asia revealed by forensic microsatellites. Ann Hum Genet 84(5):357–369. https://doi.org/10.1111/ahg.12379

Srithawong S, Muisuk K, Srikummool M, Kampuansai J, Pittayaporn P, Ruangchai S, Liu D, Kutanan W (2021) Close genetic relationship between central Thai and Mon people in Thailand revealed by autosomal microsatellites. Int J Legal Med 135(2):445–448. https://doi.org/10.1007/s00414-020-02290-4

Sufian A, Hosen MI, Fatema K, Hossain T, Hasan MM, Mazumder AK, Akhteruzzaman S (2017) Genetic diversity study on 12 X-STR loci of investigator® Argus X STR kit in Bangladeshi population. Int J Legal Med 131(4):963–965. https://doi.org/10.1007/s00414-016-1513-2

Takezaki N, Nei M, Tamura K (2014) POPTREEW: webersion of POPTREE for constructing population trees from allele frequency data and computing some other quantities. Mol Biol Evol 31(6):1622–1624. https://doi.org/10.1093/molbev/msu093

Tao R, Zhang J, Bian Y, Dong R, Liu X, Jin C, Zhu R, Zhang S, Li C (2018) Investigation of 12 X-STR loci in Mongolian and Eastern Han populations of China with comparison to other populations. Sci Rep 98(1):4287. https://doi.org/10.1038/s41598-018-22665-3

Tetzlaff S, Wegener R, Lindner I (2012) Population genetic investigation of eight X-chromosomal short tandem repeat loci from a northeast German sample. Forensic Sci Int Genet 6:e155-156. https://doi.org/10.1016/j.fsigen.2012.03.007

Than KZ, Muisuk K, Woravatin W, Suwannapoom C, Srikummool M, Srithawong S, Lorphengsy S, Kutanan W (2022) Genetic structure and forensic utility of 23 autosomal STRs of the ethnic Lao Groups From Laos and Thailand. Front Genet 13:954586. https://doi.org/10.3389/fgene.2022.954586

Uchigasaki S, Tie J, Takahashi D (2013) Genetic analysis of twelve X-chromosomal STRs in Japanese and Chinese populations. Mol Biol Rep 40(4):3193–3196. https://doi.org/10.1007/s11033-012-2394-1

Vongpaisarnsin K, Boonlert A, Rasmeepaisarn K, Dangkao P (2016) Genetic variation study of 12 X chromosomal STR in central Thailand population. Int J Leg Med 130:1497–1499. https://doi.org/10.1007/s00414-016-1363-y

Xing J, Adnan A, Rakha A, Kasim K, Noor A, Xuan J, Zhang X, Yao J, McNevin D, Wang B (2019) Genetic analysis of 12 X-STRs for forensic purposes in Liaoning Manchu population from China. Gene 683:153–158. https://doi.org/10.1016/j.gene.2018.10.020

Yang X, Zhang X, Zhu J, Chen L, Liu C, Feng X, Chen L, Wang H, Liu C (2017) Genetic analysis of 19 X chromosome STR loci for forensic purposes in four Chinese ethnic groups. Sci Rep 7:42782. https://doi.org/10.1038/srep42782

Zeng X, Ren Z, Chen J, Lv D, Tong D, Chen H, Sun HY (2011) Genetic polymorphisms of twelve X-chromosomal STR loci in Chinese Han population from Guangdong Province. Forensic Sci Int Genet 5(4):e114-116. https://doi.org/10.1016/j.fsigen.2011.03.005