in 2005 (16), various unique features of this genome had been revealed (17). One of striking findings is the abundance of tandem arrayed tRNA genes. Inergenic regions adjoin to tRNA genes in each tRNA array units contain non-coding short tandem repeats (STRs) (18) shown a high degree of polymorphism. STRs are DNA sequences of about 2 to 8 base pairs that are repeated up to 40 times in a head-to-tail manner (19). Currently, a system based on 6 tRNA-linked STRs was established for typing *E. histolytica* strains (15). A recent study of virulent and avirulent strains of *E. histolytica* from 37 Japanese *E. histolytica* samples revealed unique genotypes in asymptomatic isolates (19). However, a phylogenetic correlation remains to be established in order to understand the relationship among *E. histolytica* genotypes and the clinical symptoms. In the first year of this study, a phylogenetic assay of *E. histolytica* strains based on 6 tRNA-linked STRs was developed to elucidate the questions described as above. Furthermore, we attempted to correlate the virulent factor AIG1 gene of *E. histolytica* strains with the clinical outcome and also established some clinical *E. histolytica* strains for further study.

Materials and methods

Clinical samples and cultivation

Stool samples were collected from 122 *Entamoeba histolytica* PCR confirmed patients and then extracted DNA for the 6 STR typing. However, only 93 sampless could complete the 6 STR typing, including 22 strains from Taiwanese patients, 24 from Institutionized patients and 47 from alien workers. *E. histolytica* strain HM-1:IMSS (ATCC 30459) was grown in Diamond's TYI-S-33 medium (20) at 37°C.

Culture Methods: Xenic culture

The clinical isolates were cultured in xenic condition using Robinson's medium (21). Brief, clinical stool sample was xenically cultured at 35.5°C in basal amoebic (BR) medium with 200 μl rice starch solution (5 mg/ml) and 120 μg erythromycin for 24 hours. In the second day, remove the BR medium and keep the stool sample and rice starch powder layer in the bottle. Add the phthalate solution, bacto-peptone solution, erythromycin in to the bottle and fill with the BRS medium (BR medium with 10 % serum) to the bottle neck. The next day, take about 10 μl faeces and the starch for checking the *E. histolytica* trophozoites. If cannot see them, incubate the tube for another 24 hours. If trophozoites exist, transfer 100μl of faecal-starch layer into per new bottles and fill the bottles with phthalate solution, bacto-peptone solution, erythromycin and BRS medium (the complete BRS medium) to bottle neck. After cultured days, if there are a great mounts of trophozoites, transfer to the monoxenic culture medium.

Culture Methods: Monoxenic culture from trophozoites

These clinical isolates were cultured in monoxenic condition using yeast extract–iron–maltose–dihydroxyacetone-serum (YIMDHA-S) medium supplemented with *E. coli* (22, 23). Brief, filter the suspension from xenic culture bottles by BD filter (40 μm funnel) (put on the 50 ml tube) and transfer to glass culture tube at 35.5°C 30 minute to 1 hour for attachment. Then pipette out the BRS medium and wash the sediment by centrifugation (1200 rpm, three minute) for three times with fresh BI-S-33 medium or LYI-2 medium. The sediment is inoculated in to fresh YIMDHA-S medium (5.5 ml) containing 15% adult bovine serum, potassium penicillin G (1000 units/ml), gentamycin (130 units/ml), streptomycin sulfate (1 mg/ml). The culture tube is inoculated at stand upright position for 30 minute at 35.5°C. Centrifuge the tube (1200 rpm, 3 minute) and remove the supernatant gently. Add 5.5 ml the complete YIMDHA-S medium in the tube. The culture tube is incubated at stand upright at 35.5°C about 3 days. Observe the growth of amoeba and monitor contamination. If culture medium to be derby, on ice five minute then wash again and fill new medium. If there is a great mount of cells, put the tube on ice five minute and transfer 1 ml to 4 ml in to the tube with fresh complete YIMDHA-S medium.

Polymerase chain reaction (PCR) and DNA sequencing

Total DNA from the stool samples was automatically extracted by using the MagNA Pure Systems (Roche Diagnostics) according to the manufacturer's instructions. A total of 100 ml DNA was produced and stored at -20°C. The STR fragments were amplified using 6 *E. histolytica*-specific tRNA-linked STR primers (DA-H, AL-H, NK2-H, RR-H, SQ-H, and STGAD-H) under the conditions previously described (15). The amplified PCR products were separated using 1.5%

agarose gel and purified using QIAquick Gel Extraction Kit. Sequence analysis was performed by Genomics BioSci & Tech company (Taiwan) using the forward or reverse *E. histolytica*-specific tRNA-linked STR primers (15). Nucleotide sequences were analyzed using the BioNumerics v6.5 software (Applied Meths, Belgium) to identify the STRs.

Statistical analyses

One-way ANOVA and Post Hoc tests were used to evaluate the significance of the locus type diversity among the strains from Taiwan, Indonesia and Japan. The reliability of this phylogenetic tree built by UPGMA method was evaluated by cophenetic correlation coefficient.

Results

Analysis of Six STR locus types and genotypes of *E. histolytica*

A total of 122 *E. histolytica* strains were collected for the 6 tRNA-linked STR genotyping (locus SD, SQ, AL, NK2, RR and DA). However, only 93 strains could complete the 6 STR typing, including 22 strains from Taiwanese patients, 24 from Institutionized patients and 47 from alien workers. A genotype was assigned by combining the 6 STR locus types (Figure 1). The STR locus types had been reported previously (19, 24) or personally communicated by Dr. C. Graham Clark, London School of Hygiene and Tropical Medicine. The STR locus types were named according to his nomenclature, and newly identified STR locus types and genotypes were assigned alphanumerical codes beginning with the letter "T" and "EH" to indicate their Taiwan origin and new *E. histolytica* strains.

A total of 57 genotypes were identified (Table 1 and 2). There were 2 S^{TGA}-D locus types, 5 D-A locus types, 9 A-L locus types and 32 N-K locus types were found, whereas no new types in S-Q and R-R locus. The new STR locus type T1SD in S^{TGA}-D locus, T1DA in D-A locus, T3AL, T5AL, T10AL in A-L locus and T1NK, T2NK, T4NK, T7NK, T8NK, T11NK in N-K2 locus were found in Taiwanese strains; T2SD in S^{TGA}-D locus, T3DA, T4DA, T5DA in D-A locus, T6AL, T7AL, T9AL in A-L locus and T13NK~T27NK, T29NK~T39NK in N-K2 locus in alien worker strains. The new STR locus types were found in Taiwanese strains included: T2DA in D-A locus, T1AL, T2AL, T4AL in A-L locus and T5NK, T9NK in N-K2 locus were found in both Taiwanese and alien worker strains. Among the 6 STR loci, N-K2 locus had the most new locus types, and then A-L locus. Interestingly, almost asymptomatic strains from Indonesian contained conserved locus type 6SQ but had many different locus types in N-K2 locus.

Comparison of the diversity of the locus types among the six STR Loci

In order to understand the diversity of the locus types among the six STR Loci, we compare our locus types to other previously reported locus types (19, 24) (Table 3). The diversity of the locus types in each STR Loci was the ratio of type numbers to strain numbers. The ratio more close to 1 indicates the diversity was getting higher, whereas close to 0 indicates getting lower diversity. The diversity of the locus types from Japanese strains was A-L > N-K2 > D-A = R-R > S^{TGA}-D = S-Q locus (19), whereas from Bangladesh was N-K2 > S^{TGA}-D > R-R (24). In Taiwanese and alien worker strain, the diversity of the locus types was N-K2 > A-L > D-A > S^{TGA}-D > R-R > S-Q locus. Comparing these three countries, the N-K2 locus had highest diversity. One-way ANOVA and Post Hoc tests were used to evaluate the significance of the diversity from these three countries and indicated that N-K2 locus had higher diversity compared to S^{TGA}-D、S-Q and R-R locus ($p < 0.05$). A-L locus and D-A locus showed no significant difference. Overall, the diversity of genotypes from Taiwanese and alien worker strains was higher that Japanese and Bangladesh. However, if we only compared the genotype diversity of Taiwanese (0.609) and Japanese (0.622) strains, both were similar. The genotypes from alien worker strains especially from Indonesian were highly diverse.

Development of the phylogenetic analysis based on six tRNA-linked STR loci

After analyses of the DNA sequences of the 6 tRNA-linked STR loci, each STR locus contained conserved and variable regions. The variable regions of one STR locus from different *E. histolytica* strains were varied in numbers and type of their tandem repeats (motifs), and nucleotide mutation in sequence. A phylogenetic analysis method was developed based on the sequence of the variable regions of the six tRNA-linked STR loci. The variable regions were batchly cropped from the DNA

sequence of each amplified STR fragment by using BioNumerics v6.5 software (Applied Meths, Belgium). The pairs of 9-17 bp conserved sequences listed in Table 4 were used to locate the variable regions from each STR sequence for cropping.

The similarity between *E. histolytica* strains were calculated STR motifs of variable regions by were Pearson correlation coefficient in BioNumerics v6.5 software (Applied Meths, Belgium), and then transferred into a similar matrix. The DNA sequences of variable regions from *E. histolytica* strains were aligned and transformed to a similar matrix, too. Finally, Fifteen variable regions were cropped from 6 tRNA-linked STR loci, and produced 15 similar matrixes by the analyses of STR motifs and another 15 similar matrixes by sequence alignments. These thirty matrixes were then averaged to produce a final matrix. The final matrix used for cluster analysis and developed a phylogenetic tree by Unweighted Pair Group Method with Arithmetic Mean (UPGMA) method. The flowchart of the phylogenetic development was shown in Figure 2

Analysis of the phylogenetic tree

A phylogenetic tree had been established based on the final matrix of the six tRNA-linked STR loci from 93 Taiwanese and alien worker *E. histolytica* strains, 3 standard strains and 37 Japanese strains by UPGMA method. The *E. histolytica* strains could be grouped into three Clusters, Cluster A, B and C, shown in Figure 3. Cluster A mainly contained 30 Indonesian strains and most were asymptomatic. Cluster B mainly contained 12 Taiwanese and 18 Taiwan institution strains and had more symptomatic cases. Cluster C mainly contained Japanese trains and most were symptomatic. The reliability of this phylogenetic tree built by UPGMA method was evaluated by cophenetic correlation coefficient. The values higher than 75% on each node indicated this branch was reliable. To evaluate the reliability of grouping of

Cluster A, B and C, jackknife method was used to random sample *E. histolytica* strains from each cluster to compare strains from same Cluster to those strains from different clusters. As the result shown in Table 5, strains from Cluster A showed 85.5% similarity to themselves, Cluster B showed 98.4% and Cluster C showed 94.4%. This result indicated that the clusters of the phylogenetic tree were reliable. Although Cluster A mainly contained Indonesian strains, one sub-cluster contained 5 *E. histolytica* genotype J18 strains was also revealed. These strains came from 4 Taiwanese patients (2 MSM with HIV positive, 1 MSM and 1 without description) and 1 Japanese patient. Two patients showed no symptom, one had diarrhea, one had ALA and one had both diarrhea and ALA. Interestingly, another sub-cluster from Cluster B contained 7 genotype J20 strains from 6 Taiwanese patients (3 MSM with HIV positive, 1 MSM, 1 HIV positive and 1 without description) and 1 Japanese patient. Three patients showed no symptom, two had diarrhea and two had ALA. This result indicated that there were at least two different genotypes independently transmitted in two different HIV/MSM groups in 2009.

Analysis of the AIG1 genes in *E. histolytica* strains with different clinical symptom

Recently a group of AIG1 protein genes was identified by Drs. Kumiko Tsukui and Tomoyoshi Nozaki in NIID through the microarray comparison of *E. histolytica* strains from symptomatic and asymptomatic patients. AIG1 protein genes can be detected from symptomatic strains (like HM1:IMSS) and is absent in asymptomatic strains (like KU27). We therefore collaborated with Dr. Tsukui and begun to detect the AIG1 genes to symptomatic and asymptomatic strains from both Taiwanese and alien workers (Figure 4). However, we could only detect one AIG1 gene positive out of 14 *E. histolytica* strains from Cluster A, 6 out of 21 strains from Cluster B and 2 out of 4 strains from Cluster C. Whether the presence of AIG1 gene is related to

geographic origins or to disease outcomes need to be further investigated.

Clinical isolations

For detail genotyping and study of the strain variations, we need to enlarge the mount of clinical samples to complete the experiments. Culture of *E. histolytica* can be performed from fecal specimens, rectal biopsy specimens, or liver abscess aspirates. We have success to collect the three clinical isolations of *E. histolytica*, two from fecal samples (no.1245 amd 1249) and the other one (no.1291) is from the pus obtained from amebic liver abscess (ALA). One fecal sample (no.1245) is from the local person with diarrhea in mental handicap or retardation organization and the other (no.1249) is from an asymptomatic foreign laborer.

Discussions

Most *E. histolytica* strains (90%) of alien workers were from Indonesian patients. Many Indonesian strains caused asymptomatic feature were located in Cluster A but lees in Cluster B which had more Taiwanese strains. This may indicate that Indonesian strains are evolutionally different to Taiwanese and Japanese strains. However, interestingly almost Indonesian strains in either Cluster A or B were from asymptomatic patients and contained conserved locus type 6SQ in S-Q locus but in another N-K2 locus existed various different locus types. What the S-Q locus and N-K2 locus have different evolution rate need to be further study. However, N-K2 locus have seems to have higher evolution rate. One possible explanation for the N-K2 locus diversity is that Indonesian workers came from different islands in Indonesia and N-K2 locus is indepandent evolution in each island. To investigate these findings, the distribution of Indonesian workers in Indonesia may need to be further located.

On the other hand, so far there are no identical genotypes simultaneously exist in Taiwanese and Indonesia and indicated there is no obvious transmission between them, although two close strains ID98039 and TW098005 was seen in Cluster A. Taiwan had imported Indonesian worker for more than 10 years. Is there any genotype change during these 10 years shall be further investigated. It would be interested to compare the molecular epidemiology between Taiwanese and Indonesian strains followed the years. To further study the transmission of Indonesian strains, more epidemiological data should be collected from Indonesia to compare the data from Taiwan.

In opposition to Indonesian strains, several Japanese strains were identical to Taiwanese strains. The Japanese genotype J18 strain (JP_ALA-9) from an ALA patient was identical to 4 Taiwanese strains (TW098008, TW098018, TW098019,

TW098022) to form a sub-cluster in Cluster A. Because this sub-cluster contained 3 MSM patients, it would be very interested to know whether the JP_ALA-9 strain was come also from MSM/HIV patient, and whether there are certain *E. histolytica* strains are transmitted between Taiwanese and Japanese MSM/HIV groups.

A similar result was also seen in another sub-cluster in Cluster B, Japanese genotype J20 strain (JP_ALA-7) was identical to 6 Taiwanese strains (TW098003, TW098007, TW098013, TW098014, TW098016, TW098025). This result indicated that there were at least two different genotypes independently transmitted in two different HIV/MSM groups in 2009. In Cluster B, four Japanese genotype J20 strains (JP_ALA-3, JP_ALA-10, JP_KU3, JP_KU5) were identical to 3 Taiwanese strains (TW0981263, TW098128, TW098013, TW098129) that were isolated from the patients in same institution. Therefore, the transmission of amebiasis between Taiwanese and Japanese shall be further investigated.

Very few Taiwanese and Indonesian strains were grouped into Cluster C that mainly contained Japanese strains with severe disease manifestations. Since most Japanese strains came from patients with clinical symptoms, this may cause a grouping bias in phylogenetic analysis. Those strains were grouped into Cluster C. The AIG1 analysis also indicated that many *E. histolytica* strains (such as JP_KU48, JP_KU50, TW098006 and ID098049) in Cluster C were containing AIG1 gene and could present clinical manifestations. Whereas most *E. histolytica* strains in Cluster A did not have AIG1 gene and present asymptomatic manifestations. We found a Taiwanese HI/MSM group with same genotype to JP_ALA-9 in Cluster A is AIG1 deleted.

It will be very interested to know other Japanese strains (P_ALA-9, JP_ALA-17, JP_Ku 26, JP_Ku 14, JP_Ku 1) are also AIG1 deleted like AIG1 JP_Ku 27 in Cluster A. Another Taiwanese HI/MSM group with same genotype to JP_ALA-7 in Cluster B

is AIG1 positive. It will be very interested to know other Japanese strains (P_ALA-7, JP_ALA-12, JP_Ku 31 etc.) are also AIG1 positive in Cluster B. According the AIG1 results, the AIG1 family protein might relate to virulence especially those strains from Cluster C, but may not be the only virulent factor. It will be interesting to know the disease burden between HIV/MSM groups from Cluster A and B.

Acknowledgements

This work was supported by a grant (DOH100-DC-2028) from the Centers for Disease Control, Department of Health, Taiwan, and by a grant-in-aid of Ministry of Health, Labor and Welfare, the Government of Japan (H23-Shinkou- shitei-020).

References

1. TANYUKSEL, M., AND W. A. PETRI, JR. 2003. Laboratory diagnosis of amebiasis. Clinical Microbiology Reviews 16: 713–729.

2. WHO. 1997. Amoebiasis. Weekly Epidemiological Record 72: 97–100.

3. CHANG, S. Y., H. Y. SUN, D. D. JI, Y. C. LO, C. H.WU, P. Y.WU, W. C. LIU, C. C. HUNG, AND S. C. CHANG. 2008. Cost-effectiveness of detection of intestinal amebiasis by using serology and specificamebic-antigen assays among persons with or without human immunodeficiency virus infection. Journal of Clinical Microbiology 46: 3077–3079.

4. OHNISHI, K., Y. KATO, A. IMAMURA, M. FUKAYAMA, T. TSUNODA, Y. 5. SAKAUE, M. SAKAMOTO, AND H. SAGARA. 2004. Present characteristics of symptomatic *Entamoeba histolytica* infection in the big cities of Japan. Epidemiology and Infection 132: 57–60.

5. PARK, W. B., P. G. CHOE, J.H. JO, S.H.KIM, J.H. BANG, H. B.KIM, N. J. KIM, M. D.OH, AND K. W. CHOE. 2007. Amebic liver abscess in HIVinfected patients, Republic of Korea. Emerging Infectious Diseases 13: 516–517.

6. STARK, D., S. J. VAN HAL, G.MATTHEWS, J.HARKNESS, AND D. MARRIOTT. 2008. Invasive amebiasis in men who have sex with men, Australia. Emerging Infectious Diseases 14: 1141–1143.

7. CHAO, D. Y., P. H. WU, AND K. T. CHEN. 1997. The study of amebic infection in a provincial education and nursing institute. Epidemiology Bulletin 13: 135–144.

8. JIANG, D. S., AND K. H. CHANG. 2000. An investigation of amebiasis outbreak in one rehabilitation center for mentally retarded children. Taiwan Journal of Public Health 26: 261–270.

9. ONG, S. J., M. Y. CHENG, AND K. H. LIU. 1995. The status of parasites infection in a mental hospital—With emphasis on the amebiasis. Epidemiology Bulletin 11: 179–183.

10. HAMZAH, Z., S. PETMITR, M. MUNGTHIN, S. LEELAYOOVA, AND P. CHAVALITSHEWINKOON-PETMITR. 2006. Differential detection of *Entamoeba histolytica*, *Entamoeba dispar*, and *Entamoeba moshkovskii* by a single-round PCR assay. Journal of Clinical Microbiology 44: 3196–3200.

11. CALDERARO, A., C. GORRINI, S. BOMMEZZADRI, G. PICCOLO, G. DETTORI, AND C. CHEZZI. 2006. *Entamoeba histolytica* and *Entamoeba dispar*: Comparison of two PCR assays for diagnosis in a non-endemic setting. Transactions of the Royal Society of Tropical Medicine and Hygiene 100: 450–457.

12. Liang, S.Y., K.T. Hsia, Y.H. Chan, C.K. Fan, D. D.S. Jiang, O. Landt, and D.D. Ji. Evaluation of a new single tube multiprobe real-time PCR for the diagnosis of *Entamoeba histolytic* and *E. dispar*. J. Parasitol. 2010. 96(4):793-779.

13. SOLAYMANI-MOHAMMADI, S., M. REZAIAN, Z. BABAEI, A.RAJABPOUR, A. R. MEAMAR, A. A. POURBABAI, AND W. A. PETRI, JR. 2006. Comparison of a stool antigen detection kit and PCR for diagnosis of

Entamoeba histolytica and *Entamoeba dispar* infections in asymptomatic cyst passers in Iran. Journal of Clinical Microbiology 44: 2258–2261.

14. Gathiram, V., and T. F. H. G. Jackson. 1987. A longitudinal study of asymptomatic carriers of pathogenic zymodemes of *Entamoeba histolytica*. S. Afr. Med. J. 72:669–672.

15. Ali IKM, Zaki M, Clark CG. Use of PCR amplification of tRNA gene-linked short tandem repeats for genotyping Entamoeba histolytica. J Clin Microbiol 2005;43:5842–7.

16. Loftus B, Anderson I, Davies R, Alismark UCM, Samuelson J, Amedeo P, et al. The genome of the protist parasite *Entamoeba histolytica*. Nature 2005;433:865–8.

17. Clark CG, Cecilia U, Alsmark M, Hofer M, Saito-Nakano Y, Ali V, et al. Structure and content of the *Entamoeba histolytica* genome. Adv Parasitol 2007;65:51–190.

18. Clark CG, Ali IKM, Zaki MB, Loftus J, Hall N. Unique organization of tRNA genes in *Entamoeba histolytica*. Mol Biochem Parasitol 2006;146:24–9.

19. Cadiz, A. E., Kobayashi, S., Takeuchi, T., Tachibana, H., Nozaki, T. (2009). Identification of an avirulent *Entamoeba histolytica* strain with unique tRNA-linked

20. DIAMOND, L. S., D. R. HARLOW, AND C. A. CUNNICK. 1978. A new medium for axenic cultivation of *Entamoeba histolytica* and other *Entamoeba*. Transactions of the Royal Society of Tropical Medicine and Hygiene 72: 431–432. short tandem repeat markers. *Parasitol. Int.* 59: 75-81.

21. Robinson GL. The laboratory diagnosis of human parasitic amoebae. Trans R Soc Trop Med Hyg 1968;62:285–94.

22. Kobayashi S, Imai E, Haghighi A, Khalifa S, Tachibana H, Takeuchi T. Axenic cultivation of Entamoeba dispar in newly designed yeast extract-iron-gluconic acid-dihydroxyacetone-serum medium. J Parasitol 2005;91:1–4.

23. Suzuki J, Kobayashi S, Murata R, Tajima H, Hashizaki F, Yanagawa Y, et al. A survey of amoebic infections and differentiation of an *Entamoeba histolytica* -like variant (JSK2004) in non-human primates by a multiplex polymerase chain reaction. J Zoo Wildl Med 2008;39:370–9.

24. Tawari B, Ali IKM, Scott C, QuailMA, Berriman M,Hall N, et al. Patterns of evolution in the unique tRNA gene arrays of the genus Entamoeba.Mol Biol Evol 2008;25:187–98.

Publication list for this work

NIL

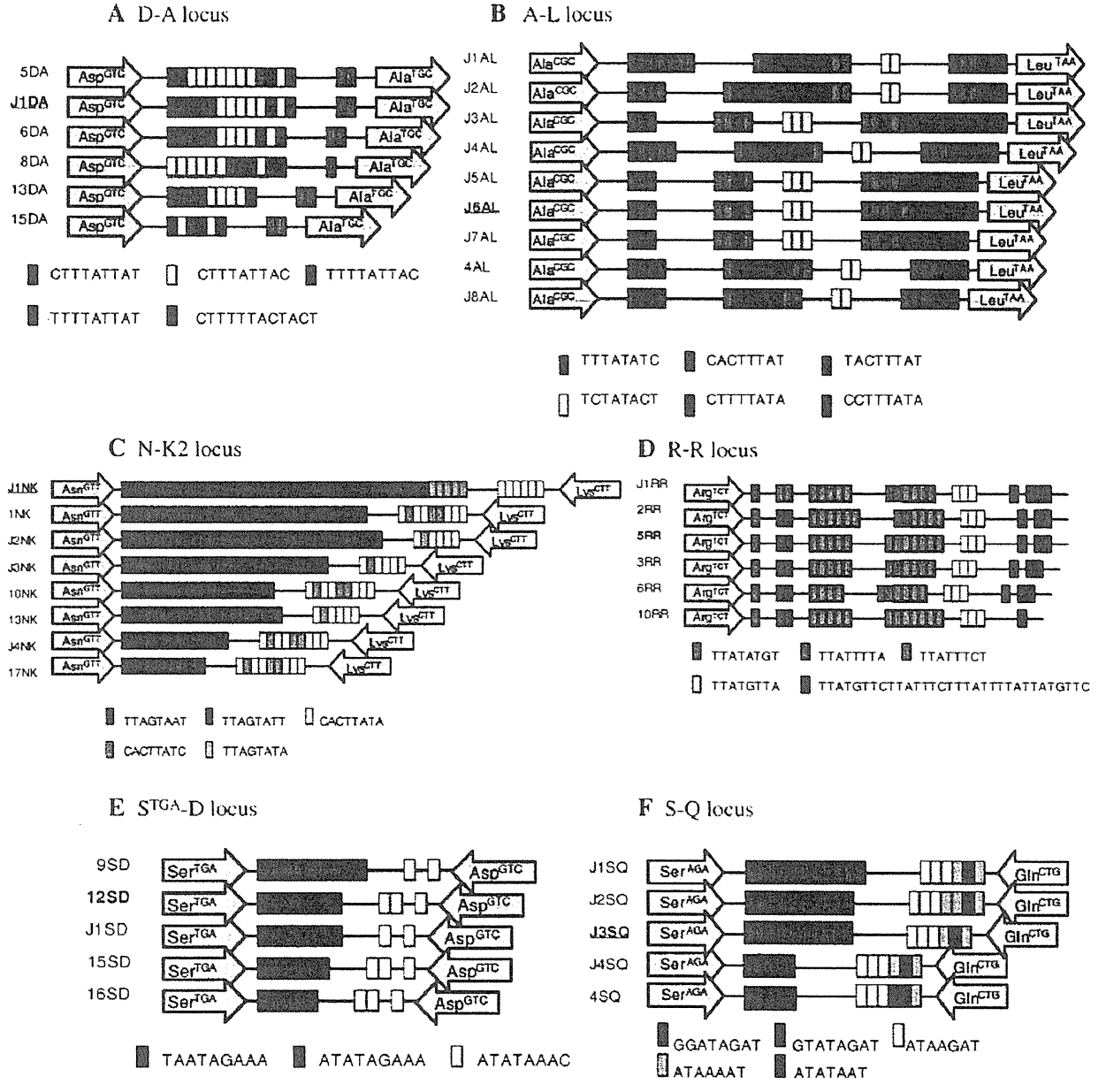

Figure 1. Schematic representation of STR types of each loci based on the nucleotide sequence of isolates found in this study. tRNA genes and STRs are depicted in arrows and rectangles, respectively, while non-tRNA, non-STR regions are shown in lines. The schematic diagrams of the D-A, A-L and S-Q loci were personally communicated by Dr. C. Graham Clark and adapted from Cadiz, A. E. *et al.* *Parasitol. Int.* 59: 75-81. 2009.

Table 1. Genotypes of Taiwanese *E. histolytica* samples using tRNA-linked STR markers.

Patient	S[tRNA]-D	S-Q	D-A	A-L	R-R	N-K2	Genotype	Symptom	Risk group	Sex	Age
TW098001	15SD	4SQ	6DA	[redacted]	2RR	[redacted]	EH1	Asymptomatic	HIV	M	27
TW098003	12SD	J2SQ	8DA	7AL	5RR	10NK	[redacted]	Asymptomatic	HIV	M	26
TW098004	15SD	4SQ	6DA	[redacted]	2RR	[redacted]	EH2	Diarrhea		[redacted]	41
TW098005	12SD	4SQ	6DA	10AL	3RR	10NK	EH3	Asymptomatic		[redacted]	61
TW098006	9SD	J1SQ	15DA	7AL	5RR	10NK	EH4	other	HIV	M	49
TW098007	12SD	J2SQ	8DA	7AL	5RR	10NK	[redacted]	Diarrhea		M	45
TW098008	15SD	4SQ	6DA	8AL	7RR	J4NK	[redacted]	Diarrhea	HIV	M	26
TW098009	15SD	4SQ	6DA	[redacted]	5RR	10NK	[redacted]	Diarrhea		M	34
TW098010	10SD	J3SQ	[redacted]	[redacted]	6RR	10NK	EH6	Diarrhea		[redacted]	76
TW098011	12SD	4SQ	5DA	4AL	6RR	1NK	EH7	Asymptomatic	HIV	M	27
TW098013	12SD	J2SQ	8DA	7AL	5RR	10NK	[redacted]	Asymptomatic	HIV	M	23
TW098014	12SD	J2SQ	8DA	7AL	5RR	10NK	[redacted]	Diarrhea	HIV	M	29
TW098015	15SD	4SQ	5DA	10AL	5RR	[redacted]	EH8	other		[redacted]	25
TW098016	12SD	J2SQ	8DA	7AL	5RR	10NK	[redacted]	Asymptomatic		M	37
TW098017	15SD	6SQ	6DA	[redacted]	5RR	[redacted]	[redacted]	Asymptomatic		[redacted]	82
TW098018	15SD	4SQ	6DA	8AL	7RR	J4NK	[redacted]	Asymptomatic	HIV	M	33
TW098019	15SD	4SQ	6DA	8AL	7RR	J4NK	[redacted]	Asymptomatic		M	25
TW098021	15SD	4SQ	3DA	8AL	7RR	J4NK	EH10	[redacted]		M	28
TW098022	15SD	4SQ	6DA	8AL	7RR	J4NK	[redacted]	D / A		M	51
TW098023	15SD	6SQ	6DA	[redacted]	5RR	[redacted]	[redacted]	Diarrhea		M	33
TW098024	12SD	4SQ	5DA	10AL	6RR	1NK	EH11	D / A	HIV	M	34
TW098025	12SD	J2SQ	8DA	7AL	5RR	10NK	[redacted]	[redacted]	HIV	M	31
TW098103	15SD	4SQ	6DA	[redacted]	8RR	13NK	[redacted]	Asymptomatic	Psychopathy	M	54
TW098104	15SD	4SQ	6DA	[redacted]	5RR	17NK	EH13	Asymptomatic	Psychopathy	M	[redacted]
TW098105	15SD	4SQ	6DA	[redacted]	5RR	[redacted]	EH14	Asymptomatic	Psychopathy	M	[redacted]
TW098106	15SD	4SQ	13DA	12AL	5RR	17NK	EH15	Asymptomatic	Psychopathy	M	[redacted]
TW098107	15SD	4SQ	6DA	10AL	5RR	[redacted]	EH16	Asymptomatic	Psychopathy	M	56
TW098108	15SD	4SQ	6DA	[redacted]	8RR	17NK	EH17	Asymptomatic	Psychopathy	M	64
TW098109	15SD	4SQ	6DA	[redacted]	5RR	[redacted]	EH18	Asymptomatic	Psychopathy	[redacted]	57
TW098110	15SD	4SQ	6DA	[redacted]	8RR	13NK	[redacted]	Asymptomatic	Psychopathy	[redacted]	60
TW098111	15SD	4SQ	6DA	[redacted]	8RR	13NK	[redacted]	Asymptomatic	Psychopathy	[redacted]	69
TW098113	15SD	4SQ	6DA	[redacted]	8RR	13NK	[redacted]	Asymptomatic	Psychopathy	[redacted]	54
TW098114	9SD	4SQ	9DA	[redacted]	8RR	17NK	EH19	Asymptomatic	Psychopathy	[redacted]	70
TW098116	15SD	4SQ	6DA	[redacted]	8RR	[redacted]	EH20	Asymptomatic	Psychopathy	M	58
TW098117	15SD	4SQ	6DA	[redacted]	8RR	13NK	[redacted]	Asymptomatic	Psychopathy	M	72
TW098118	15SD	4SQ	6DA	[redacted]	8RR	13NK	[redacted]	Asymptomatic	Psychopathy	M	51
TW098119	15SD	4SQ	6DA	[redacted]	8RR	13NK	[redacted]	Asymptomatic	Psychopathy	M	78
TW098122	[redacted]	4SQ	6DA	5AL	5RR	[redacted]	EH21	Asymptomatic	Psychopathy	M	72
TW098124	[redacted]	4SQ	6DA	5AL	7RR	[redacted]	EH22	Asymptomatic	Psychopathy	M	[redacted]
TW098125	15SD	6SQ	5DA	4AL	5RR	14NK	EH23	other	Psychopathy	M	[redacted]
TW098126	15SD	4SQ	5DA	4AL	10RR	10NK	[redacted]	Diarrhea	Psychopathy	M	[redacted]
TW098128	15SD	4SQ	5DA	4AL	10RR	10NK	[redacted]	Asymptomatic	Psychopathy	M	64
TW098129	15SD	4SQ	5DA	4AL	10RR	10NK	[redacted]	Asymptomatic	Psychopathy	M	[redacted]
TW098130	15SD	6SQ	3DA	4AL	10RR	10NK	[redacted]	Asymptomatic	Psychopathy	M	51
TW098131	15SD	6SQ	3DA	4AL	10RR	10NK	[redacted]	Asymptomatic	Psychopathy	M	[redacted]
TW098133	15SD	4SQ	9DA	4AL	10RR	10NK	EH25	Asymptomatic	Psychopathy	M	79

Taiwanese-Native
Taiwanese-Institution
New locus type
Genotype Genotype identical
M Male
Female
Age Median age
Age Age higher than median
Age Age less than median

Table 2. Genotypes of alien worker *E. histolytica* samples using tRNA-linked STR markers.

Patient	S[TM1]-D	S-Q	D-A	A-L	R-R	N-K2	Genotype	Symptom	Risk group	Sex	Age
ID098001	12SD	6SQ	3DA	8AL	5RR	T13NK	EH26	Asymptomatic	Foreign labor	M	25
ID098002	16SD	6SQ	15DA	5AL	5RR	T3NK	EH27	Asymptomatic	Foreign labor		23
ID098004	15SD	6SQ	6DA	11AL	5RR	T15NK	EH28	Asymptomatic	Foreign labor		33
ID098006	12SD	4SQ	5DA	T6AL	2RR	T1NK	EH29	Asymptomatic	Foreign labor		26
ID098007	15SD	6SQ	5DA	10AL	8RR	T4NK	EH30	Asymptomatic	Foreign labor		37
ID098008	15SD	4SQ	3DA	11AL	2RR	9NK	EH31	Asymptomatic	Foreign labor		33
ID098009	15SD	4SQ	5DA	10AL	5RR	T1NK	EH32	Asymptomatic	Foreign labor		22
ID098010	16SD	6SQ	15DA	4AL	6RR	17NK	EH33	Asymptomatic	Foreign labor		32
ID098013	15SD	6SQ	T4DA	T3AL	5RR	17NK	EH34	Asymptomatic	Foreign labor		27
ID098014	8SD	6SQ	11DA	1AL	10RR	T9NK	EH35	Asymptomatic	Foreign labor	M	28
ID098016	15SD	6SQ	3DA	5AL	5RR	T1NK	EH36	Asymptomatic	Foreign labor		22
ID098017	15SD	6SQ	3DA	10AL	8RR	T4NK	EH37	other	Foreign labor		31
ID098018	15SD	6SQ	3DA	T2AL	2RR	T2NK	EH38	Asymptomatic	Foreign labor		32
ID098019	T2SD	J3SQ	3DA	5AL	5RR	10NK	EH39	Asymptomatic	Foreign labor		34
ID098020	12SD	6SQ	3DA	11AL	3RR	T2NK	EH40	Asymptomatic	Foreign labor		28
ID098021	16SD	6SQ	3DA	11AL	2RR	T3NK	EH41	Asymptomatic	Foreign labor		30
ID098022	12SD	6SQ	3DA	11AL	5RR	10NK	EH42	Asymptomatic	Foreign labor		30
ID098023	15SD	J2SQ	9DA	10AL	5RR	T2NK	EH43	Asymptomatic	Foreign labor		21
ID098025	16SD	6SQ	9DA	5AL	5RR	17NK	EH44	Asymptomatic	Foreign labor		29
ID098026	16SD	6SQ	3DA	5AL	5RR	10NK	EH45	Asymptomatic	Foreign labor		33
ID098027	16SD	6SQ	T5DA	4AL	6RR	17NK	EH46	Asymptomatic	Foreign labor		31
ID098028	5SD	6SQ	3DA	10AL	5RR	T14NK	EH47	Asymptomatic	Foreign labor		37
ID098029	15SD	6SQ	3DA	5AL	5RR	T25NK	EH48	Asymptomatic	Foreign labor		34
ID098030	15SD	4SQ	9DA	5AL	5RR	12NK	EH49	Asymptomatic	Foreign labor		35
ID098031	15SD	J3SQ	3DA	11AL	5RR	T1NK	EH50	Asymptomatic	Foreign labor		26
ID098032	15SD	4SQ	6DA	T2AL	2RR	T2NK	EH51	Asymptomatic	Foreign labor		37
ID098033	15SD	6SQ	5DA	T2AL	8RR	T3NK	EH52	Asymptomatic	Foreign labor		24
ID098034	15SD	4SQ	3DA	11AL	5RR	T25NK	EH53	Asymptomatic	Foreign labor	M	27
ID098037	15SD	6SQ	T4DA	5AL	5RR	17NK	EH54	Asymptomatic	Foreign labor		24
ID098038	16SD	6SQ	15DA	4AL	6RR	T26NK	EH55	Asymptomatic	Foreign labor		33
ID098039	12SD	4SQ	6DA	8AL	3RR	T27NK	EH56	Asymptomatic	Foreign labor		28
ID098041	15SD	4SQ	T4DA	T6AL	3RR	17NK	EH57	Asymptomatic	Foreign labor		28
ID098042	4SD	6SQ	3DA	5AL	5RR	T2NK	EH58	Asymptomatic	Foreign labor		26
ID098043	15SD	6SQ	6DA	10AL	5RR	T3NK	EH59	Asymptomatic	Foreign labor		35
ID098044	12SD	6SQ	T3DA	12AL	5RR	T3NK	EH60	Asymptomatic	Foreign labor		31
ID098046	15SD	6SQ	3DA	4AL	5RR	T32NK	EH61	Asymptomatic	Foreign labor		37
ID098048	15SD	4SQ	3DA	5AL	5RR	T39NK	EH62	Asymptomatic	Foreign labor		33
ID098049	1SD	6SQ	5DA	T2AL	8RR	T35NK	EH63	other	Foreign labor		31
ID098050	16SD	6SQ	15DA	5AL	8RR	T3NK	EH64	Asymptomatic	Foreign labor		33
ID098051	12SD	4SQ	3DA	5AL	3RR	T29NK	EH65	Asymptomatic	Foreign labor	M	25
ID098052	15SD	6SQ	3DA	T9AL	5RR	T3NK	EH66	Asymptomatic	Foreign labor	M	26
PH098001	15SD	4SQ	6DA	T4AL	5RR	10NK	EH5	Asymptomatic	Foreign labor		33
PH098002	15SD	4SQ	6DA	T4AL	5RR	10NK	EH5	Asymptomatic	Foreign labor	M	40
PH098003	15SD	6SQ	6DA	T4AL	7RR	10NK	EH67	other	Foreign labor		34
PH098004	15SD	4SQ	6DA	T4AL	5RR	10NK	EH5	Asymptomatic	Foreign labor		38
PH098005	15SD	4SQ	6DA	10AL	5RR	10NK	EH68	Asymptomatic	Foreign labor	M	24
VN098001	15SD	3SQ	T2DA	10AL	5RR	T16NK	EH69	Asymptomatic	Foreign labor		29

Indonesian
Filipino
Vietnamese

New locus type

Genotype Genotype identical

M Male
Female

Age Median age
Age Age higher than median
Age Age less than median

Table 3. Comparison of the diversity of the locus types in each STR Loci.

		Diversity (type/sample number)					
Origin of samples		Japan, Thailand, Bangladesh	Primarily Bangladesh	Japan	Taiwan, Indonesia, Philipoines, Vietnam	Taiwan [a]	Indonesia [a]
Isolate date		1987-2002	N/A	1985-2006	2009	2009	2009
tRNA-linked STR loci	S^{TGA}-D		17/128 (0.133)	5/37 (0.135)	11/93 (0.118)	5/46 (0.109)	8/41 (0.195)
	S-Q			5/37 (0.135)	6/93 (0.065)	5/46 (0.109)	4/41 (0.096)
	D-A	13/79 (0.165)		6/37 (0.162)	13/93 (0.140)	8/46 (0.174)	9/41 (0.220)
	A-L			9/37 (0.243)	17/93 (0.183)	12/46 (0.261)	11/41 (0.268)
	R-R		12/136 (0.088)	6/37 (0.162)	7/93 (0.075)	7/46 (0.152)	6/41 (0.146)
	*N-K2		18/ 53 (0.340)	8/37 (0.216)	40/93 (0.430)	14/46 (0.304)	29/41 (0.707)
	genotype			23/37 (0.622)	73/93 (0.785)	28/46 (0.609)	41/41 (1.000)
Reference		Haghighi et al. (2003)	Tawari et al. (2008)	Cadiz et al. (2010)			

Table 4. The pairs of 9-17 bp conserved sequences used to locate the variable regions from each STR sequence.

Locus	Spacer reigon			
S^{TGA}-D	Block A spacer	Block B spacer		
Begin	ACATAGAAA	TAGTATAAT		
End	TAATAAAAA	ATATAAAAG		
S-Q	Block A spacer	Block B spacer		
Begin	TTAACATAA	ATAAAGAGA		
End	GGATATAAA	AGAAAGAAT		
D-A	Block A spacer	Block B spacer		
Begin	TTACTCCCTAT	ATACGTACT		
End	ATCTATTCT	CTTCTTACT		
A-L	Block A spacer	Block B spacer	Block C spacer	Block D spacer
Begin	TTCTACTTA	TTATCTTAT	CTATATACT	TACTTTGAC
End	TATATCCCT	TATCTTTAC	ATATGTGTG	TTGTATCTT
R-R	Block A spacer	Block B spacer	Block C spacer	
Begin	ACTATATGT	TTATGTTC	TTGTTTTTTTTATGTTA	
End	CTATGTATG	TTATCCTATTATG	TTATGCTATTATTTC	
N-K2	Block A spacer	Block B spacer		
Begin	TATACCTCC	CTTTATATC		
End	TTACTGTCT	CTCTTCCCC		

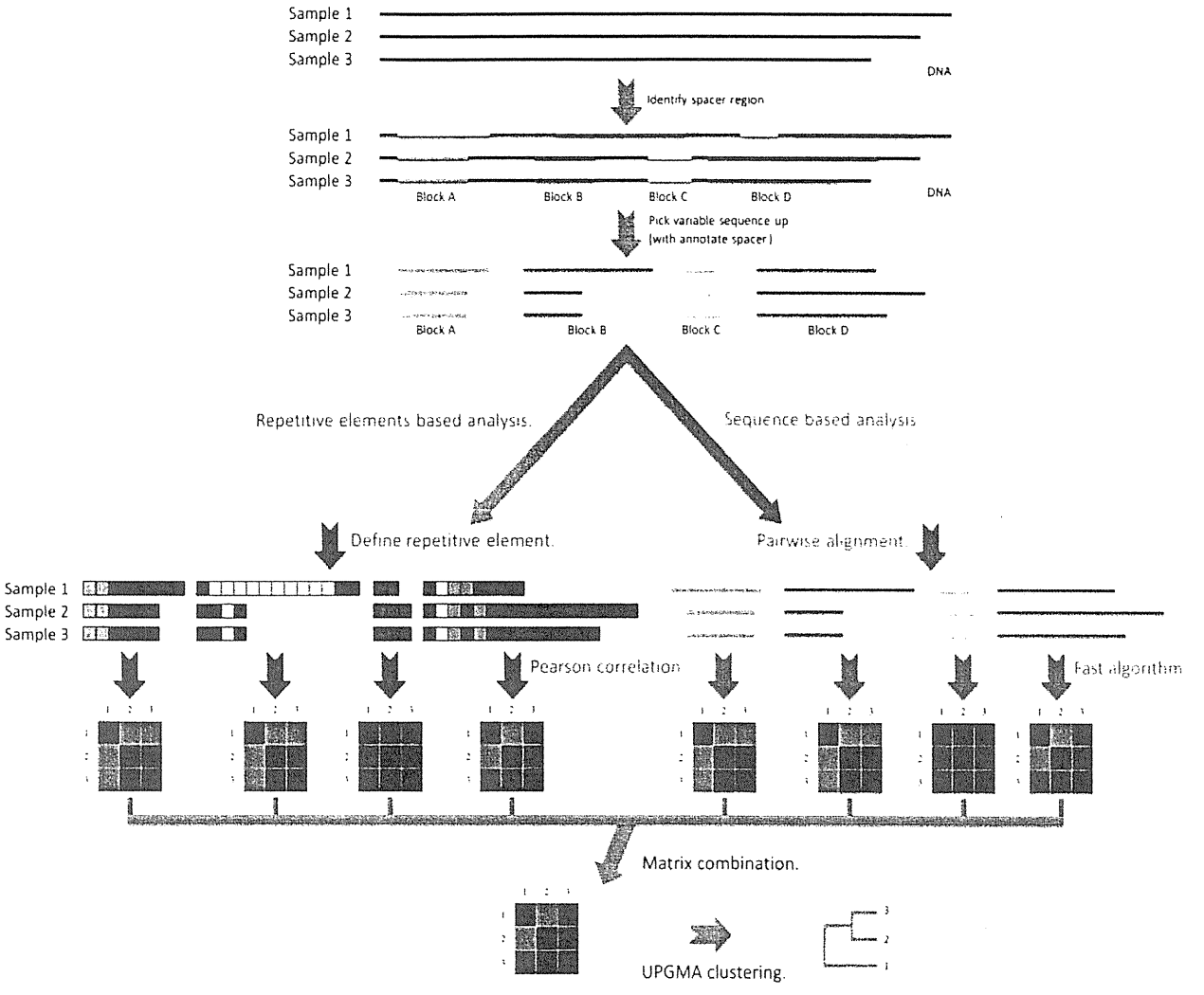

Figure 2. The flowchart of the phylogenetic development based on the six tRNA-linked STR loci.