UNIVERSITY OF SZEGED

Examination of somatic and Y chromosomal STR markers in human populations

Summary of PhD thesis

Written by
Zsuzsanna Beer

Supervisor:
Prof. Dr. Tibor Varga

Performed at:
University of Szeged
Szentgyörgyi Albert Medical Faculty
Department of Forensic Medicine
2004

1. Introduction

The early home of ancient Hungarians was in West Siberia. In the II. millennia before Christ ancient Hungarians started to move towards Southwest direction. During migration they met other people (Mongolians, Turkish people) and mixed with them. After centuries of migration they settled down in the Karpathian-basin in the IX. century after Christ. It is generally admitted that Szekely people were migrated together with ancient Hungarians and they settled down together with them. The hypothesis of double conquest of Hungary says that Szekely people arrived at the Karpathian-basin some hundred years before the Hungarians. Nowadays Szekelys are known as a Hungarian people but their customs and subculture makes them a little bit different from other Hungarian populations. Szekelys think that they are originated not from the people of Arpad but from people of Attila.

The origin of different populations has been the field of intense research for more decades. The examinations of different genetic markers have already given some results. Markers which has high mutation rate but are not impressed by natural selection so they stay with their changes in the genome are useful for examining the evolution in short period of time. These markers are the hypervariable region of mitochondrial DNA (mtDNA) and short tandem repeats (STR) of nuclear DNA. In STRs the 2-6 bp long repeat units are present in 6-40 tandem copies. Alleles of STR markers differ from each other in the number of repeat units. STRs are present on every chromosome in the non-coding region of inter- and intragenic sequences. In the practise of forensic genetics polymorphic markers are used which have several alleles and the distribution of alleles causes diversity in the population. Though somatic STR markers are usually under recombination they are useful for detection of personal identification. In forensic genetic practice it is necessary to establish the polymorphism of populations in somatic and Y chromosome STR markers. Y-STR markers are more useful for examination of relationships of different populations because they don't have recombination and they are linked. Mutation of STR systems causes changes in the number of repeat units. This speciality can be used for the calculation of genetic distance between populations.

The aim of my work was to determine the allele distribution of nine somatic (D3S1358, D5S818, D7S820, D8S1179, D13S317, D18S51, D21S11, FGA, VWA) and eight Y chromosome (DYS19, DYS385, DYS389/I,II, DYS390, DYS391, DYS392, DYS393) STR systems in Hungarian and Szekely populations and to establish the genetic distance between these populations. For the comparison it was necessary to use the data of other European populations. Earlier results showed that Hungarian population is genetically closer to Asian populations than other European populations we decided to examine the present Mongolian people.

The other important aim was for our result to be involved to the Y-STR Haplotype Reference Database where data of other populations from all over the world are reachable for everyone.

2. Materials and methods

Blood samples were collected from Hungarian population (100 males, 101 females) from Szeged and its environment; from isolated Szekely population (99 males) populated Coround in Transylvania and from Mongolians (56 males, 48 females) living in Hungary but originated from Mongolia. Blood samples were transferred to cotton swab for storing. DNA extraction was performed from the whole blood in case of Hungarian samples and from bloodstains in case of Szekely and Mongolian samples.

From whole blood samples DNA extraction was performed with the salting out method of Miller et al. From bloodstain samples DNA was extracted partly with Chelex 100 solution and partly with phenol-chloroform methodology. Concentration of DNA was measured by DyNA Quant 2000 Fluorometer (Hoefer).

DNA fragments of somatic STR markers were amplified in multiplex PCR reaction with fluorescently labelled primers using Profiler Plus (Applied Biosystems) and AmpFlSTR Identifiler (Applied Biosystems) kits. Y-STR markers were amplified either by singleplex or multiplex reactions. For multiplex PCR fluorescently labelled primers were used in kits of Y-PLEX^{TM}6 (Reliagene), DYSplex-1, DYSplex-2 (Serac) kits.

Allele separation was performed in polyacrylamid gel. Non-labelled fragments were separated with horizontal gel electrophoresis and fluorescently labelled fragments were separated with capillary gel electrophoresis (ABI Prism 310 and 3100 Genetic Analyzer). Pherograms were analysed by GeneScan software ver. 2.1..

For describing the examined populations, allele frequencies, power of discrimination (PD), power of exclusion (PE) and polymorphism informed concern (PIC) were determined by PowerStatsV12.xls program. Determination of population structures were performed with AMOVA analysis by ARLEQUIN software. Genetic distances between populations were calculated from pairwise F_{ST} values involving published data of German, Romanian and Japanese populations. A tree was constructed by Phylip program to illustrate the genetic relationship between populations.

Our haplotype data were compared to other European and Asian data in the Y-STR Haplotype Reference Database.

3. Results and discussion

Somatic markers

The result of examination of somatic STR markers shows that all the three examined populations are in Hardy-Weinberg equilibrium for all markers except one somatic marker in Székely population (D7S820) which could be due to isolation of populations.

In this work we determined the allele and genotype frequency distribution of nine somatic STR markers (D3S1358, D5S818, D7S820, D8S1179, D13S317, D18S51, D21S11, FGA, VWA) in Hungarian population living in Szeged and its environment and in Szekely population living in Korond and of 14 somatic STR markers (D3S1358, CSF1PO, D5S818, D7S820, D8S1179, D13S317, D18S51, D21S11, FGA, VWA, D16S539, TH01, TPOX, D2S1338) in cooperation with the DNA group of Louis Pasteur University in a Mongolian population working in Budapest.

The allele frequency values of somatic markers in Hungarian population are usually in all Europe. In paternity cases which belong to Szeged area in calculation of possibility it is proposed counting with these results.

The most frequent allele was different in four from the nine somatic STR markers between Hungarian and Szekely populations (D3S1358, VWA, D18S51, D7S820) and only 3 were the same in Hungarian-Mongolian (D8S1179, D13S317, VWA and Szekely-Mongolian comparison (D8S1179, D7S820, D3S317).

The observed genotypes of the examined populations were compared to each other on order to establish the genetic distance between them.

AMOVA analysis showed that although variance within populations was much higher than variance between populations there is significant difference between Hungarian-Szekely and Szekely-Mongolian populations. The value of F_{ST} were only 0,008 between Hungarian-Szekely populations like it was expected because earlier it has been showed that the genetic distance between European population by somatic STR markers are very law [Budowle 2000]. Differentiation between Hungarians and Mongolians is higher than between Mongolians and Szekelys by somatic STR markers.

Y chromosome markers

Allele- and haplotype distribution of eight Y chromosomal STR markers (DYS19, DYS385, DYS389/I, DYS389/II, DYS390, DYS391, DYS392, DYS393) were determined in Hungarian population living in Szeged and its environment and in Szekely population living in Korond. Allele- and haplotype distribution of 8 Y chromosomal STR markers (DYS19, DYS389/I, DYS389/II, DYS390, DYS391, DYS392, DYS393, YCAII) in the Mongolian population were determined in cooperation with the French DNA group.

The most frequent allele was different in two Y chromosome STR markers between Hungarian and Szekely populations. There is meaningful difference between these populations in DYS389/I and DYS389/II markers where the difference between the most frequent alleles were two repeat units. The most frequent allél of DYS391 and DYS392 were the same in these populations but in case of DYS391 the frequency of the allele 10 was 59,48% in the Hungarian population and more than 72% in the Szekely population. In case of DYS392 marker the frequency of allél 11 was 72% in Hungarians and only 40% in the Szekely population.
The most frequent allele of Y-STR markers in Mongolian population was different from Hungarian and Szekely population in 3 STR loci (Hungarian: DYS19, DYS389/I, DYS389/II; Szekely: DYS19, DYS389/I, DYS30). In case of DYS19 and DYS 389/I the difference was 2 repeat units.

Y-STR haplotype data of the examined Hungarian, Szekely and Mongolian populations were compared in order to determine the level of genetic relatedness and to establish the genetic distance between them. For the comparison data of other populations (German, Romanian and Japanese) were involved for control. The earlier examination of 3 Y chromosomal STR

markers showed that the population living in Szeged and its environment differs from other Caucasoid people [Csete 1999].
The analysis of molecular variance by Y-STR markers showed that there is significant difference between the examined populations. The variance between populations was very low (6,21% of the total variance) but the value was higher than in case of somatic markers which means that Y-STR markers can differentiate the populations more than somatic markers.
The F_{ST} values showed that Mongolians differ from every of the examined populations in the highest level, but the Hungarian and Szekely populations are the closest to them. Japanese and Mongolian populations are significantly different with a quite high Φ_{ST} value. This value could be due to the big difference between sample number in the populations. The genetic distance between Szekely, Hungarian, Romanian and German populations are close to zero. It is difficult to say clear conclusions from the results because the control populations may are mixed populations. For example the samples of Romanian populations were from Bukarest where big part of inhabitants has Hungarian origin. There is no data from isolated Romanian population.
The result of the comparison shows that Hungarian and Szekely population is genetically closer to Asian people than to German population like it was expected because earlier it was determined that there are some special forms of isoenzimes and STR markers in Hungarian population which are frequent in Asian people [Lahermo 1999].

The results of the comparison in Y-STR database shows that the haplotypes of the examined populations are different from the typical European haplotypes. The haplotypes were frequent in German and Polish populations due to the fact that large numbers of samples of these populations are present in the database. Also Baltic and Turkish populations had the same haplotypes like the examined Hungarian and Szekely populations. One of the haplotypes found in Szekely population is frequent in Finnish population also if earlier it was said that there is no genetic relationship between Finnish and Hungarian people [Lahermo 1999].

The results of our examination indicate that although polymorphic TR systems can make the populations being differentiated because of their structure and speciality of their inheritance and mutation, they are useful in analysis of short period evolution.

Description of Szekely population by eight Y chromosome STR markers and Mongolian population by somatic and Y-STR markers have already published.

After our laboratory successfully made quality control for Y chromosome STR markers the haplotypes of Hungarian and Szekely populations have been deposited in the Y-STR Haplotype Reference Database.

List of congresses related to the subject of the thesis

Oral presentations

Beer Zs., Csete K., Varga T.
Comparison of two isolated "Hungarian" populations to Szegedian (mixed) population by somatic- and Y-chromosomes
Allpe-Adria-Pannonia Congress, Croatia, Opatija, 2001. június 23-26

Beer Zs., Pénzes Zs., Farkas-Hegyi., Csete K., Varga T.
Analysis of nine STR loci in an isolated Szekely population
Alpe-Adria-Pannonia Congress, Visegrád, 2002. május 3-5.

Varga T., Keyser C., **Beer Zs**., Pénzes Zs., Pamzsav H., Csete K., Ludes B
STR data analyses in Mongolian population
Alpe-Adria-Pannonia Congress, Visegrád, 2002. május 3-5.

T. Varga, C. Keyser, **Zs. Beer** , Zsolt Penzes, H. Pamzsav, K. Csete, B. Ludes
STR data analysis in Mongolian population
Fifth International Symposium on Advances in Legal Medicine, 2002. Okt. 1-4 Takayama, Japán.

K. Csete, **Zs. Beer**, T. Varga
DNA Untersuchungen in Ungarischen Paternitätsfälle
Alpe-Adria-Pannonia Congress, Visegrád, 2002. május 3-5.

C. Keyser-Tracqui, D. Montagnon, P. Blandin, **Z. Beer**, H. Pamjav, E. Crubézy and B. Ludes
A comparison of Y-chromosomal and autosomal STR data in past and present-day Mongolian populations
III. International Forensic Y-User Workshop, 2002. november 7-9.
Porto, Portugalia

K. Csete, **Zs. Beer**, T. Varga
Prenatal and newborn paternity testing with DNA analysis
Alpe-Adria-Pannonia Congress, Rogaska Slatina, Slovenia, 2003

Zs. Beer, E. Frakas-Hegyi, K. Csete, T. Varga
Y-chromosome STR analysis in a Szekely population
Alpe-Adria-Pannonia Congress, Rogaska Slatina, Slovenia, 2003

Posters:
Beer Zs., Pénzes Zs., Csete K., Varga T.
Comparison of two isolated „Hungarian" populations to Szegedian (mixed) population by Y-chromosomes
ISFG Congress, Germany, Muenster, 2001. aug. 28-szeptember 1.

Beer Zs., Csete K., Varga T.
Examination of the genetic distance between a mixed Hungarian and two isolated Szekely population by Y-chromosome STR systems
The second European-American training course in Forensic Genetics Croatia, Dubrovnik, 2001. szept. 3-14

List of publications related to the subject of the thesis

Beer Zs., Csete K., Varga T.
Comparison of two isolated „Hungarian" populations to Szegedian (mixed) population by somatic- and Y-chromosomes
Croatian Medical Association 2001, 115-119

Beer Zs., Pénzes Zs., Farkas-Hegyi., Csete K., Varga T.
Analysis of nine STR loci in an isolated Szekely population
11th International meeting on Forensic Medicine Alpe-Adria-Pannonia Proceedings, 2002, 17-20

Varga T., Keyser C., **Beer Zs**., Pénzes Zs., Pamzsav H., Csete K., Ludes B
STR data analyses in Mongolian population
11th International meeting on Forensic Medicine Alpe-Adria-Pannonia Proceedings, 2002, 21-24

Beer Zs., Pénzes Zs., Csete K., Varga T.
Comparison of two isolated „Hungarian" populations to Szegedian (mixed) population by Y-chromosomes
International Congress Series (2003): 1239:473-480

Beer Zs., Pénzes Zs., Farkas-Hegyi É., Csete K., Varga T
Population genetics of nine STR loci – D3S1358, VWA, FGA, D8S1179, D21S11, D18S51, D5S818, D13S317 AND D7S820 – in a Szekely population from Transsylvania
Forensic Science International (submitted for publication)

Zs. Beer, K. Csete, T. Varga
Y-chromosomal STR haplotype in Szekely population
Forensic Science International (2004) 139(2-3):155-158

T. Varga, C. Keyser, **Zs. Beer** , Zs. Penzes, H. Pamzsav, K. Csete, B. Ludes
STR data analysis in Mongolian population
Legal Medicine 5 (2003) S156-159

Zs. Pénzes, Gy. Csanádi, G. M. Kovács, **Zs. Beer**
Molecular markers in ecology
Tiscia 33:9-30 (2002))